,

Distinguishing *b*-quark and gluon jets with a tagged *b*-hadron

Robin Linten

Based on arxiv:1512.05265

21st April 2016

Content

Motivation

- b-jets important final state at the LHC
- SM: $H \to b\bar{b}$, $t \to Wb$, ...
- BSM: decay cascades of heavy particles, interactions with the (SM) Higgs, ...
- large QCD backgrounds from $g \to b\bar{b}$ in the same jet - not a b-jet, rather a gluon jet
- identifying double b-jets offers handle on this background
- large uncertainties on $g \to b\bar{b}$-splittings in the shower

Double b-tagging

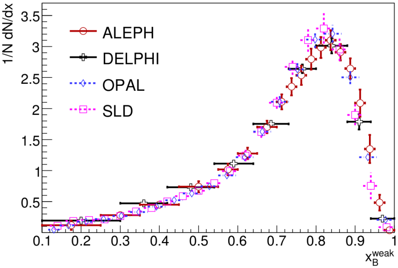
- need a way to identify double b-hadron jets
- experimentally, there are a few techniques:
  - find two b-vertices: difficult due to opening angle and background [1]
  - substructure observables: build MVA's - train on b-jet samples from dijets or radion decay [2]
- moderate success in the identification - not standard usage yet (changed for Run 2?)

[1] D. Acosta et al. Measurements of $b\bar{b}$ azimuthal production correlations in $p\bar{p}$ collisions at $\sqrt{(s)} = 1.8$ TeV. Phys. Rev., D71:092001, 2005.

[2] ATLAS-CONF-2012-100, CMS DP-2015/038

Shaping *b*-jets

- in order to find strong observables, consider QCD radiation pattern
- fragmentation function of *b*-quarks is relatively hard
- collinear divergence in gluon emissions being shielded by the mass
- roughly symmetric energy share between *b*-quarks from gluon splitting



Eur.Phys.J.C71:1557,2011, arxiv:1102.4748

Shaping b-jets

Strong observables incorporate these considerations, e. g.:

- number of tracks:

$$n_{ch} = \sum_{i \in J} i$$

- girth aka jet width:

$$g = \sum_{i \in J} \frac{p_\perp^i \Delta R_{iJ}}{p_\perp^J}$$

- energy fraction:

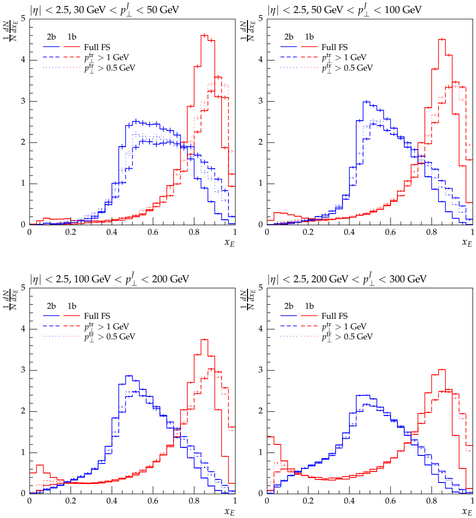
$$x_E = \frac{E_B}{E_J}$$

- (subjettiness, planar flow, eccentricity, Fox-Wolfram moments, ...)

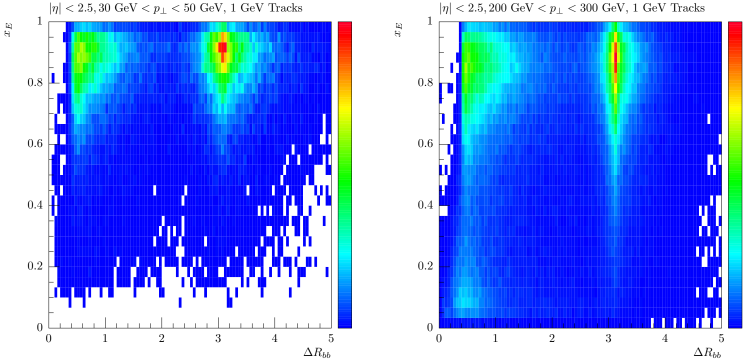
Setup

- Pure QCD $pp \to jets$ sample at $\sqrt{s} = 13$ TeV
- Use SHERPA, with $2 \to 2$ matrix elements at LO, CSS shower, cluster hadronization, UE
- Analysis performed using Rivet:
  - anti-k_T jets with $R = 0.4$, $p_\perp^J > 30$ GeV, $|\eta_J| < 2.5$
  - define charged tracks with minimum $p_\perp^{\mathrm{track}} \geq 0.5$ GeV or 1 GeV
- Check against HERWIG++ and PYTHIA 8 with respective default setup - found good agreement

