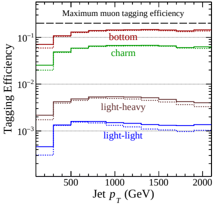
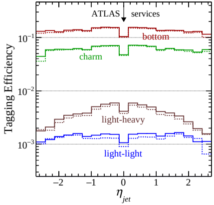
Boosted-bottom jet tagging and BSM searches

Zack Sullivan

Illinois Institute of Technology
CTEQ Collaboration

March 21, 2016





In collaboration with Keith Pedersen and Daniel Duffty
Pedersen, Z.S., Phys. Rev. D 93, 014014 (2016) [arXiv:1511.05990]
Duffty, Z.S., Phys. Rev. D 90, 015031 (2014) [arXiv:1307.1820]

(and Pedersen, Z.S., arXiv:1606:xxxxx)

Outline

1. **Introduction**
   - Multi-TeV $W' \to tb$ and boosted tags
2. **The μ_x boosted-b tag**
3. **Search for a multi-TeV leptophobic $Z' \to b\bar{b}$**
4. **TeV-scale charged Higgs ($tH^{\pm} \to t + \mathrm{boosted}(tb)$)**
5. **Conclusions**

TeV-scale vector currents are generic in BSM

Most extensions of the Standard Model predict heavy, narrow particles which couple via a **vector current** ... the W' and Z'

Some model classes

Sequential Standard Model

Left-right symmetric models: Broken $\mathrm{SU}(2)_L \times \mathrm{SU}(2)_R$

- Generic mixing of W_L–W_R
  R. N. Mohapatra, J. Pati, A. Salam, G. Senjanovic, ...
- Orbifold-breaking — suppressed mixing, enhanced couplings
  Y. Mimura, S. Nandi, ...
- Supersymmetric L–R models
  M. Cvetic, J. Pati, ...

Models with additional left-handed W' and Z'

- Little Higgs: SU(5)/SO(5), SU(6)/SP(6), SU(N)/SU(N–1), ...
  T. Gregoire, N. Arkani-Hamed, S. Chang, H. C. Cheng, A. Cohen, I. Low, D. E. Kaplan, E. Katz, O. C. Kong, A. Nelson, M. Schmaltz, W. Skiba, D. Smith, J. Terning, J. Wacker, ...
- Topcolor — topflavor, leptophobic topflavor seesaw, generic mixing
  H. Georgi, H. J. He, E. Jenkins, X. Li, E. Ma, E. Malkawi, D. Muller, S. Nandi, E. Simmons, T. Tait, C. P. Yuan, ...
- Extra dimensions: Kaluza-Klein modes of the W and Z
  A. Datta, P. O'Donnell, T. Huang, Z. Lin, X. Zhang, ...
- Non-commuting extended technicolor
  R. Chivukula, E. Simmons, J. Terning, ...

+ 1000's more

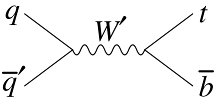
Model-independent $W' \to tb$ searches at the LHC

The cross section factorizes:

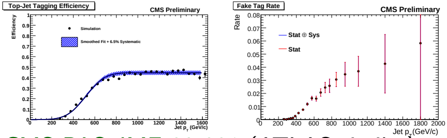
$$\sigma_{NLO} = (g'/g_{SM})^2 \times \sigma_{NLO}^{SM}$$

To extend the mass reach > 1.5 TeV we proposed a boosted-top search

Duffty, Z.S., PRD 90, 015031 (2014)



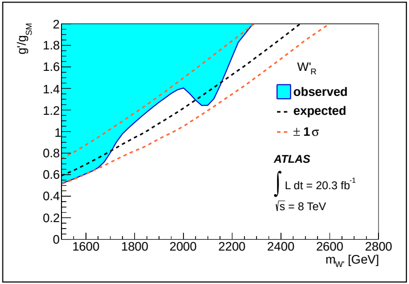
We use CMS boosted top efficiencies



CMS PAS JME-09-001 (ATLAS similar)

Warning: *fast* montecarlo tends to underestimate fake rates.

EPJC**75**(15)165[1408.0886] Fig. 8b



ATLAS performed this study and improved the reach $M_{W'} > 1.8$ TeV, but not as much as we hoped.

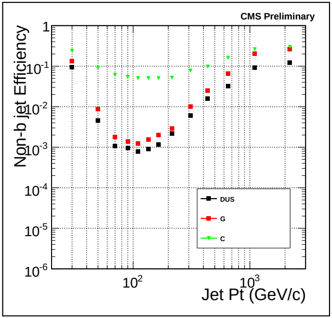
Why? You have to tag the recoil boosted-b jet as well...

