N-subjettiness “tagger” with grooming

Comparisons of performance with various grooming tools: Pruning, trimming and filtering

David W. Miller

Enrico Fermi Institute

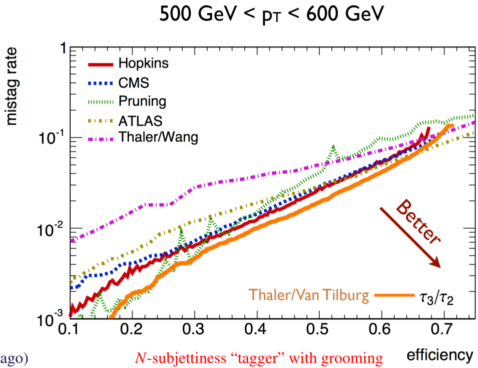
THE UNIVERSITY OF CHICAGO

July 27, 2011

N-subjettiness, τ_N

BOOST 2011 offered an excellent opportunity to present the fantastic new results on jet mass to the wider community. It also allowed for in depth discussion that highlighted what that community might most like to see measured and understood in the data.

The list is long, but it is clear that N-subjettiness is near the top of that list for one simple reason:



Grooming comparisons

Some initial indications suggest that some of the **grooming** procedures may affect the ability to utilize N-subjettiness as a top-tagger. The goal of these plots is to demonstrate what that affect may be with the BOOST 2010 benchmark samples, and then to see how much pile-up affects that result.

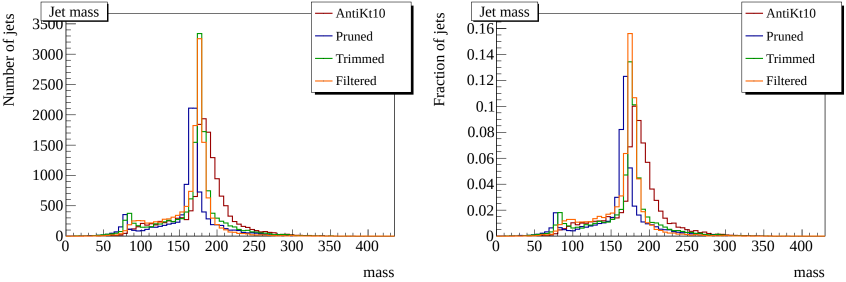
Configuration using `TopTag.py`:

- **Un-groomed jets:** anti-k_t, $R = 1.0$ jets
- **Pruned jets:** anti-k_t, $R = 1.0$ parent jets, pruned using C/A with $z_{\mathrm{cut}} = 0.05$ and $R_{\mathrm{cut}} = 0.1$
- **Filtered jets:** C/A, $R = 1.0$ parent jets, filtered using $R_{\mathrm{filt}} = 0.35$ and $N_{\mathrm{subjets}} = 3$.
  - However, in the implementation in SPARTYJET, I do not see y_{cut} or μ (the symmetry and mass drop cuts)....
- **Trimmed jets:** anti-k_t, $R = 1.0$ parent jets trimmed using $R_{\mathrm{subjet}} = 0.35$ and $p_{T,j1}/p_{T,J} < 0.05$.

Mass distributions

$500 < p_{\mathrm{T}} < 600$ *GeV*

Using the $t\bar{t}$ and di-jet samples found on the TWiki, here, plot just the raw jet mass for each collection.

