Flavor Tagging TeV Jets for BSM and QCD

Keith Pedersen
(kpeders1@hawk.iit.edu)

ILLINOIS INSTITUTE OF TECHNOLOGY

In collaboration with **Zack Sullivan**
To appear in arXiv:1508.xxxxx

Z' Models

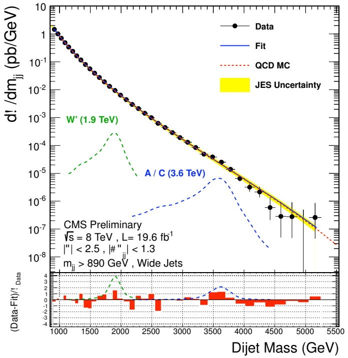
- One of the simplest BSM models is an additional $U(1)$ symmetry, mediated by a neutral heavy boson (Z').
  - Experiments don't see heavy resonance in dilepton channels – what if the Z' is "leptophobic"?
  - Dijets are discovery channel!

Z' Models

- One of the simplest BSM models is an additional $U(1)$ symmetry, mediated by a neutral heavy boson (Z').
  - Experiments don't see heavy resonance in dilepton channels – what if the Z' is "leptophobic"?
  - Dijets are discovery channel!
- *Dobrescu* and *Yu* [arXiv:1306.2629 and 1506.04435] outline simple, renormalizable, leptophobic Z'_B
  - No tree-level couplings to SM leptons and bosons, only SM quarks are charged. This implies association with *baryon* number B.
  - Coupling to quarks is flavor independent,
    - $\frac{1}{6} g_B Z'_{B\mu} \bar{q} \gamma^{\mu} q$
  - Narrow width:
    - $\Gamma_{Z'}/M_{Z'} \approx \frac{1}{6}\alpha_B \left(1 + \frac{\alpha_S}{\pi}\right) \approx$1–5% (depending on g_B)
  - Model needs vector-like fermions (anomalons); we assume they're "kinematically inaccessible".

- Dijet search gives *lots of data*, but
  - Strong QCD BG (all quarks are "massless" at 2 TeV).
  - Precise reconstructing of radiation and $t\bar{t}$ uses *wide* jets ... pileup smears mass resolution.

CMS-PAS-EXO-12-059 (Fig. 1)

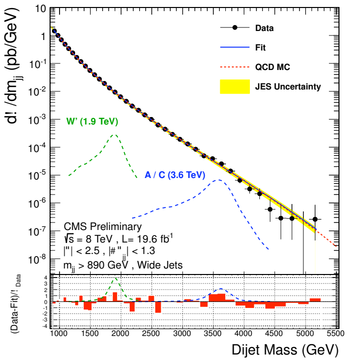


- Wide jets ($R = 1.1$)
- $|\eta_{jet}| \leq 2.5, \ |\Delta\eta_{jj}| \leq 1.3$

Z' Dijet Searches

- Dijet search gives *lots of data*, but
  - Strong QCD BG (all quarks are "massless" at 2 TeV).
  - Precise reconstructing of radiation and $t\bar{t}$ uses *wide* jets ... pileup smears mass resolution.
- Leading processes at α^2?
  - $\boldsymbol{Z'}$: $q\bar{q} \xrightarrow{s} q'\bar{q}'$
  - **QCD heavy dijets (cbt)**: $gg \xrightarrow{t} q\bar{q} \mid q\bar{q} \xrightarrow{s} q'\bar{q}'$
  - **QCD light dijets (dusg)**: $gq \xrightarrow{t} gq \mid q\bar{q} \xrightarrow{t} q\bar{q} \mid gg \xrightarrow{s} gg$

CMS-PAS-EXO-12-059 (Fig. 1)

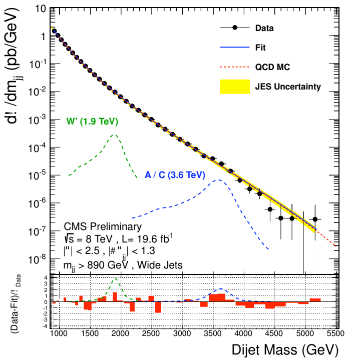


- Wide jets ($R = 1.1$)
- $\left|\eta_{jet}\right| \leq 2.5, \left|\Delta\eta_{jj}\right| \leq 1.3$

