Analytic understanding of boosted top tagging with N-subjettiness

J. Helliwell
with: M. Dasgupta

Work in progress

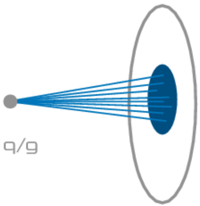
MANCHESTER 1824
The University of Manchester

Jets and their substructure
02/06/2021

Outline

Introduction

- Tagging boosted hadronically decaying top quarks, which are reconstructed as a single jet.
- Jet shapes, prong finding, grooming used in combination ((ATLAS 2016, CMS 2016))
- Understand what drives this type of tagging procedure
- Focus on Y_m-Splitter (M Dasgupta, M Guzzi, J Rawling, G Soyez, 2018) with a cut on the N-subjettiness ratio τ_{32} (J. Thaler, K, Tilburg 2011), and grooming with mMDT/ Soft drop (M. Dasgupta, A Fregoso, S Marzani, G Salam, 2013) (A. Larkoski, S Marzani, G Soyez, J. Thaler, 2014)



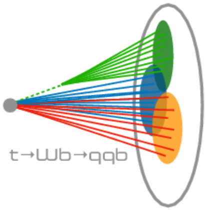
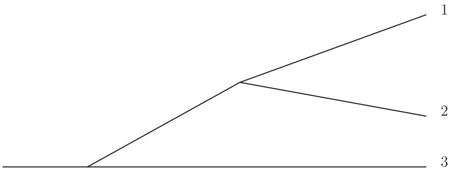


figure adapted from arXiv:1909.12285FERMILAB-PUB-19-492-CMS-E

Y_m-Splitter

- Cluster jet using Gen-$k_t(p = 1/2)$ (M. Cacciari, G. Salam, G. Soyez, 2011) $d_{ij} = \min(z_i, z_j)\theta_{ij}^2$
- Undo last **two** clusterings

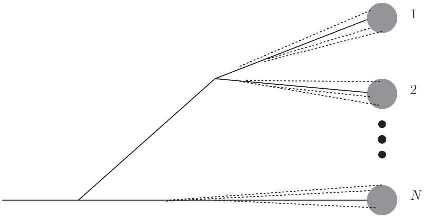


$$\min(m_{12}, m_{13}, m_{23}) > m_{\text{min}}$$

$$m_{\text{min}} \lesssim m_{\text{w}}$$

$$\min(z_1, z_2, z_3) > \zeta$$

$$\tau_N^{(\beta)} = \frac{1}{p_t R_{jet}^{\beta}} \sum_i p_t^i \min\left((\Delta R_{1,i})^{\beta}, (\Delta R_{2,i})^{\beta}, ..., (\Delta R_{N,i})^{\beta}\right)$$



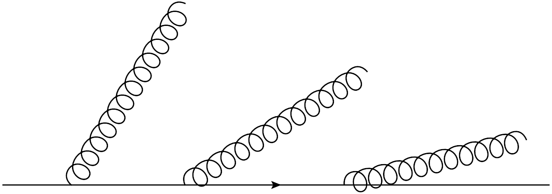
$$\tau_N^{(2)} \simeq \sum_i z_i \min\left((\theta_{1,i})^2, (\theta_{2,i})^2, ..., (\theta_{N,i})^2\right)$$

Use Gen-k$_t(p = {}^1\!/_2)$ axes
$(d_{ij} = \min(z_i, z_j)\theta_{ij}^2)$

$$\tau_{32}^{(2)} = \frac{\tau_3^{(2)}}{\tau_2^{(2)}}$$

1. Cluster jet using Cambridge Aachen algorithm
2. Undo last clustering
3. Check if lower p_t sub-jet has energy fraction, $z > z_{cut}\theta^{\beta}$. If not, discard this sub jet and go back to 2.
4. If it does, this is the jet.