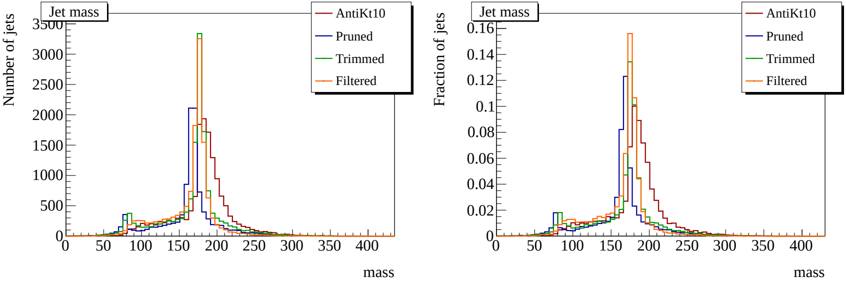


Unnormalized

Normalized

Mass distributions

$500 < p_{\mathrm{T}} < 600$ *GeV*



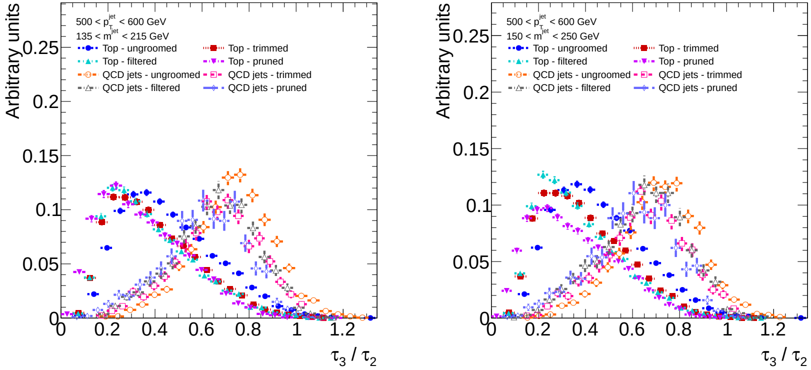
Unnormalized Normalized

Efficiency of the mass range preselection is fairly stable:

- $150 < m^{\mathrm{jet}} < 250$ GeV: 80% (ungroomed), 76-78% (groomed)
- $135 < m^{\mathrm{jet}} < 215$ GeV: 78% (ungroomed), 76-78% (groomed)

τ_3/τ_2 distributions

$500 < p_{\mathrm{T}} < 600$ GeV



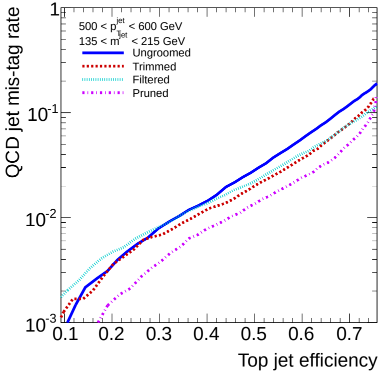
$135 < m^{\mathrm{jet}} < 215$ $150 < m^{\mathrm{jet}} < 250$

The relative positions of the peaks shift slightly, both with grooming and with the different mass windows applied to the parent jets.

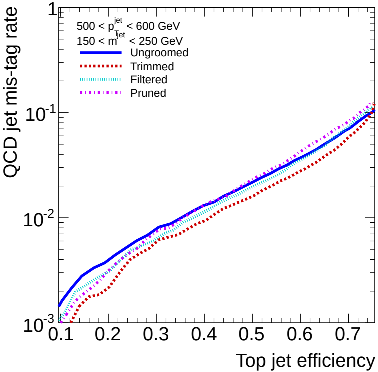
Tagger performance (I)

$500 < p_{\mathrm{T}} < 600$ *GeV*

Efficiency is now calculated on a per jet basis with respect to the total number of jets **before** the mass selection (earlier version was w.r.t. **after** the mass selection).



$135 < m^{\mathrm{jet}} < 215$ (all jets)

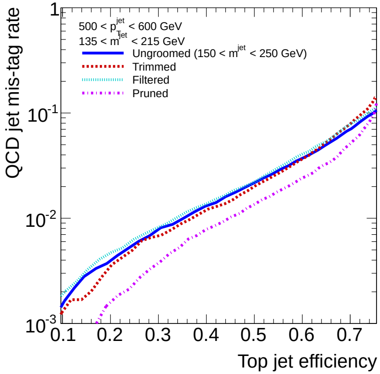


$150 < m^{\mathrm{jet}} < 250$ (all jets)

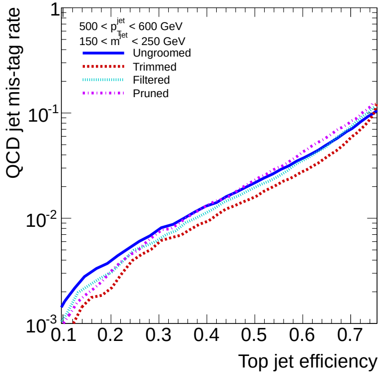
Achieve a very similar performance to Thaler/Van Tilburg for the full range.