Z' Dijet Searches

- Dijet search gives *lots of data*, but
  - Strong QCD BG (all quarks are "massless" at 2 TeV).
  - Precise reconstructing of radiation and $t\bar{t}$ uses *wide* jets ... pileup smears mass resolution.
- Leading processes at α^2?
  - **Z'**: $q\bar{q} \stackrel{s}{\to} q'\bar{q}'$
  - **QCD heavy dijets (cbt)**: $gg \stackrel{t}{\to} q\bar{q} \mid q\bar{q} \stackrel{s}{\to} q'\bar{q}'$
  - **QCD light dijets (dusg)**: $gq \stackrel{t}{\to} gq \mid q\bar{q} \stackrel{t}{\to} q\bar{q} \mid gg \stackrel{s}{\to} gg$
- $\frac{S}{\sqrt{BG}} \approx \frac{Z'}{\sqrt{\text{QCD light}}}$
  - *Tag b/c-jets to drive down light-jet QCD BG!*

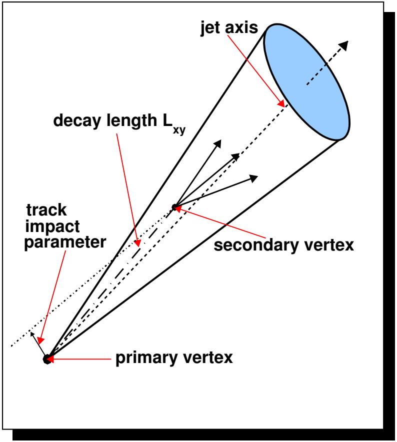
CMS-PAS-EXO-12-059 (Fig. 1)



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

Track-based *b*-tag

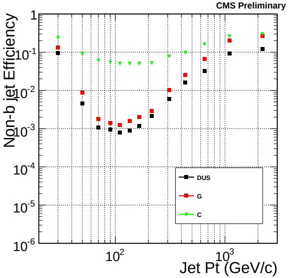
- *b*-hadrons have long lifetimes and decay at a secondary vertex (**SV**). Use tracks to find an **SV** inside a jet.
  - High efficiency for *b*-jets (.5-.8).
  - Also tag **charm** and **light** jets.



Track-based b-tag

- b-hadrons have long lifetimes and decay at a secondary vertex (**SV**). Use tracks to find an **SV** inside a jet.
    - High efficiency for b-jets (.5-.8).
    - Also tag **charm** and **light** jets.
- Probability to tag light jets increases severely as jet p_T approaches TeV.
    - 10^{-3} at 100 GeV $\rightarrow 10^{-1}$ at 1 TeV

CMS PAS BTV-09-001 (Fig. 12)

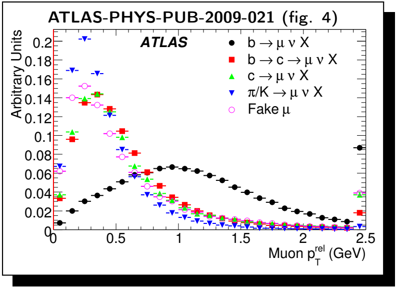


Maintaining 50% b-jet efficiency

- Resolving **SV** inside high p_T jets is *limited by tracking performance*.
    - Tracks have higher p_T and bend less – harder to constrain.
    - Tracks are more collimated – increased sensitivity to tracker resolution.
    - Dense tracking environment – fake tracks *and* false duplicates.
- Gluons split $(g \rightarrow b\bar{b})$ more often at high p_T – light flavor BG!

Muon-based *b*-tag

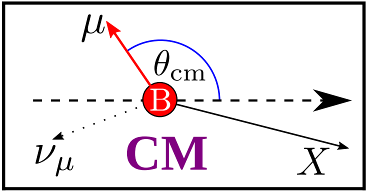
- *b*/*c*-hadrons frequently decay semi-leptonically (where $l \in \{e, \mu\}$):
  - $BR(b \to l\nu_l X) \approx 11\%$
  - $BR(c \to l\nu_l X) \approx 10\%$
  - 20% of *b*-jets have $N_{muon} \geq 1$
- Hard to see electrons inside jets; we're lucky that someone ordered the *muon chamber*!