We need updated public efficiencies from CMS and ATLAS.

When b jets are boosted, the fake rates kill you

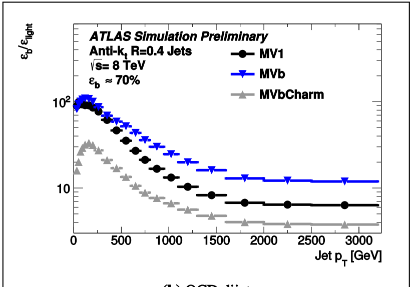
- Probability to tag light flavors *rises dramatically* for boosted jets!
  - Light jet = no b or c **hadrons**; experiments can't differentiate b-initiated jets and $g \to b\bar{b}$ jets.
- Huge (40%) systematic uncertainties in tagging efficiency can dominate experimental results/exclusions.

[CMS PAS BTV-09-001] Fig. 12



Maintaining 50% b jet efficiency

[ATL-PHYS-PUB-2014-014] Fig. 14b

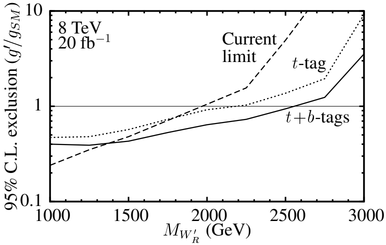


(b) QCD dijet

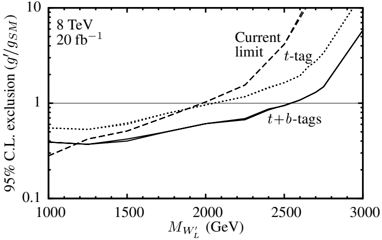
W' searches w/ boosted-top and boosted-bottom

D. Duffty, ZS, Phys. Rev. D 90, 015031 (2014) [1307.1820]

Right-handed W'



Left-handed W'



- We find factor 2-3 improvements in g'/g_{SM} above 1.5 TeV can be obtained with existing data.
- We can extend the mass reach near $g'/g_{SM} = 1$ to 2.5 TeV.
- We have reach for some perturbative models up to 3 TeV.
- The critical feature to improve the reach was a newly proposed boosted-bottom-jet tag . . .

Outline

1. Introduction
   - Multi-TeV $W' \to tb$ and boosted tags
2. The μ_x boosted-b tag
3. Search for a multi-TeV leptophobic $Z' \to b\bar{b}$
4. TeV-scale charged Higgs ($tH^{\pm} \to t + \text{boosted}(tb)$)
5. Conclusions

μ_x boosted-b jets, and substructure variables

Pedersen, ZS, Phys. Rev. D 93, 014014 (2016) [arXiv:1511.05990]

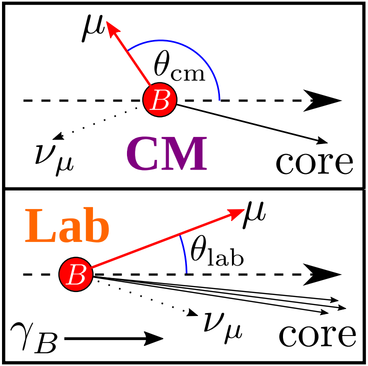
- **CM**: The muon is emitted with speed $\beta_{\mu,\mathrm{cm}}$ at angle θ_{cm}.
- **Lab**: Muon is detected at angle θ_{lab} w.r.t. the centroid of the *decay subjet* (boosted by γ_B).

$$p_{\mathrm{subjet}} = p_\mu + p_{\nu_\mu} + p_{\mathrm{core}}$$

- Defining $\kappa \equiv \beta_{\mathrm{B}}/\beta_{\mu,\mathrm{cm}}$, construct

$$x \equiv \gamma_B \tan(\theta_{\mathrm{lab}}) = \frac{\sin(\theta_{\mathrm{cm}})}{\kappa + \cos(\theta_{\mathrm{cm}})}$$

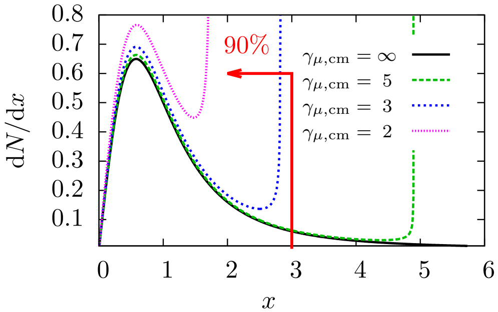
$$\boxed{x \approx \tan(\theta_{\mathrm{cm}}/2) \qquad (\text{when } \kappa \approx 1)}$$