Tagger performance (II)

$500 < p_{\mathrm{T}} < 600$ GeV



$135 < m^{\mathrm{jet}} < 215$ (groomed jets)

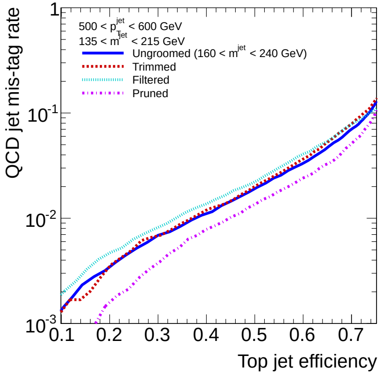


$150 < m^{\mathrm{jet}} < 250$ (all jets)

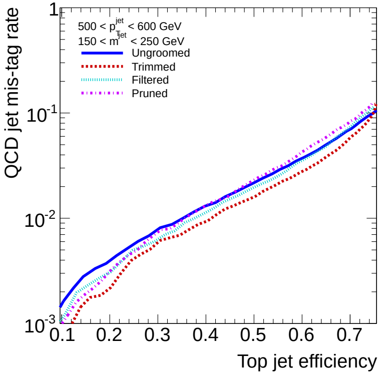
The smaller mass window is more appropriate for groomed jets, while the larger (and higher) window is better for the ungroomed jets.

Tagger performance (III)

$500 < p_{\rm T} < 600$ *GeV*



$135 < m^{\rm jet} < 215$ (groomed jets)

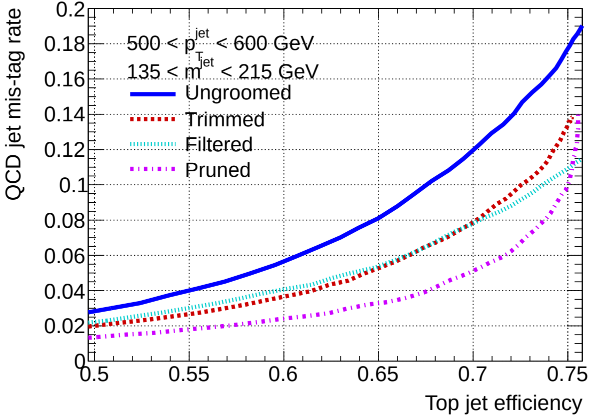


$150 < m^{\rm jet} < 250$ (all jets)

The mass window used by Jesse and Ken is even slightly better for the intermediate efficiency range ($\sim 50\%$).

Tagger performance (IV)

$500 < p_{\mathrm{T}} < 600$ *GeV*

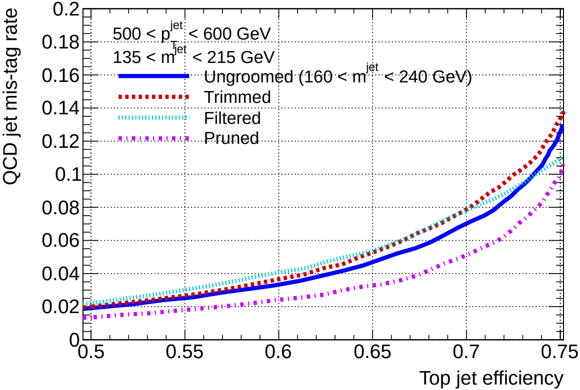


With pruning and the same mass window for both groomed and ungroomed jets:

- Reduction of up to 30-50% in mis-tag rate for the same efficiency.
- 5-12% absolute improvement in efficiency for the same mis-tag rate.

Tagger performance (V)

$500 < p_{\mathrm{T}} < 600$ *GeV, with optimized mass window for ungroomed jets*



With grooming and an optimized mass window for ungroomed jets:

- Improvement is more modest, and only achieved for the more “aggressive” pruning.
- 10-20% reduction in mis-tag rate for the same efficiency.
- 5% absolute improvement in efficiency for the same mis-tag rate.

Summary and next steps

So far...

- Have performed a first look at the impact of grooming on *N*-subjettiness
  - Non-optimal configuration for filtering.
- Somewhat significant relative improvement is observed for pruning, but details depend on the mass window around the top.
  - Implies some sensitivity to the detector-level mass resolution.
  - An optimized mass window for the ungroomed jets compared to that for groomed jets reduces the magnitude of the improvement.

Moving forward

- The real aim is to assess the impact of pile-up on *N*-subjettiness and then how the various grooming tools mitigate those affects.
- Can more aggressive filtering / trimming perform similar to pruning?
- How does grooming **after** computing τ_3/τ_2 perform?
- In the process of organizing pile-up samples with Stefan and Steffen. Hopefully we can mix them up soon!