- The angle muons make with boost axis of mother should be larger if its mother had a high mass (and thus a heavy flavor).
  - Look at p_T^{rel}(muon momentum transverse to the centroid of its jet)
  - p_T^{rel} loses efficiency for boosted decays because *the centroid of the entire jet* is too coarse a measure – need *subjet of decay*.

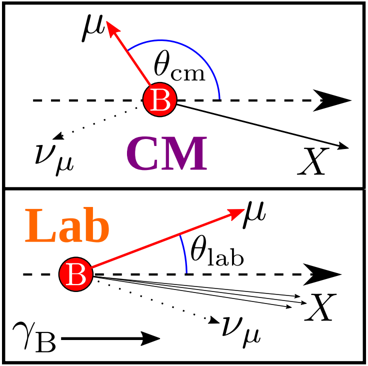
Muon Emission vs. Semi-leptonic Subjet

- **CM**: Muon is emitted with $\gamma_{\mu,\mathrm{cm}}$ at angle θ_{cm} w.r.t. the boost axis.



Muon Emission vs. Semi-leptonic Subjet

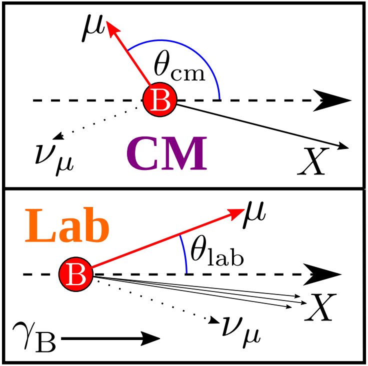
- **CM**: Muon is emitted with $\gamma_{\mu,\mathrm{cm}}$ at angle θ_{cm} w.r.t. the boost axis.
- **Lab**: Muon is detected at angle θ_{lab} w.r.t. the centroid of the subjet (which has boost γ_{B}).
  - Define $\kappa \equiv \beta_{\mathrm{B}}/\beta_{\mu,\mathrm{cm}}$



Muon Emission vs. Semi-leptonic Subjet

- **CM**: Muon is emitted with $\gamma_{\mu,\mathrm{cm}}$ at angle θ_{cm} w.r.t. the boost axis.
- **Lab**: Muon is detected at angle θ_{lab} w.r.t. the centroid of the subjet (which has boost γ_{B}).
  - Define $\kappa \equiv \beta_{\mathrm{B}}/\beta_{\mu,\mathrm{cm}}$

$$x \equiv \gamma_{\mathrm{B}} \tan(\theta_{\mathrm{lab}}) = \frac{\sin(\theta_{\mathrm{cm}})}{\kappa + \cos(\theta_{\mathrm{cm}})} \quad (1)$$

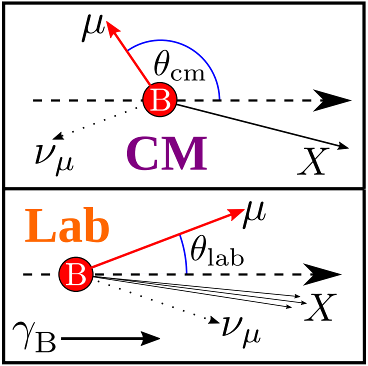


Muon Emission vs. Semi-leptonic Subjet

- **CM**: Muon is emitted with $\gamma_{\mu,\mathrm{cm}}$ at angle θ_{cm} w.r.t. the boost axis.
- **Lab**: Muon is detected at angle θ_{lab} w.r.t. the centroid of the subjet (which has boost γ_{B}).
    - Define $\kappa \equiv \beta_{\mathrm{B}}/\beta_{\mu,\mathrm{cm}}$