Use $z_{\text{cut}} = \zeta = 0.05$ throughout this work



Monte Carlo Study

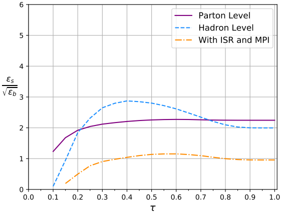
- Performance

- Impact of non-perturbative effects

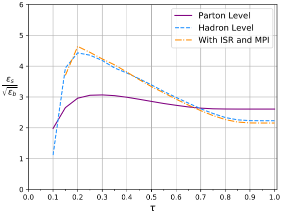
- Using Pythia 8

- 160 GeV $< m_{\text{jet}} < 225$ GeV

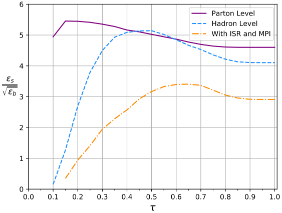
Signal significance



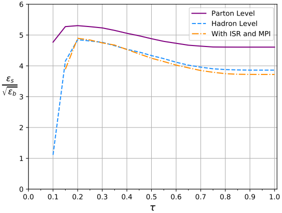
Cut on τ_{32}



Cut on τ_{32} after grooming with mMDT.



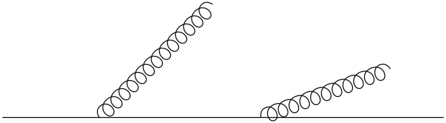
Cut on τ_{32} after application of $\mathrm{Y_m}$-Splitter.



Cut on τ_{32} after application of $\mathrm{Y_m}$-Splitter and grooming with mMDT.

Background Distribution at LO

- Y_m-Splitter requires 3 particles (2 emissions)
- $\rho = \frac{m^2}{p_t^2 R^2}$
- $\tau_{32} = 0$ at $\mathcal{O}(\alpha_s^2)$.
- double logarithmic accuracy



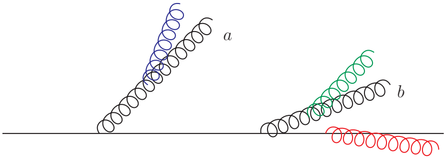
$$\rho \frac{\mathrm{d}\Sigma}{\mathrm{d}\rho} \overset{\zeta > \frac{\rho_{min}}{\rho}}{=} \left(\frac{\alpha_s C_F}{\pi}\right)^2 \ln\left(\frac{\rho}{\rho_{min}}\right) \ln^2(\zeta)\,.$$

$$\overset{\zeta < \frac{\rho_{min}}{\rho}}{=} \left(\frac{\alpha_s C_F}{\pi}\right)^2 \ln^2\left(\frac{\rho_{min}}{\rho}\right)\left(\frac{3}{2}\ln(\zeta) + \ln\left(\frac{\rho_{min}}{\rho}\right)\right)$$

(M Dasgupta, M Guzzi, J Rawling, G Soyez, 2018)

- Leading log accuracy
- Strong ordering in angle and mass
- $\tau_2 = \min(z_1\theta_1^2, z_2\theta_2^2)$, $\tau_3 = \sum_{i=3}^{\infty} z_i\theta_i^2$
- LO result multiplied by Sudakov factor

$$\rho_i = z_i\theta_i^2 \qquad \rho_a = \max(z_1\theta_1^2, z_2\theta_2^2) \qquad \rho_b = \min(z_1\theta_1^2, z_2\theta_2^2)$$



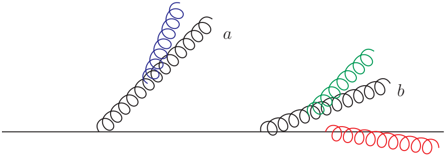
$$S^{\text{(fixed-coupling, soft)}} = \exp\left[-\frac{C_F\alpha_s}{2\pi}\ln^2\frac{1}{\tau\rho_b} - \frac{C_A\alpha_s}{2\pi}\ln^2\frac{\rho_a}{\tau\rho_b} - \frac{C_A\alpha_s}{2\pi}\ln^2\frac{1}{\tau}\right],$$

Accuracy goal of $\alpha_s^n L_\rho^n f_n(\tau)$.