Additional Material

N-subjettiness, τ_N

Measures the extent to which a given jet is likely to be composed of N subjets by first identifying a set of subjets and then comparing the energy flow in the jet to the direction of these subjets [1].

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min \{\Delta R_{1,k}, \Delta R_{2,k}, \cdots, \Delta R_{N,k}\} \tag{1}$$

$$d_0 = \sum_k p_{T,k} R \tag{2}$$

The sum runs over the k constituent particles in a given jet where $p_{T,k}$ are their transverse momenta, and $\Delta R_{j1,k} = \sqrt{\delta y_{j1,k}^2 + \delta \phi_{j1,k}^2}$ is the distance in $\Delta y \times \Delta \phi$ between a candidate subjet $j1$ and a constituent particle k.

→Uses ratios of momenta and exclusive subjets

→Sensitive to subjet multiplicity, not so much kinematics

Efficiency vs. mis-tag for ungroomed jets

```
     Bin #  |  Bin index |  Bin value |         Eff |     Mis-tag
-------------------------------------------------------------------
         0  |          0 |  -0.014000 |    0.000000 |    0.000000
         1  |          1 |   0.000000 |    0.000000 |    0.000000
         2  |          2 |   0.014000 |    0.000000 |    0.000000
         3  |          3 |   0.028000 |    0.000000 |    0.000000
         4  |          4 |   0.042000 |    0.000000 |    0.000000
         5  |          5 |   0.056000 |    0.000068 |    0.000000
         6  |          6 |   0.070000 |    0.000339 |    0.000000
         7  |          7 |   0.084000 |    0.000543 |    0.000077
         8  |          8 |   0.098000 |    0.001696 |    0.000077
         9  |          9 |   0.112000 |    0.003257 |    0.000077
        10  |         10 |   0.126000 |    0.005496 |    0.000232
        11  |         11 |   0.140000 |    0.009974 |    0.000232
        12  |         12 |   0.154000 |    0.018387 |    0.000232
        13  |         13 |   0.168000 |    0.027275 |    0.000310
        14  |         14 |   0.182000 |    0.039555 |    0.000310
        15  |         15 |   0.196000 |    0.052649 |    0.000620
        16  |         16 |   0.210000 |    0.067779 |    0.000775
        17  |         17 |   0.224000 |    0.085080 |    0.001007
        18  |         18 |   0.238000 |    0.102314 |    0.001394
        19  |         19 |   0.252000 |    0.122803 |    0.001782
        20  |         20 |   0.266000 |    0.143022 |    0.002324
        21  |         21 |   0.280000 |    0.167718 |    0.002789
        22  |         22 |   0.294000 |    0.190108 |    0.003176
        23  |         23 |   0.308000 |    0.211615 |    0.003796
        24  |         24 |   0.322000 |    0.231902 |    0.004415
        25  |         25 |   0.336000 |    0.255648 |    0.005190
        26  |         26 |   0.350000 |    0.277292 |    0.005887
        27  |         27 |   0.364000 |    0.300156 |    0.006894
        28  |         28 |   0.378000 |    0.324038 |    0.007436
        29  |         29 |   0.392000 |    0.346021 |    0.008366
        30  |         30 |   0.406000 |    0.368885 |    0.009605
        31  |         31 |   0.420000 |    0.390257 |    0.010689
        32  |         32 |   0.434000 |    0.411900 |    0.011541
        33  |         33 |   0.448000 |    0.430898 |    0.013168
        34  |         34 |   0.462000 |    0.451387 |    0.014562
        35  |         35 |   0.476000 |    0.469978 |    0.016112
        36  |         36 |   0.490000 |    0.487686 |    0.017738
        37  |         37 |   0.504000 |    0.504987 |    0.019675
        38  |         38 |   0.518000 |    0.522898 |    0.021611
        39  |         39 |   0.532000 |    0.538639 |    0.024012
        40  |         40 |   0.546000 |    0.554040 |    0.025639
        41  |         41 |   0.560000 |    0.569102 |    0.028428
        42  |         42 |   0.574000 |    0.583757 |    0.030674
        43  |         43 |   0.588000 |    0.597123 |    0.032765
        44  |         44 |   0.602000 |    0.610421 |    0.035399