Theoretical lab frame muon distributions

- We are interested in specific boosted subjets ...
  - boosted b jets ($p_T \geq 300$ GeV $\Longrightarrow \gamma_B \gtrsim 60$).
  - b hadron decays ($\gamma_{\mu,\mathrm{cm}} \leq \frac{m_B}{2\,m_\mu} \lesssim 25$)
- Muon Number density $\frac{dN_\mu}{dx} = \frac{2x}{(x^2+1)^2} K(x,\kappa)$ (when $\kappa \geq 1$)

  ($K(x,\kappa)$ restricts muons to boost cone boundary ($x \leq 1/\sqrt{\kappa^2-1}$))



- Once $\gamma_{\mu,\mathrm{cm}} \gtrsim 3$, **lab** muons approach a universal boosted shape.

The μ_x boosted b tag: two simple cuts

$$p_{\rm subjet} = p_\mu + p_{\nu_\mu} + p_{\rm core}$$

- $x \leq 3$ only indicates the muon is consistent with a *boosted* decay.
- It's heavy-flavor origin can be confirmed via a complementary measurement … the stiff fragmentation function for $b \to B$ causes the B subjet to carry a *large fraction* of its jet's momentum.

$$x \equiv \gamma_B \tan(\theta_{\rm lab}) \leq 3 \qquad\qquad f_{\rm subjet} \equiv \frac{p_{T,\rm subjet}}{p_{T,\rm jet}} \geq 0.5$$

But, half the muons in b jets come from secondary c hadron decays!
Is γ_B a valid observable?

No … but we can observe $\gamma_{\rm subjet}$.

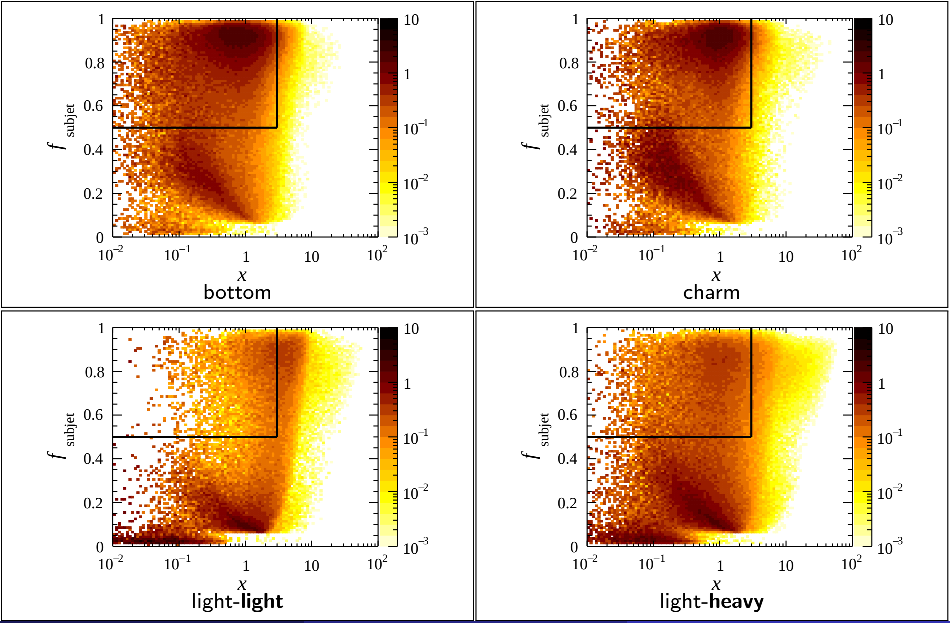
Reconstructing the boosted subjet

Anti-k_T jets are clustered with $\mathrm{R} = 0.4$. Allowing muons to participate lets *hard muons* seed jet formation.