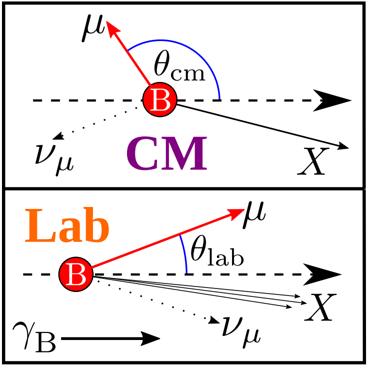
$$x \equiv \gamma_{\mathrm{B}} \tan(\theta_{\mathrm{lab}}) = \frac{\sin(\theta_{\mathrm{cm}})}{\kappa + \cos(\theta_{\mathrm{cm}})} \quad (1)$$

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



Muon Emission vs. Semi-leptonic Subjet

- **CM**: Muon is emitted with $\gamma_{\mu,\mathrm{cm}}$ at angle θ_{cm} w.r.t. the boost axis.
- **Lab**: Muon is detected at angle θ_{lab} w.r.t. the centroid of the subjet (which has boost γ_{B}).
  - Define $\kappa \equiv \beta_{\mathrm{B}}/\beta_{\mu,\mathrm{cm}}$



$$x \equiv \gamma_{\mathrm{B}} \tan(\theta_{\mathrm{lab}}) = \frac{\sin(\theta_{\mathrm{cm}})}{\kappa + \cos(\theta_{\mathrm{cm}})} \quad (1)$$

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

$$\frac{dN}{dx} = 4\pi \frac{2x}{(x^2+1)^2} K(x,\kappa) \quad (\text{when } \gamma_{\mathrm{B}} \gg \gamma_{\mu,\mathrm{cm}} \to \kappa > 1) \qquad (3)$$

x Marks the Heavy Flavor tag

- We can calculate the **lab frame** region $[0, x_\delta]$ which captures $(1-\delta)$ of the CM emission. It's not pretty

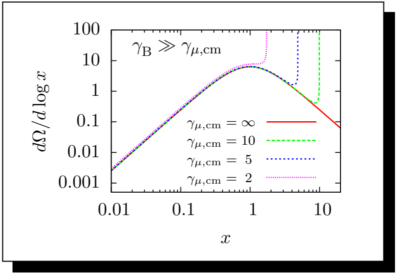
$$x_\delta = \sqrt{\frac{1-\kappa^2-2\delta^2+\sqrt{(\kappa^2-1)^2+4\kappa^2\delta^2}}{2\delta^2}} \tag{4}$$

- Taking the limit as $\delta \to 0$ gives $x_{\max} = 1/\sqrt{\kappa^2-1}$.
- More important is the region which captures 90% of emitted muons

$$x_{90\%} \approx 3\left(1-\frac{5}{2}\gamma_{\mu,\mathrm{cm}}^{-2}\right) \approx 3 \tag{5}$$

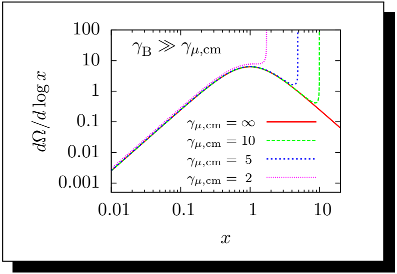
- If you find a muon in a jet, ask:
  - "Is the muon consistent with a *very boosted emission* from the jet's **primary hadron** (i.e. is $x \le x_{90\%}$)?"
- If so, use the muon to tag the jet – *this is the* μ_x *tag*.

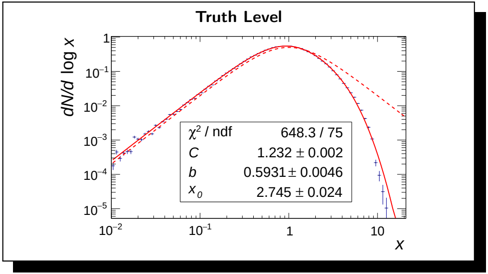
Testing dN/dx



- Muons restricted to $x_{\max}$, defined by $\gamma_{\mu,\mathrm{cm}}$.
- $K(x, \kappa)$ corrects for muon speed; small corrections for ultra-relativistic muons.

Testing dN/dx





- Muons restricted to $x_{\max}$, defined by $\gamma_{\mu,\mathrm{cm}}$.
- $K(x,\kappa)$ corrects for muon speed; small corrections for ultra-relativistic muons.