Single Observables: Energy fraction x_E

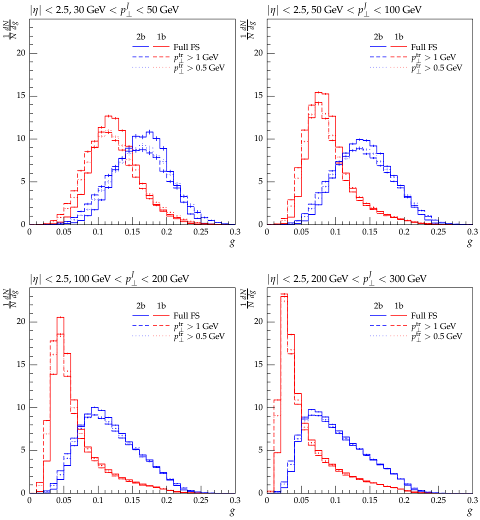


But what about those low x_E events? Soft radiation leaking in

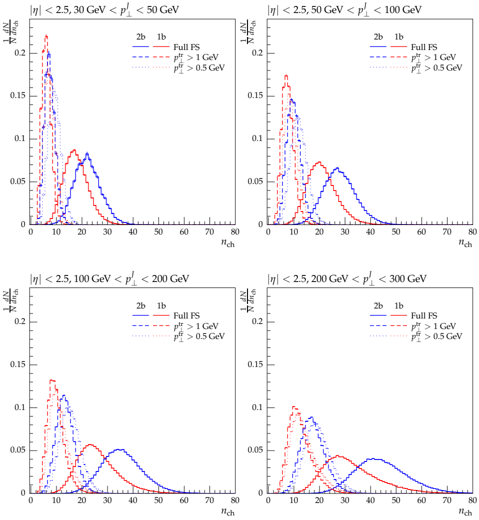


- choose to place cut for $x_E < 0.3$

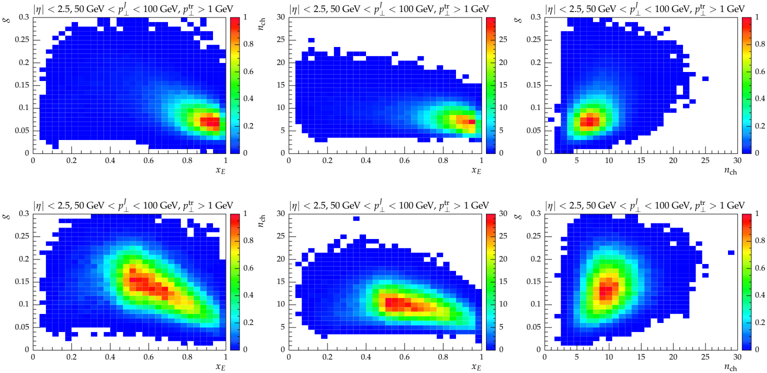
Single Observables: Girth g



Single Observables: Multiplicity n_{ch}



Combining Observables: Correlations (top: 1b, bottom: 2b)

