Better use of Y-splitter for boosted object tagging

Lais Schunk

Work currently being finalized with Mrinal Dasgupta, Alexander Powling and Gregory Soyez.

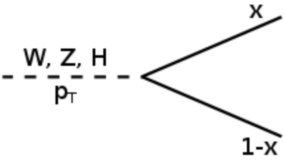
IPhT - CEA Saclay

PSR2016, 4th July, Paris.

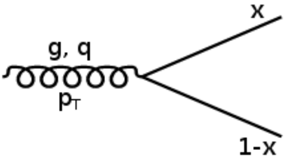
Boosted heavy particles

- At the LHC : production of boosted heavy particles ($p_T \gg m$);
  $\rightarrow$ decay in collimated final states;
  $\rightarrow$ clustered in a single jet.
- QCD jets are also collimated (collinear divergence).

Boosted Z, W, H



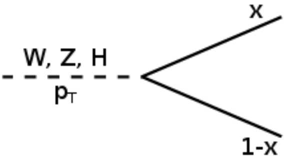
"Standard jet"



Boosted heavy particles

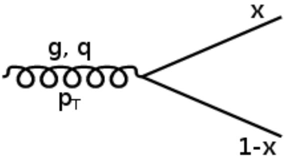
- At the LHC : production of boosted heavy particles ($p_T \gg m$);
  $\rightarrow$ decay in collimated final states;
  $\rightarrow$ clustered in a single jet.
- QCD jets are also collimated (collinear divergence).

Boosted Z, W, H



Signal

"Standard jet"



Background

- How to discriminate QCD jets and hadronic decay jets?

Jet Substructure

- Use *jet substructure* techniques;
- Different techniques are available:

Taggers	find hard prongs in the jets, usually signal has 2 symmetric prongs and background has only 1; *e.g.* Y-splitter, trimming.
Shapes	constrain soft gluon radiation, signal is colorless and has different radiation patterns; *e.g.* Energy correlation, N-subjettiness.
Grooming	"clean" soft gluons at large angles, decreases NP effects; *e.g.* Soft Drop, mMDT, trimming.

Y-splitter

- Undo last step of k_T clustering, i.e. using the distance $d_{ij} = \min\left(p_{Ti}^2, p_{Tj}^2\right)\theta_{ij}^2$;
  Butterworth, Cox and Forshaw (2002)
- Impose the condition $d_{ij}/m_{jet}^2 > y$;
- Retains symmetric 2-pronged structures (keeps signal while eliminating QCD background).

Triming and mMDT

- **Trimming** :

  Krohn, Thaler, Wang (2009)

  1. Recluster jets using $R_{trim} < R$;
  2. Retains only subjets $p_T^{sub} > f_{trim} p_T^{jet}$.

- **(modified) Mass Drop Tagger**:

  Butterworth, Davison, Rubin and Salam, (2008)

  Dasgupta, Fregoso, Marzani and Salam, (2013)

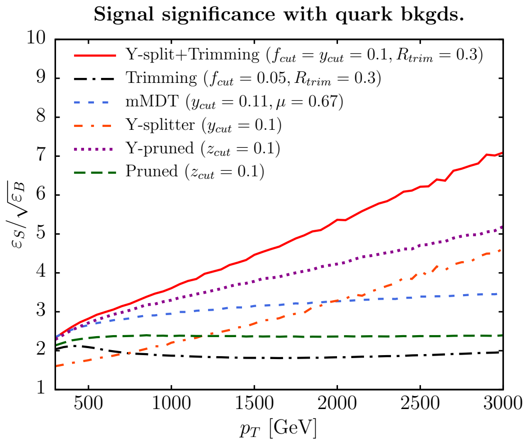
  Apply recursion:

  1. Split jet in two $j \to j_i + j_j$;
  2. Keep splittings that are not too asymmetric $\min\left(p_{Ti}^2, p_{Tj}^2\right)\Delta_{ij}^2/m_{jet}^2 > y_{mMDT}$.

Motivation

- MC studies show an increase in performance for the combination of Y-splitter and trimming;

Dasgupta, Powling, Siodmok (2015)



Our work

- Study Y-splitter and Y-splitter with grooming;
- Compute analytically $\frac{m}{\sigma}\frac{d\sigma}{dm}$ with a cut y;
- Focus on QCD background;
- Compare with Monte Carlo, study impact of NP effects;
- Consider the parameters $f_{trim} = y_{mMDT} = y$;
- Study variants and improvements.