- **Exact**: Integrate over $\gamma_{\mu,\mathrm{cm}}$ and distribute asymptote.
- **Approximate**: apply cut-off function to asymptotic $\frac{d\Omega}{d\log x}$
  - $\frac{2x^2}{(x^2+1)^2} \times C\,\frac{(1+\exp(-bx_0))}{(1+\exp(b(x-x_0)))}$

Measuring x

$$\mathbf{p}_{\text{subjet}} = \mathbf{p}_\mu + \mathbf{p}_{\text{core}} + \mathbf{p}_{\nu_\mu}$$

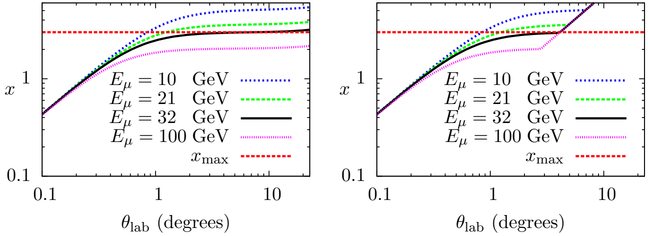
- Require **muons** to pass $p_{T,\mu}^{\min}$ (we use 10 GeV).
  - Allow them to participate in initial jet clustering (*hard muon seeds*).
- Isolate subjet's hadronic/EM energy (the **core**).
  - Re-cluster jets to find *extremely thin* core:
    - $R \approx 1/\gamma = \mathcal{O}(.01)$ (larger in practice)
  - γ_{subjet} needs ***mass*** of core – very poorly measured.
    - Constrain mass to *best guess* (e.g. $m_{D^\pm/D^0} \approx 2$ GeV).
  - Choose core which gets subjet mass closest to $m_{B^\pm/B^0} \approx 5.3$ GeV.
- Subjet's **neutrino**:
  - System is under-determined. Simplest estimate: *add muon again to simulate neutrino*.

What is μ_x Doing?

$$m_{\text{subjet}} \approx \sqrt{m_{\text{core}}^2 + 4E_{\text{core}}E_\mu\left(1 - \cos\left(\theta_{\text{lab}}\right)\right)} \quad (6)$$

$$x \approx \frac{E_{\text{core}} + 2E_\mu}{\min\left(m_{\text{subjet}}, m_{\text{subjet}}^{\text{max}}\right)} \tan\left(\theta_{\text{lab}}\right) \quad (7)$$

- When θ_{lab} begins to be dominate m_{subjet}, x flattens. If E_μ is also large, it pushes the plateau *below* x_{max} — any θ_{lab} causes tag.
- Using $m_{\text{subjet}}^{\text{max}}$ prevents arbitrarily large θ_{lab} from passing the x cut.