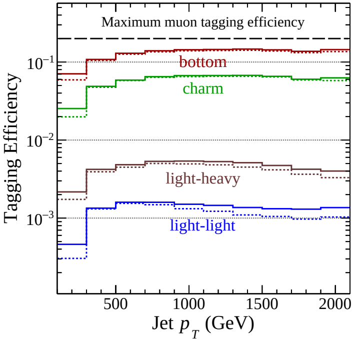
$$p_{\rm subjet} = p_\mu + p_{\nu_\mu} + p_{\rm core}$$

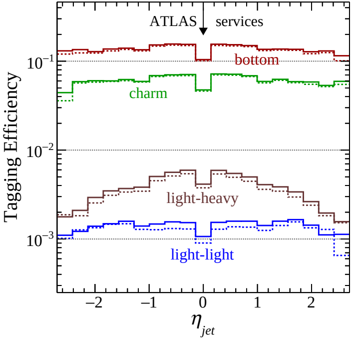
- *Taggable* **muons** must pass a quality cut ($p_T \geq 10\ \mathrm{GeV}$).
- The **core** (the hadronic remnants of the semi-leptonic decay).
  - Re-cluster jet using $R = 0.04$ to localize core (3×3 grid)
  - $\gamma_{\rm subjet}$ needs mass of core — very poorly measured. Core mass is constrained to *best guess* (e.g. $m_D \approx 2$ GeV).
  - The "correct" core brings $\sqrt{p^2_{\rm subjet}}$ closest to $m_B \approx 5.3$ GeV.
- Subjet's **neutrino**:
  - System is under-determined. Simplest estimate: *add muon a second time to simulate neutrino* ($p_{\nu_\mu} = p_\mu$).

μ_x tags applied to b, c, and light initiated jets



μ_x boosted b tagging efficiencies at LHC





- Boosted kinematics turn on at 300 GeV.
- **Light jets** classified by hadronic origin of taggable muon (light-heavy are b jets from gluon splitting — rapidity handle?)

Signal efficiencies

- $\sim$ 14% of b-jets
- $\sim$ 6.5% of c-jets

Light jet fake rate

- Light-light $\mathcal{O}(0.1\%)$
- All light $\mathcal{O}(0.5\%)$

Pileup helps (a bit)

- *Solid*: no pileup
- *Dotted*: $\mu = 40$

Outline

Searching for a leptophobic $Z' \to b\bar{b}$

We begin with *Dobrescu and Yu* [1306.2629, 1506.04435] — a simple, renormalizable, leptophobic Z'_B

- Coupling to quarks is flavor independent, $\mathscr{L} = \frac{g_B}{6} Z'_{B\mu} \bar{q}\gamma^\mu q + \ldots$
- Narrow width: $\Gamma_{Z'}/M_{Z'} \approx \frac{1}{6}\alpha_B \left(1 + \frac{\alpha_S}{\pi}\right) \approx 1$–$5\%$

MadGraph5 (w/ CT14llo) → Pythia 8 → Delphes 3 (w/ `FastJet 3`)

- Generate MLM matched Z'_B samples for a variety of $M_{Z'_B}$,
  $pp \to Z'_B \to b\bar{b}/c\bar{c}(+j)$.
- QCD dominates background: $pp \to b\bar{b}/c\bar{c}/j\bar{j}(+j)$, $jq_h \to jq_h(+j)$
- Look for signal excess in $d\sigma/dM_{jj}$ of width $[0.85, 1.25] \times M_{Z'}$ in 2-tag and 1-tag inclusive classes.
- We developed a custom Delphes module (`MuXboostedBTagging`) to implement μ_x tagging, available on GitHub:

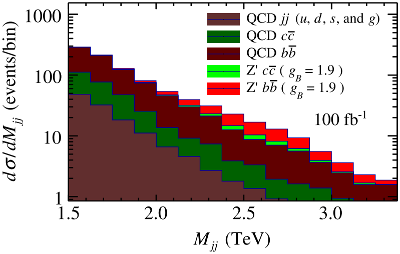
  https://github.com/keith-pedersen/delphes/tree/MuXboostedBTagging

Leptophobic Z' discovery signal at LHC

Pedersen, ZS, Phys. Rev. D 93, 014014 (2016) [arXiv:1511.05990]

2 μ_x-tag discovery

($M_{Z'} = 2.5$ TeV, $g_B = 1.9$)

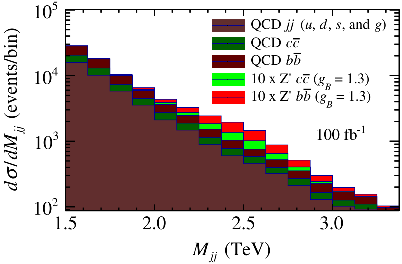


Mass window: 2.125–3.125 TeV

A clean channel with $S/B = 1/2$.