Plain Y-splitter

Leading Order

- We define $\rho = m^2/(p_T^2 R^2) \ll 1$;
- One gluon emission in soft-collinear limit;
- The energy fraction of the gluon x, angle of emission θ;

$$\frac{1}{\sigma}\frac{d\sigma}{d\rho}\bigg|_{\text{plain},y_{cut}}^{\text{LO,soft-coll.}} = \frac{C_F\alpha_s}{\pi}\int_0^1 \frac{dx}{x}\int_0^1 \frac{d\theta^2}{\theta^2}\delta\left(\rho - x\theta^2\right)\Theta\left(x-y\right)$$
$$= \frac{C_F\alpha_s}{\pi}\left(\log\frac{1}{y}\Theta\left(y-\rho\right) + \log\frac{1}{\rho}\Theta\left(\rho-y\right)\right).$$

- For small ρ : $\log(1/\rho) \to \log(1/y)$, reduction of QCD background.

Next-to-Leading Order

- Two gluon emissions, can be virtual or real;
- Y-splitter condition is given by $x_2^2\theta_2^2/(x_1\theta_1^2) < y$;
- We find

$$\frac{\rho}{\sigma}\frac{d\sigma}{d\rho}\bigg|_{\mathrm{plain},y_{cut}}^{\text{NLO, LL}} \overset{\rho<y}{=} \frac{C_F\alpha_s}{\pi}\log\frac{1}{y} \times \left[\frac{C_F\alpha_s}{2\pi}\log^2\frac{1}{\rho} + \ldots\right].$$

Resummation

- For $\rho < y$, we can resum $\log(1/\rho)$ (leading-log)

$$\frac{\rho}{\sigma}\frac{d\sigma}{d\rho}\bigg|^{\text{LL}}_{\text{plain},y_{cut}} = \frac{C_F\alpha_s}{\pi}\log\frac{1}{y} \times \exp\left[-\frac{C_F\alpha_s}{2\pi}\log^2\frac{1}{\rho}\right];$$

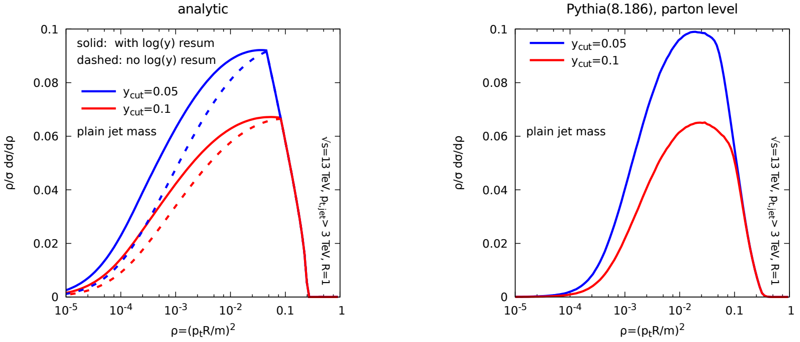
- Compared with plain mass distribution

$$\frac{\rho}{\sigma}\frac{d\sigma}{d\rho}\bigg|^{\text{LL}}_{\text{plain}} = \frac{C_F\alpha_s}{\pi}\log\frac{1}{\rho} \times \exp\left[-\frac{C_F\alpha_s}{2\pi}\log^2\frac{1}{\rho}\right];$$

- Same Sudakov suppression, different pre-factors;
- It is possible to resum $log(1/y)$, expression more complicated.

Comparison with Monte Carlo

- Mass cross section, analytical vs. Monte Carlo simulation.



- Adding the $\log(1/y)$ resummation changes the curve, but not fundamentally.

Y-splitter with Grooming

Trimming at Fixed Order

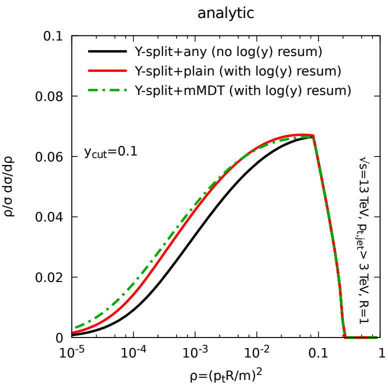
- LO (1 gluon): same thing as Y-splitter.
- NLO (2 gluons): multiple regions depending on y and $r = R_{trim}/R$;
- We can write :