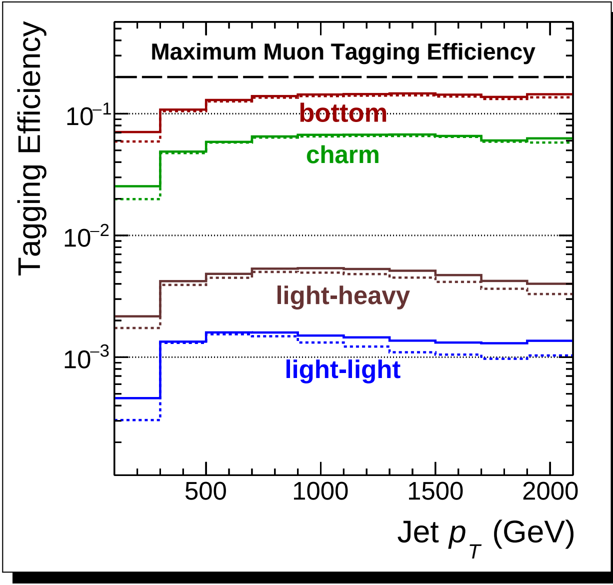
$E_{\text{subjet}} = 500$ GeV

$m_{\text{core}} = 2$ GeV

$m_{\text{subjet}}^{\text{max}} = 12$ GeV

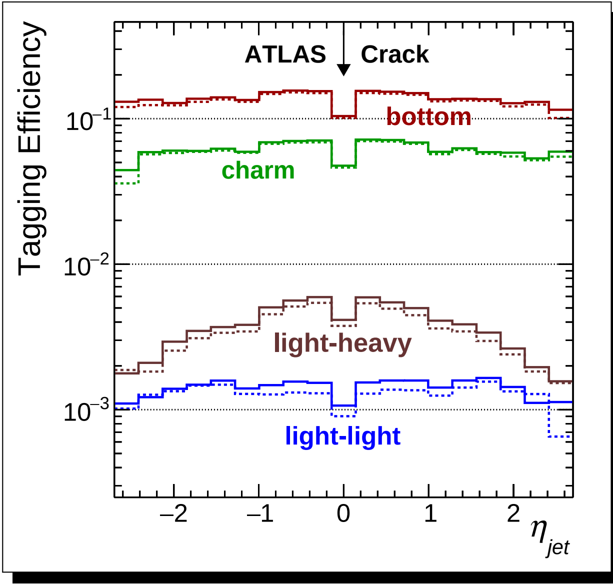
Tagging Efficiency (Jet p_T)

- All studies simulated at ATLAS at $\sqrt{s} = 13$ TeV.
- Boosted kinematics turn on at 300 GeV.
- **light-heavy**:
  light jet fragments to heavy quark
- **light-light**:
  light jet stays light
- *Solid* line:
  no pileup
- *Dotted* line:
  $\langle \mu \rangle = 40$



Tagging Efficiency (Jet η)

- Tag jets with $x \leq 3$
  - $\sim$ 14% of b-jets
  - $\sim$ 6.5% of c-jets
  - Mis-tag $\sim$ 0.65% of light jets
- η dependence *effectively flat* for bottom and charm.
  - Slight dip in endcap ($\eta > 1$).
  - Large dip at ATLAS detector services crack.
- Light jets lose extra forward efficiency.



Simulated Z' Bump Hunt

MadGraph 5 → Pythia 8 → Delphes 3 (with `FastJet` 3)

- Plot $d\sigma/dM$ for dijet systems where *both* top jets are tagged.
  - Proof of concept for μ_x tag; Delphes simulates a very simple detector.
  - Need member of ATLAS or CMS to implement in μ_x Geant4.

Simulated Z' Bump Hunt

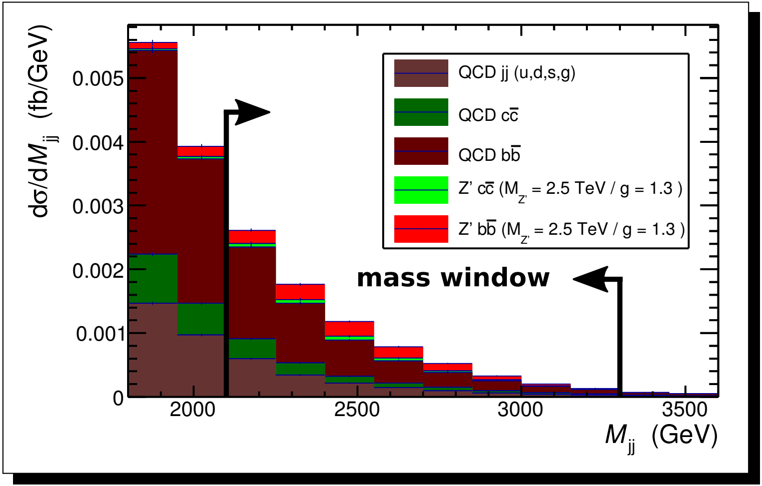
MadGraph 5 → Pythia 8 → Delphes 3 (with FastJet 3)

- Plot $d\sigma/dM$ for dijet systems where *both* top jets are tagged.
  - Proof of concept for μ_x tag; Delphes simulates a very simple detector.
  - Need member of ATLAS or CMS to implement in μ_x Geant4.
- Plenty of muons in **heavy** jets, good statistics for double-tagging.
- The double-tag rate for **light** jets $\left((\sim 6 \times 10^{-3})^2 \approx 4 \times 10^{-5}\right)$ is *too small* (he complained) — need $\mathbf{10^9}$ events for good statistics
  - When one jet is tagged, scale event weight by one half times the probability to tag the recoil jet (from tagging efficiency survey)
  - Either jet could produce the real tag (probabilistic OR):

$$\left(1-(1-\epsilon_j)^2\right) \times \tfrac{1}{2}\epsilon_j = \left(2\epsilon_j - \cancelto{0}{\epsilon_j^2}\right) \tfrac{1}{2}\epsilon_j \approx \epsilon_j^2$$