L_ρ is a log or a mass scale i.e. $\ln(\rho)$ or $\ln(\rho_{\text{min}})$

Strong ordering in angle only

$$\tau_{32}(\{p_i\}) = \frac{\sum_{i=1} \rho_i - \rho_a - \rho_b}{\sum_{i=1} \rho_i - \rho_a}$$



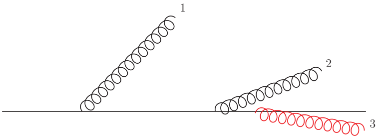
Result valid for $\tau < \frac{1}{2}$

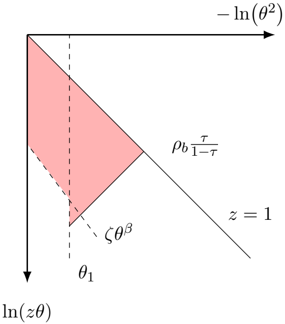
$$\frac{\rho}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\rho} \overset{\tau<1/2}{=} \mathrm{LO} \times {}_2F_1(1, R', 1+R', \tau)\frac{\exp[-R-\gamma_E R']}{\Gamma[1+R']}$$

$$R^{\text{fixed coupling}} = \frac{\alpha_s C_F}{2\pi}\ln^2\left(\rho_b \frac{\tau}{1-\tau}\right)$$

Following method of (D. Napoletano, G. Soyez: 2018)

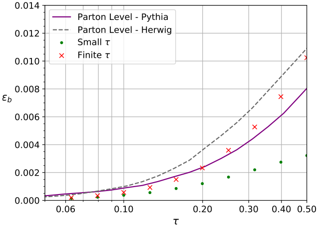
- Sudakov modified as soft wide angle emissions removed
- Complicated by emission 3 setting τ_3 but emission 1 being able to stop the groomer
- Grooming only affects primary emissions (Angular ordering)
- Removes sensitivity to most non perturbative region of phase space



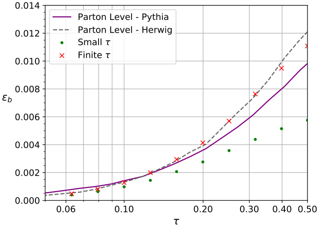


$$R^{(\text{primary})}(\rho_b, \theta_a) = R_{\text{SD}}(\rho_b) + R_{\text{SD}}^{\text{angle}}(\theta_a, \rho_b).$$

Background: Results



No Grooming.



With $\beta = 2$ soft drop

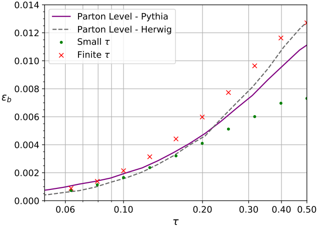
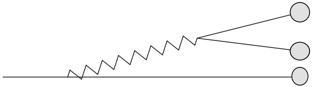
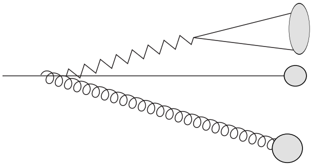


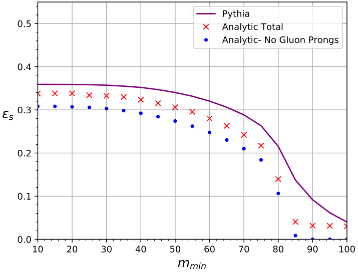
Figure 5: With mMDT

Signal distribution: Y_m-Splitter

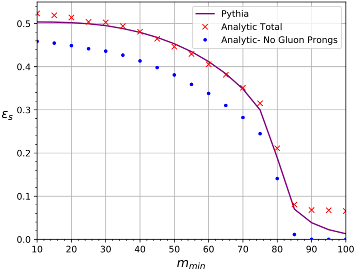


(M Dasgupta, M Guzzi, J Rawling, G Soyez, 2018)