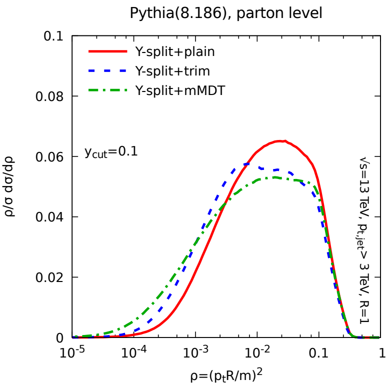
$$\frac{1}{\sigma}\frac{d\sigma}{d\rho}\Big|_{\mathrm{trim},y_{cut}}^{\mathrm{NLO}} = \frac{1}{\sigma}\frac{d\sigma}{d\rho}\Big|_{\mathrm{plain},y_{cut}}^{\mathrm{NLO}} + \mathcal{F}^{\mathrm{trim}}.$$

- When $\rho \to 0$, trimming correction to Y-splitter vanishes , background mostly untouched (LL in $\log(1/\rho)$);
- For other regions, correction is always sub-dominant in $\log(1/\rho)$.

Comparison to Monte Carlo

- Both mMDT and trimming are the same than YS at LL $\log(1/\rho)$;
- mMDT has less transition points;
- Resummation of $\log(1/y)$ is possible (mMDT in paper).





Variants

Mass declustering

- Replace kt declustering by generalized-kt with p = 1/2 (mass-like ordering);
- Removes regime "dominant mass" $\neq$ "dominant k_T";
- Small improvement in performance;
- Same Sudakov exponent as previously (LL).

z cut Condition

- Condition replace by $\min(p_{t1}, p_{t2})/(p_{t1} + p_{t2}) > z_{cut}$;
- Easier to resum higher orders;
- Small loss in performance, less NP effects;
- Same Sudakov exponent (LL).

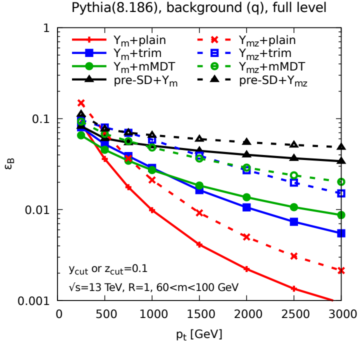
Pre-grooming (SoftDrop)

- Groom jet before Y-splitter;
- Smaller Sudakov suppression (region $x < z_{cut}\theta^{\beta}$ no longer vetoed);
- Loss in performance, less NP effects.

MC study and NP effects

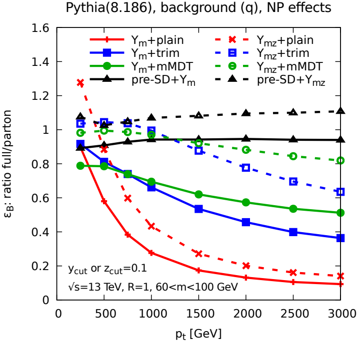
MC simulations and NP effects

Efficiency (ε_B)



Smaller is better.

Sensitivity to NP effects($\varepsilon_B^{full}/\varepsilon_B^{parton}$)

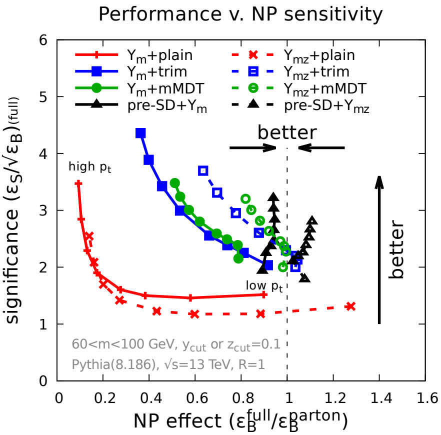


Closer to 1 is better.

- Trade-off between performance and insensitivity to NP effects.

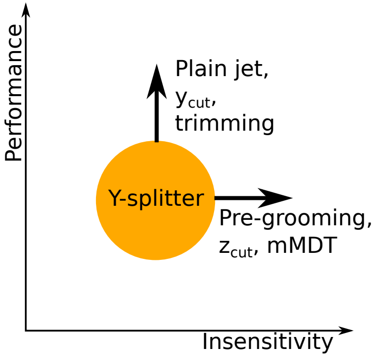
Observations

- Add grooming to Y-splitter increases performance;
- Using z_{cut} or pre-grooming decreases performance and reduces NP effects;
- mMDT slightly less NP effects than trimming, but slightly less discriminant.