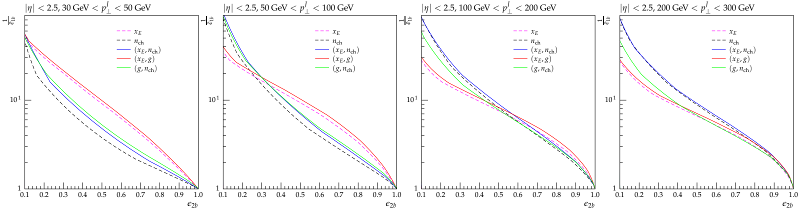


- have fairly strong correlation between observables - but in a usable way

Efficiencies

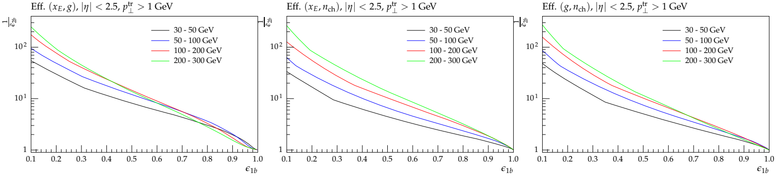
- efficiencies defined via ROC curve: cut on the observable, slide this along the observable to get efficiencies
- generalize to 2D distributions: rotate axis onto the largest eigenvector of the correlation - slide cut along this axis
- simple but useful even if observables are strongly correlated

Comparison of observables



- visible improvement from our simple combination of two observables
- best combinations of observables involve x_E

Performance for different $p_\perp^J$ slices

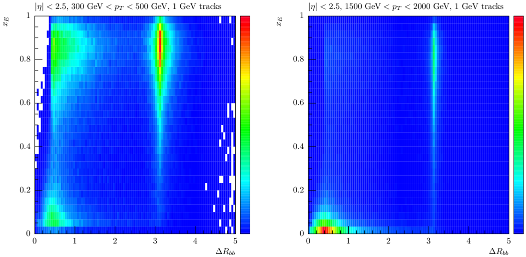


- performance improves for larger $p_\perp^J$

Conclusions

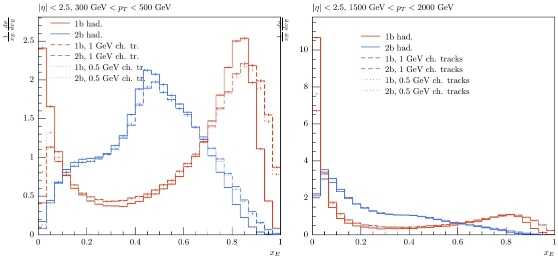
- studies requiring multiple *b*-jets will become increasingly important - so will identifying "legitimate" *b*-jets
- substructure observables prove valuable for this purpose
- knowledge of QCD radiation properties helps choose strong observables
- new observable considered: energy fraction x_E
- simple combination of correlated observables can lead to improvements in efficiencies

Backup



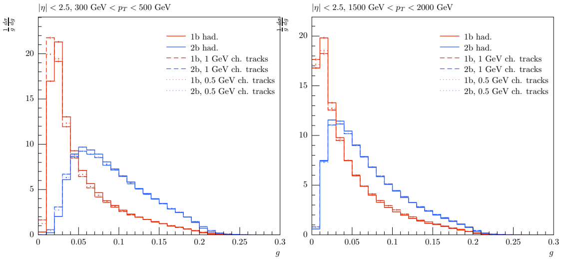
|η| < 2.5, 300 GeV < p_T < 500 GeV, 1 GeV tracks
x_E
ΔR_{bb}
0
0.2
0.4
0.6
0.8
1
0
1
2
3
4
5
|η| < 2.5, 1500 GeV < p_T < 2000 GeV, 1 GeV tracks
x_E
ΔR_{bb}
0
0.2
0.4
0.6
0.8
1
0
1
2
3
4
5

Backup



|η| < 2.5, 300 GeV < p_T < 500 GeV
1b had.
2b had.
1b, 1 GeV ch. tr.
2b, 1 GeV ch. tr.
1b, 0.5 GeV ch. tr.
2b, 0.5 GeV ch. tr.
|η| < 2.5, 1500 GeV < p_T < 2000 GeV
1b had.
2b had.
1b, 1 GeV ch. tracks
2b, 1 GeV ch. tracks
1b, 0.5 GeV ch. tracks
2b, 0.5 GeV ch. tracks
x_E

Backup


$|\eta| < 2.5$, 300 GeV $< p_T <$ 500 GeV
$|\eta| < 2.5$, 1500 GeV $< p_T <$ 2000 GeV
$\frac{1}{\sigma}\frac{d\sigma}{dg}$
1b had.
2b had.
1b, 1 GeV ch. tracks
2b, 1 GeV ch. tracks
1b, 0.5 GeV ch. tracks
2b, 0.5 GeV ch. tracks
g