        45  |         45 |   0.616000 |    0.622295 |    0.038497
        46  |         46 |   0.630000 |    0.634168 |    0.041673
        47  |         47 |   0.644000 |    0.644820 |    0.044849
        48  |         48 |   0.658000 |    0.653776 |    0.048412
        49  |         49 |   0.672000 |    0.663274 |    0.052208
        50  |         50 |   0.686000 |    0.672637 |    0.055229
        51  |         51 |   0.700000 |    0.680779 |    0.058869
```

```
    Bin # |  Bin index |  Bin value |        Eff |    Mis-tag
--------------------------------------------------------------
        0 |          1 |   0.000000 |   0.000000 |   0.000000
        1 |          2 |   0.013000 |   0.000093 |   0.000000
        2 |          3 |   0.026000 |   0.000841 |   0.000000
        3 |          4 |   0.039000 |   0.002709 |   0.000000
        4 |          5 |   0.052000 |   0.007193 |   0.000000
        5 |          6 |   0.065000 |   0.014573 |   0.000000
        6 |          7 |   0.078000 |   0.022326 |   0.000000
        7 |          8 |   0.091000 |   0.034844 |   0.000132
        8 |          9 |   0.104000 |   0.048762 |   0.000397
        9 |         10 |   0.117000 |   0.065577 |   0.000397
       10 |         11 |   0.130000 |   0.085100 |   0.000397
       11 |         12 |   0.143000 |   0.104157 |   0.000397
       12 |         13 |   0.156000 |   0.125922 |   0.000794
       13 |         14 |   0.169000 |   0.145446 |   0.000794
       14 |         15 |   0.182000 |   0.166371 |   0.000926
       15 |         16 |   0.195000 |   0.190659 |   0.001455
       16 |         17 |   0.208000 |   0.215787 |   0.001852
       17 |         18 |   0.221000 |   0.238580 |   0.002117
       18 |         19 |   0.234000 |   0.262774 |   0.002778
       19 |         20 |   0.247000 |   0.283045 |   0.003307
       20 |         21 |   0.260000 |   0.303129 |   0.003836
       21 |         22 |   0.273000 |   0.323494 |   0.004630
       22 |         23 |   0.286000 |   0.344325 |   0.005292
       23 |         24 |   0.299000 |   0.362447 |   0.006350
       24 |         25 |   0.312000 |   0.380757 |   0.006879
       25 |         26 |   0.325000 |   0.398786 |   0.007805
       26 |         27 |   0.338000 |   0.417095 |   0.008467
       27 |         28 |   0.351000 |   0.434657 |   0.009260
       28 |         29 |   0.364000 |   0.452032 |   0.010319
       29 |         30 |   0.377000 |   0.469594 |   0.011113
       30 |         31 |   0.390000 |   0.484073 |   0.012436
       31 |         32 |   0.403000 |   0.501355 |   0.013626
       32 |         33 |   0.416000 |   0.516674 |   0.015081
       33 |         34 |   0.429000 |   0.531154 |   0.016007
       34 |         35 |   0.442000 |   0.545446 |   0.017595
       35 |         36 |   0.455000 |   0.558991 |   0.018918
       36 |         37 |   0.468000 |   0.571789 |   0.020241
       37 |         38 |   0.481000 |   0.586081 |   0.022093
       38 |         39 |   0.494000 |   0.598319 |   0.023945
       39 |         40 |   0.507000 |   0.610929 |   0.025400
       40 |         41 |   0.520000 |   0.623914 |   0.027252
       41 |         42 |   0.533000 |   0.633816 |   0.029898
       42 |         43 |   0.546000 |   0.644839 |   0.032147
       43 |         44 |   0.559000 |   0.654647 |   0.033470
       44 |         45 |   0.572000 |   0.664549 |   0.035851
       45 |         46 |   0.585000 |   0.673891 |   0.038894
       46 |         47 |   0.598000 |   0.680897 |   0.042334
       47 |         48 |   0.611000 |   0.688277 |   0.046170
       48 |         49 |   0.624000 |   0.696030 |   0.049213
       49 |         50 |   0.637000 |   0.702943 |   0.052652
       50 |         51 |   0.650000 |   0.709201 |   0.056092
       51 |         52 |   0.663000 |   0.715367 |   0.059003
```

J. Thaler and K. Van Tilburg.
Identifying Boosted Objects with *N*-subjettiness.
page 26, November 2010.