Conclusion

- Analytic results for Y-splitter (with or without grooming);
- Suggested variations with various advantages;
- Trade off between performance and insensitivity to NP effects;
- For NP insensitive cases, an optimization based on parton level results would be possible.



Backup slides

Trimming at Fixed Order

- We note p_1 the particle with a dominant mass, region a is $k_{T1} > k_{T2}$ and region b is $k_{T1} < k_{T2}$;

$$\frac{1}{\sigma}\frac{d\sigma}{d\rho}^{\mathrm{NLO}} = \frac{1}{\sigma}\frac{d\sigma}{d\rho}^{\mathrm{NLO,YS}} + \mathcal{F}^{\mathrm{trim,a}} + \mathcal{F}^{\mathrm{trim,b}}.$$

- In the region $\underline{\rho < y^2 r^2}$:

$$\mathcal{F}^{\mathrm{trim,a}} = \frac{1}{\rho}\left(\frac{C_F\alpha_s}{\pi}\right)^2 \ln\frac{1}{r^2}\left(\frac{1}{2}\ln^2 y\right)$$
$$\mathcal{F}^{\mathrm{trim,b}} = -\frac{1}{\rho}\left(\frac{C_F\alpha_s}{\pi}\right)^2 \ln\frac{1}{r^2}\left(\frac{1}{2}\ln^2 y\right)$$
$$\mathcal{F}^{\mathrm{trim,a}} + \mathcal{F}^{\mathrm{trim,b}} = 0.$$

Trimming at Fixed Order

- $\underline{y^2 r^2 < \rho < yr^2}$: $\mathcal{F}^{\rm trim,b}$ remains the same,

$$\mathcal{F}^{\rm trim,a} = \frac{1}{\rho}\left(\frac{C_F\alpha_s}{\pi}\right)^2 \Bigg[-\frac{1}{2}\ln^2\frac{yr^2}{\rho}\ln\frac{\rho}{y(\rho+yr^2)} - \frac{1}{2}\ln^2\frac{1}{y}\ln\frac{y^2}{\rho} + \ln\frac{1}{r^2}\ln\frac{yr^2}{\rho}\ln\left(1+\frac{yr^2}{\rho}\right) - \frac{2}{3}\ln^3\frac{1}{y} + \ln^2\frac{1}{y}\ln\frac{y}{\rho} - \ln\frac{y}{\rho}\ln\frac{yr^2}{\rho}\ln\left(1+\frac{yr^2}{\rho}\right) + \frac{1}{2}\ln^2\frac{yr^2}{\rho}\ln\left(1+\frac{yr^2}{\rho}\right)\Bigg]$$

Trimming at Fixed Order

- $\underline{y^2 > \rho > yr^2}$

$$\mathcal{F}^{\mathrm{trim,a}} = \frac{1}{\rho}\left(\frac{C_F\alpha_s}{\pi}\right)^2\left[\frac{1}{2}\ln\frac{y}{\rho}\ln^2\frac{1}{y} - \frac{1}{6}\ln^3\frac{1}{y}\right]$$

$$\mathcal{F}^{\mathrm{trim,b}} = -\frac{1}{\rho}\left(\frac{C_F\alpha_s}{\pi}\right)^2\frac{1}{2}\ln\frac{y}{\rho}\ln^2\frac{1}{y}$$

- $\underline{y > \rho > y^2}$: $\mathcal{F}^{\mathrm{trim,b}}$ remains the same,

$$\mathcal{F}^{\mathrm{trim,a}} = \frac{1}{\rho}\left(\frac{C_F\alpha_s}{\pi}\right)^2\frac{1}{2}\ln^2\frac{y}{\rho}\ln\frac{1}{y}$$

- $\underline{\rho > y}$: no Y-splitter effects.

mMDT at Fixed Order

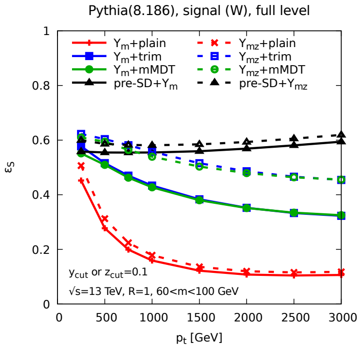
- For $\rho < y^2$:

$$\mathcal{F}^{\mathrm{mMDT}} = \frac{-1}{6} \ln^3 \frac{1}{y}.$$

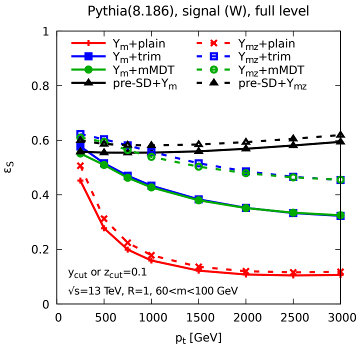
- For $y > \rho > y^2$: sum of contributions a and b of trimming in same region.