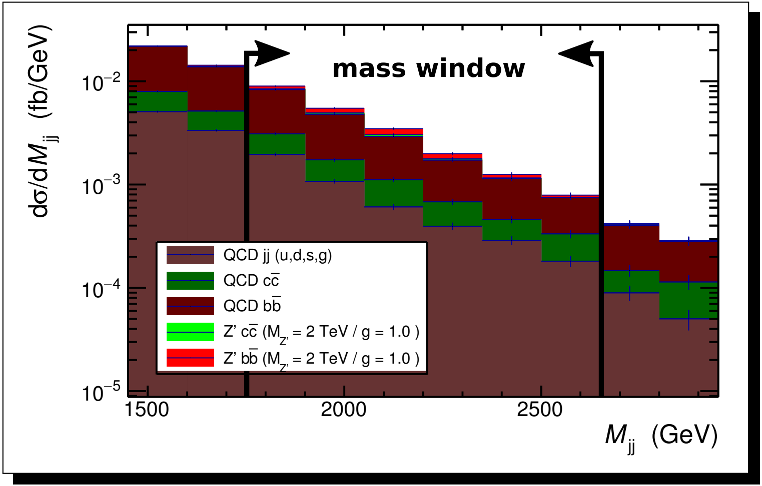
  - **Reject** events with two real tags, otherwise double counting.

$M_{Z'} = 2.5$ TeV $\quad g_B = 1.3$



- Require both jets to be tagged.
- Mass window: $M_{Z'} \times [0.85, 1.25]$
- Narrow jets ($R = 0.4$)
- $\left|\eta_{jet}\right| \leq 2.7, \ \left|\Delta\eta_{jj}\right| \leq 1.5$