1 μ_x-tag inclusive discovery

($M_{Z'} = 2.5$ TeV, $g_B = 1.3$)

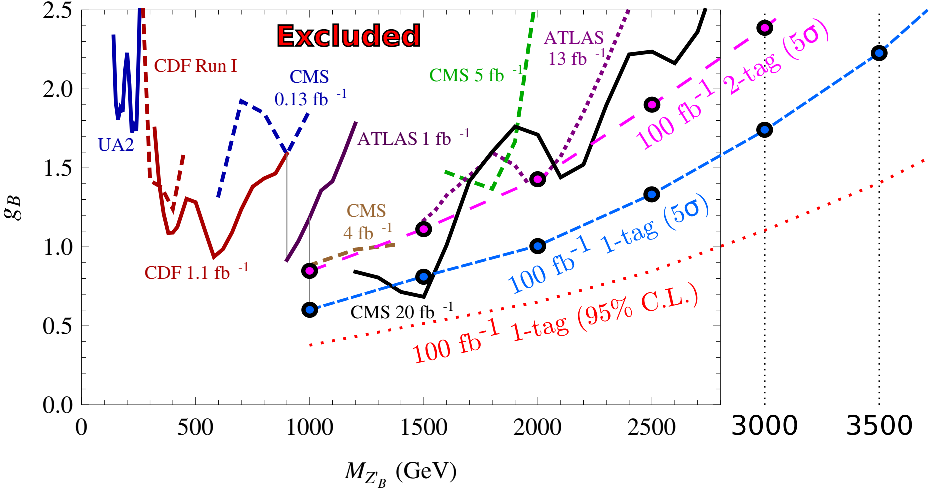


$|\eta_{\rm jet}| \leq 2.7, \quad |\Delta\eta_{jj}| \leq 1.5$

$S/B = 1/12.5$, but with $12\times$ the 2-tag signal, discovery is possible for smaller coupling.

Leptophobic Z' reach at 13 TeV LHC

Compared to PRD**88**(13)035021[1306.2629] Fig. 1

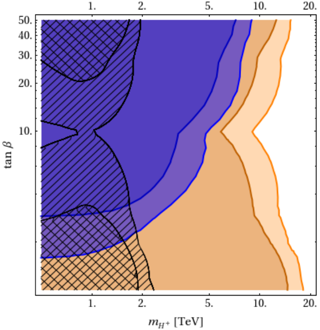


Outline

Heavy (charged) Higgs: $pp \to \bar{t}bH^+ \to \bar{t}b(t\bar{b})$

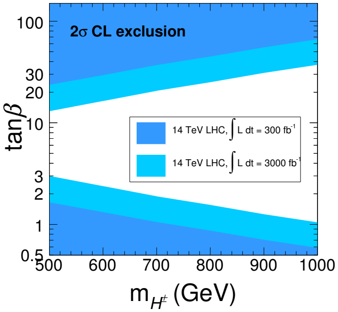
Associated production of MSSM Higgs produces a final state rich in *b*-jets

Hajer et al., JHEP 11(2015)124



Exclusion potential $tbH^\pm \to \bar{t}b(t\bar{b})$, at $\sqrt{s} = 100(14)$ TeV (hatched) $b\bar{b}t\bar{t}$ (orange) and $tb\tau\nu$ (blue), in 3/30 ab^{-1} (0.3/3 ab^{-1}).

Craig et al., 1504.04630



Based on personal communication, we believe the $t\&b$ tagging efficiencies and fake rates were **over-optimistic**.

Much *less* optimistic in key "wedge" region (stopped at 1 TeV).
Fake rates still underestimated.

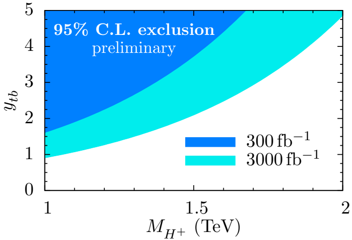
- Can we reach 2 TeV wedge?
- How well can μ_x tagging do?

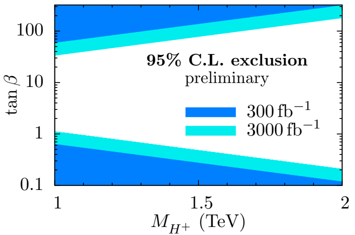
Reexamining charged Higgs: $pp \to \bar{t}H^+ \to \bar{t}(t\bar{b})$

Pedersen, ZS, [arXiv:1606.xxxxx]

FeynRules → MadGraph5 → Pythia 8 → Delphes 3 (w/ `FastJet` 3)