MC simulations and NP effects (signal)

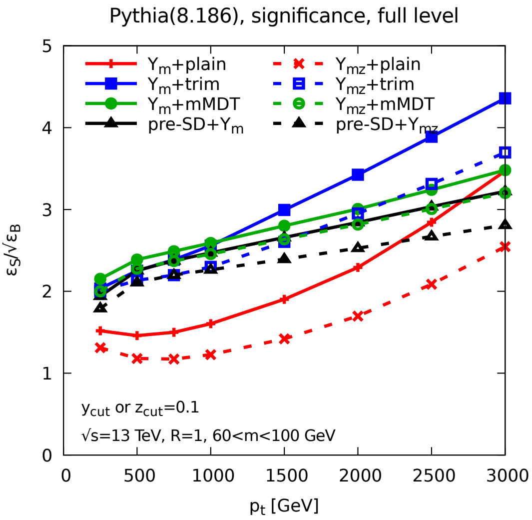
Efficiency (ε_B)



Sensitivity to NP effects($\varepsilon_B^{full}/\varepsilon_B^{parton}$)

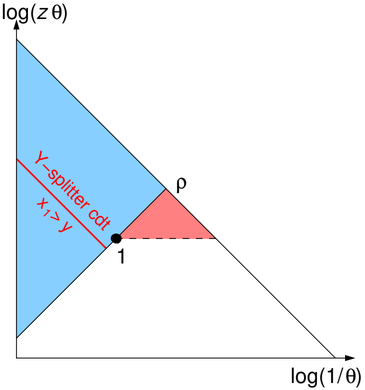


Significance

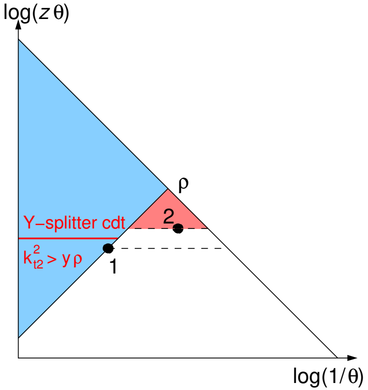


Lund diagram Y-splitter

Emission that dominates the jet mass also has the largest k_t.



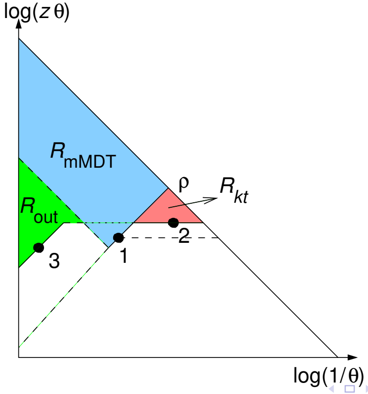
Emission with larger k_t than the k_t of the emission that dominates the mass.



Resummation for Y-splitter with mMDT

$$\frac{\sigma}{\rho}\frac{d\sigma}{d\rho} = \int_y^1 dx_1\, P(x_1)\,\frac{\alpha_s(\rho x_1)}{2\pi}\, e^{-R_{\text{mMDT}(\rho)}}$$

$$\left[e^{-R_{k_t}(k_{t1};\rho)-R_{\text{out}}(k_{t1}^2/y)} + \int_{k_{t1}}^{\sqrt{\rho}} \frac{dk_{t2}}{k_{t2}} R'_{k_t}(k_{t2};\rho) e^{-R_{k_t}(k_{t2};\rho)-R_{\text{out}}(k_{t2}^2/y)}\right]$$



Y-splitter with pre-grooming

- Y-splitter applied on a pre-groomed jet with SoftDrop.
- Shadowed area : the region allowed by SoftDrop and entering into the Sudakov factor.
- Dashed (red) line: to the Y-splitter (z_m) condition.