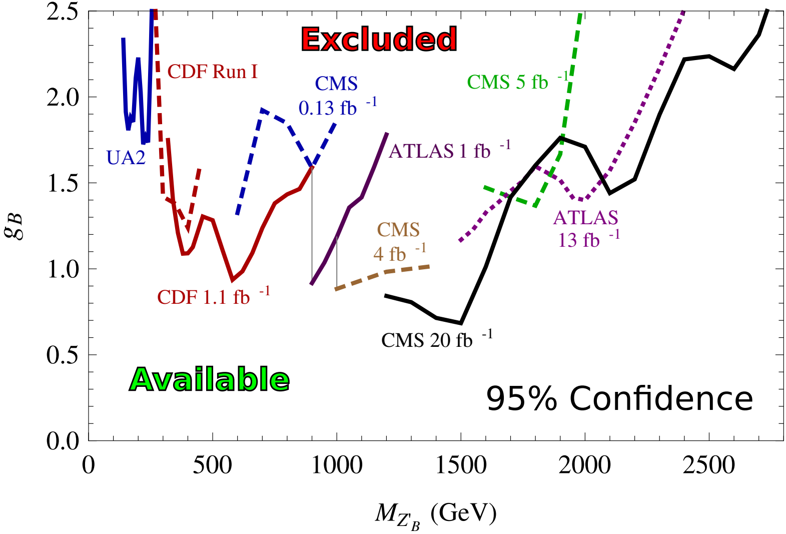
$M_{Z'} = 2$ TeV $\quad g_B = 1$



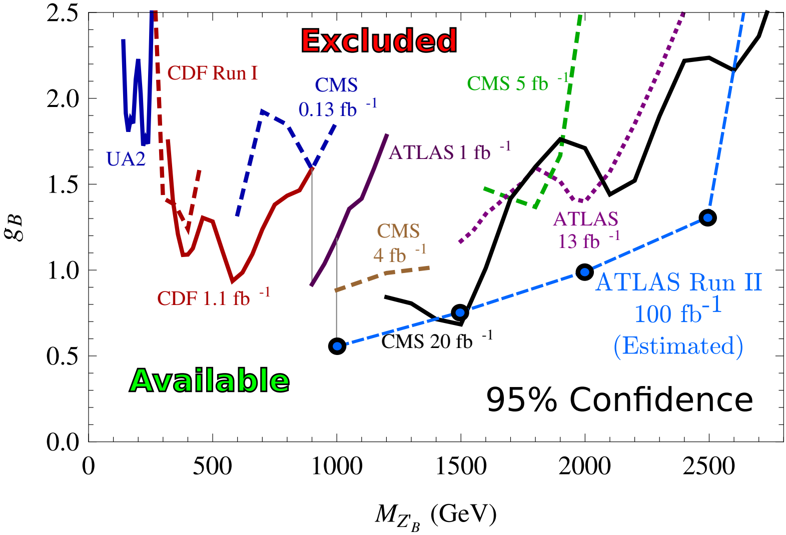
- Require both jets to be tagged.
- Mass window: $M_{Z'} \times [0.85, 1.25]$
- Narrow jets ($R = 0.4$)
- $\left|\eta_{jet}\right| \leq 2.7$, $\left|\Delta\eta_{jj}\right| \leq 1.5$

Uncharted Waters

Dobrescu & Yu [arXiv:1306.2629, fig. 1]



Dobrescu & Yu [arXiv:1306.2629, fig. 1]

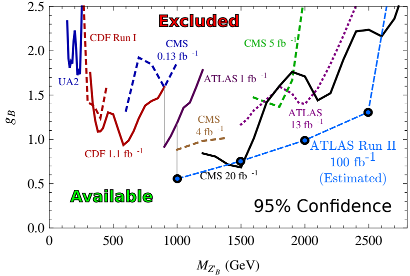


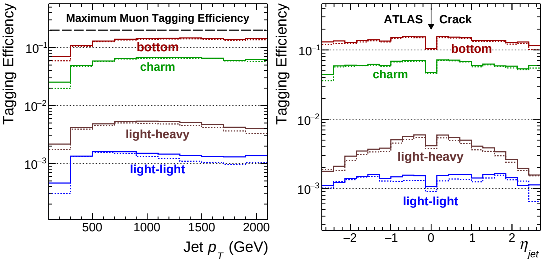
Moving Forward

- Double-tagged exclusion is systematically limited — must detect an integer number of signal events.
  - A *single-tag* could offer a competitive (possibly better) $S/\sqrt{BG}$, with 15 times more signal.
- New paper by Dobrescu and Liu (arXiv: 1506.06736) proposes a Z' in 3.4–4.5 TeV range.
  - Heavy partner of possible W´ seen at 1.8–2 TeV in ATLAS WZ channel.
- Seeing/excluding a bump in all three classes (0-tag / 1-tag / 2-tag) makes the strongest statement.

Conclusion

- μ_X tags heavy flavor jets at the TeV scale.
    - ***b*-jet**: 14%
    - **light-jet**: $< 0.7\%$
- Flat signal response in p_T and η.
- Minimal pileup sensitivity.
- Should offer significant improvements in detecting/excluding Z' and other high energy phenomenon.





$K(x,\kappa)$ and $x_{\max}$

$$\frac{dN}{dx} = 4\pi \, \frac{2x}{(x^2+1)^2} \, K(x,\kappa) \tag{8}$$

where $K(x,\kappa)$ corrects the shape

$$K(x,\kappa) = \frac{(1+\kappa^2) + x^2(1-\kappa^2)}{2\sqrt{1+x^2(1-\kappa^2)}} \, \Theta(x_{\max} - x) \tag{9}$$

We can calculate the maximum emission

$$x_{\max} = 1/\sqrt{\kappa^2 - 1} \approx \sqrt{\gamma_{\mu,\mathrm{cm}}^2 - 1} \tag{10}$$

which, in our regime of interest $\left(\gamma_{\mathrm{B}} \gg \gamma_{\mu,\mathrm{cm}} \gg 1\right)$ is

$$x_{\max} \approx \sqrt{\gamma_{\mu,\mathrm{cm}}^2 - 1} \tag{11}$$