- Generate MLM matched $tH^\pm \to t\bar{t}b(+j)$ samples at **14 TeV** LHC.
- Signal: 1 boosted-t, 1 μ_x boosted-b, 1 reconstructed $t \to b_{\rm norm} l\nu$
- ttj dominates background: (boosted b fake), followed by tjj (+boosted-t fake — using CMS measured tagging and fake rates)
- Cut on boosted jets: $\frac{P_{Tt_b}-P_{Tb_b}}{P_{Tt_b}+P_{Tb_b}}$, $\Delta\phi(P_{Tt_b}, P_{Tb_b}) \Rightarrow S/B \sim 1/10$



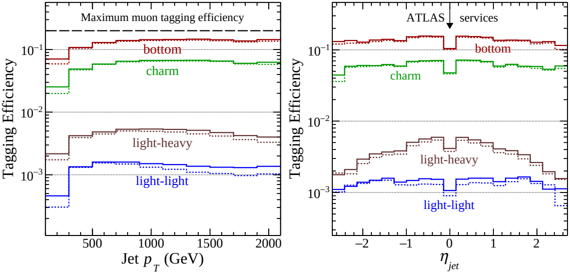


Results so far look like extension of Craig et al. to higher mass.

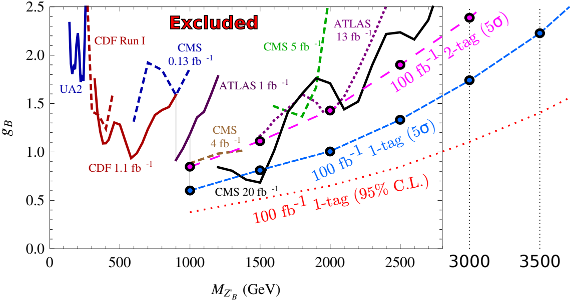
It seems we should focus on a 100 TeV collider for $H^+ \to tb$.

Conclusions

- μ_x tags heavy jets at the TeV scale.
  - **b jet**: $\sim$14%
  - **light-light**: $\sim$0.1%
- Flat p_T/η_{jet} response & minimal pileup sensitivity

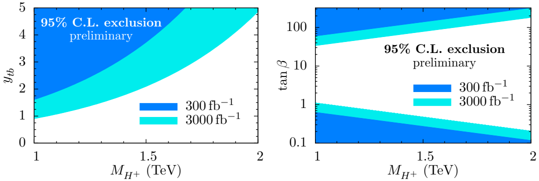


- $W' \to (\text{boosted})tb$ can improve limits
- μ_x boosted-b tagging offers a significant improvement for leptophobic Z' searches



Conclusions

- $t + \text{boosted}(tb)$ can improve charged Higgs reach at 14 TeV.
- "Wedge region" will be a 100 TeV measurement.



- A DELPHES module for μ_x boosted-b tagging can be found at `https://github.com/keith-pedersen/delphes/tree/MuXboostedBTagging`

We look forward to continue working with the CMS and ATLAS Collaborations to implement these ideas.

THANK YOU

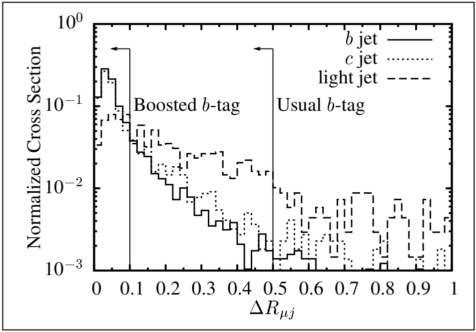
Backup Slides

A muon-based boosted-b tag

A boosted-b tag was proposed by Duffty and Sullivan in PRD **90**, 015031 (2014)

- Muon ($p_T \geq 20$ GeV) within a cone of $\Delta R = 0.1$ around jet's centroid.
  - A purely angular cut.
  - Compare to muon p_T^{rel}, which conflates well-measured angle with poorly measured energy (radiative losses at high-p_T).
- *But* a jet's centroid is **coarse** (QCD radiation, UE, pileup ...).
  - We can do better by studying boosted b tagging in the context of jet substructure.

PRD**90**(14)015031[1307.1820] Fig. 2



Tagging efficiencies

Jet type	100 GeV	400 GeV	1000 GeV
b	4.8%	11.8%	15.0%
c	2.1%	5.5%	7.5%
light	0.1%	0.4%	0.6%

Aside: What is x?

Given $p_{\nu_\mu} = p_\mu$, we can imagine reconstructing an arbitrary subjet:

$$p_{\rm subjet} = 2p_\mu + p_{\rm core}$$

What x will we measure? Let's express it in terms of **direct** observables.