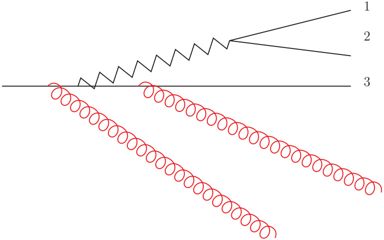
No Grooming.



With mMDT

Key difference between signal and background:

- At LO invariant mass is just the top mass
- No soft/collinear divergences at LO



$$\Sigma^{\tau<\frac{1}{2}}(\rho_{\mathsf{max}}, \rho_{min}, \zeta, \tau) =$$

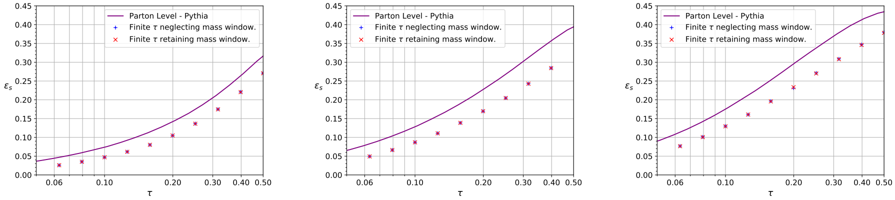
$$\frac{1}{\sigma_0}\int |M_{t\to bq\bar{q}}|^2 \mathrm{d}\Phi_3 \delta\left(\frac{s_{123}}{R^2 p_t^2} - \rho_t\right) \Theta_{\mathsf{Clust}} \Theta_{\mathsf{Y_m\text{-}Splitter}} \frac{e^{-R-\gamma_E R'}}{\Gamma[1+R']}$$

$$R = \frac{\alpha_s C_F}{2\pi} \ln^2\left(\min\left(\frac{\tau}{1-\tau}\min(d_{12}, d_{13}, d_{23}), \rho_{\mathsf{max}} - \rho_t\right)\right)$$

Condition for τ constraint to be stronger than jet mass constraint:

$$\tau_{32} < \frac{M_{\max}^2 - M_T^2}{M_{\max}^2 - M_T^2 + p_t^2 \mathrm{Min}(d_{12}, d_{13}, d_{23})}$$

For $M_{\max} = 225$ GeV the jet mass constraint does not affect the signal distribution for $\tau \lesssim 0.76$



No Grooming. With $\beta = 2$ soft drop With mMDT

- Reducing M_{max} suppresses the background with very little effect on the signal

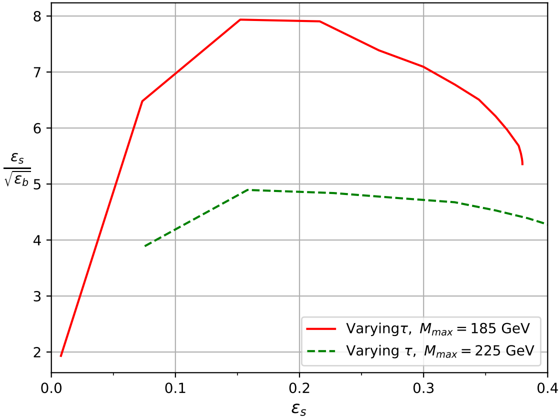


Figure 9: Jets pre-groomed with mMDT, tagged with $\mathrm{Y_m}$-Splitter and a cut on τ_{32}

Summary

- mMDT + Y_m-Splitter + τ_{32} is an effective top-tagger and resilient to non-perturbative effects.
- Used analytic calculations to understand the physics driving this tagging procedure.
- Understanding the interplay between cuts allowed us to use the mass cut to greatest effect.