$\gamma_{\rm core}$	$\lambda = \frac{2E_\mu}{E_{\rm core}}$	ξ (the angle between muon and **core**)

If $\beta \to 1$ for both the muon and the core,

$$x(\xi) \approx \underbrace{\gamma_{\rm core} \frac{1+\lambda}{\sqrt{1+2\lambda\,\gamma_{\rm core}^2[1-\cos(\xi)]}}}_{\gamma_{\rm subjet}} \underbrace{\frac{\sin(\xi)}{\cos(\xi)+\lambda}}_{\tan(\theta_{\rm lab})}$$

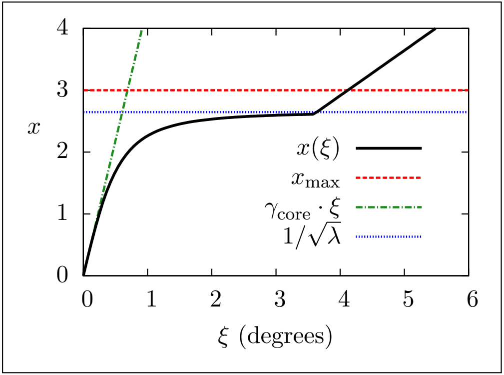
Angle where ξ dominates $m_{\rm subjet}$
$\xi_m = \sqrt{\frac{m_{\rm core}^2}{2E_{\rm core}\,E_\mu}}$

$\xi < \xi_m$	$\xi \gtrsim \xi_m$
$x(\xi) \approx \gamma_{\rm core} \cdot \xi$	$x \approx 1/\sqrt{\lambda}$

Aside: What is x? It is a "smart" angle

A poorly reconstructed $m_{\rm subjet}$ is inevitable; a **large** $m_{\rm subjet}$ is *inconsistent* with heavy-hadron decay. So we implement a ceiling

$$m_{\rm subjet} = \min(\sqrt{p^2_{\rm subjet}}, 12~{\rm GeV})$$



Subjet with a **hard muon** ($\gamma_{\rm core} = 250$, $\lambda = 1/7$)

Solve for $\xi_{\rm max}$, the largest ξ which keeps $x \leq 3$.

Hard muons ($\lambda \geq 1/9$)

$$\xi^{\textbf{hard}}_{\rm max} \approx \frac{18}{\gamma_{\rm core}}$$

Soft muons ($\lambda < 1/9$)

$$\xi^{\textbf{soft}}_{\rm max} \approx \frac{3}{\gamma_{\rm core}} \left(\frac{1}{\sqrt{1 - 9\lambda}} \right)$$

Flavor tagging

heavy jets (b or c initiated) $\overset{\textbf{distinguish?}}{\Longleftrightarrow}$ **light** jets (d, u, s, or g initiated).

- Heavy quark ($m \gtrsim \Lambda_{\mathrm{QCD}}$) *decay functions* peak near $z = 1$ (versus $z = 0$ for light partons).
  - Heavy quarks spawn heavy hadrons carrying a large fraction of $\vec{p}_{\mathrm{jet}}$.
- b/c hadrons decay at a secondary vertex (SV)...
  - Far enough from the primary vertex to be resolved
  - Close enough to rule out other particles (e.g. K_S^0)
- b/c hadrons have a modest rate of semi-leptonic decay ($l \in \{e, \mu\}$):
  - $\mathcal{B}(b \to l\nu_l X) \approx 11\%$
  - $\mathcal{B}(c \to l\nu_l X) \approx 10\%$ (thus 20% of b jets have $N_{muon} \geq 1$)

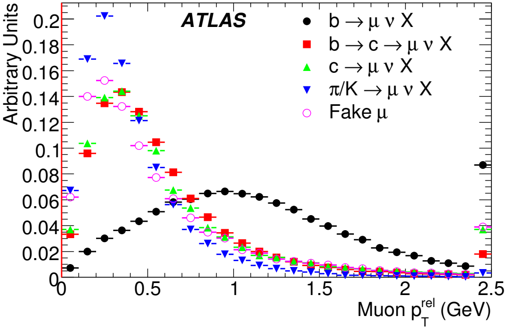
b hadrons $\to$ c hadrons, so generally ...

heavy-flavor tags $\to$ b tags

$p_T^{\rm rel}$ muon tagging

- 20% of b jets have $N_{muon} \geq 1$
- Electrons in jets are hard to identify; luckily someone ordered the *muon chamber!*
- Previous studies have investigated $\mathbf{p_T^{rel}}$: muon momentum transverse to the *centroid* of its jet.

[ATLAS-PHYS-PUB-2009-021] Fig. 4



- b hadrons → large mass, hard muons → higher $p_T^{\rm rel}$.
  - $\epsilon_b = \mathcal{O}(10\%)$, light jet fake rate $= \mathcal{O}(0.3\%)$.
- $p_T^{\rm rel}$ stops working when jet p_T exceeds 140 GeV.
  - *Is this a problem of definition?*

The direction of the core is extremely important!

- Tracks provide the best angular information, *but* ...
  - Accurately tracking boosted jet constituents in a *fast detector* simulator is not possible; we only track "standalone" muons.
  - Jets are clustered from **Cal towers** and **muons**.

- **Trimming:** Before reclustering, discard Cal towers with low jet p_T fraction (we choose $f_{\rm tower}^{\rm min} = 0.05$). This reduces the core's sensitivity to *pileup*, *UE*, soft *QCD* , etc.

- **ECal pointing:** Use the segmentation of the ECal to orient the combined (ECal+HCal) tower. This creates a *minimal angular resolution* independent of track reconstruction efficiencies.
  - We use the dimensions of ATLAS ECal L2: $(\Delta\, r\phi \times \Delta\eta = \mathbf{0.025} \times \mathbf{0.025})$
  - Also ran coarser ($\mathbf{0.05} \times \mathbf{0.05}$); no degradation of heavy jet efficiency, the light jet fake rate is 1.2 times larger at jet $p_T = 600$ GeV, but *no enhancement* in fake rate at $p_T = 2$ TeV.

Z' bosons

Dileptons are the "golden channel"
— clean reconstruction, low background

- Current limits $Z' \to l^+l^-$ (SM-like)
  $M_{Z'} > 2.9$ TeV
  (ATLAS/CMS, $\sqrt{S} = 8$ TeV)

Many models decouple from leptons

Some limits on *leptophobic* Z'

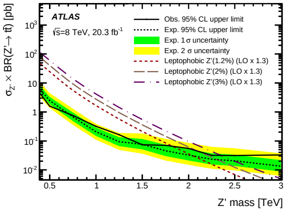
- Topcolor-assisted technicolor $Z' \to t\bar{t}$
  — $M_{Z'} > 1.8$ TeV
- Sequential SM (leptophobic) $Z' \to b\bar{b}$
  — $M_{Z'} > 1.7$ TeV

Regardless, leptophobic means a *dijet* signal, and the dreaded QCD background.

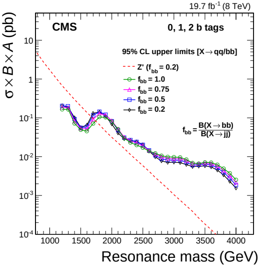
- We must flavor tag the jets (t or b)!

Leptophobic Z' limits

JHEP**1508**(15)148[1505.07018] Fig. 11a



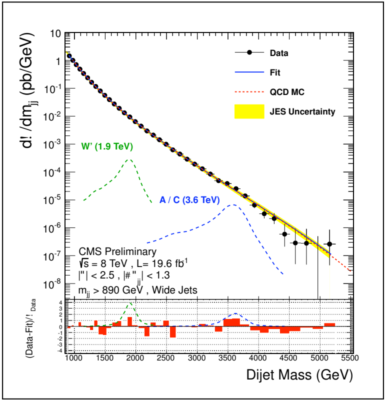
PRD **91**(15)052009[1501.04198] Fig. 4c



Z' Dijet Searches

- Tons of dijet data, but
  - QCD is *strong*, huge BG.
  - *Wide* jets capture QCD radiation; pileup smears dijet mass resolution.
- Leading processes at α^2?
  - **Z′**: $q\bar{q} \xrightarrow{s} q'\bar{q}'$
  - **QCD heavy dijets (c, b)**: $gg \xrightarrow{t} q\bar{q} \mid q\bar{q} \xrightarrow{s} q\bar{q}$
  - **QCD light dijets (d, u, s, g)**: $gq \xrightarrow{t} gq \mid q\bar{q} \xrightarrow{t} q\bar{q} \mid gg \xrightarrow{s} gg$
- $\dfrac{S}{\sqrt{BG}} \approx \dfrac{Z'}{\sqrt{\text{QCD light}}}$
  - *Tag b/c-jets to drive down light-jet QCD BG!*

[CMS-PAS-EXO-12-059] Fig. 1



- Wide jets ($R = 1.1$)
- $|\eta_{\text{jet}}| \le 2.5$, $|\Delta\eta_{jj}| \le 1.3$