

ATLAS NOTE

August 4, 2015



Trigger studies for LHC Run-2 $W' \rightarrow tb \rightarrow qqbb$ search

D. Wilbern^a and J. Love^b

^aDepartment of Physics, Northern Michigan University, 1401 Presque Isle Ave, Marquette, MI 49855, USA ^bHEP Division, Argonne National Laboratory, 9700 S. Cass, Argonne, IL 60439, USA

Abstract

A preliminary study of usable triggers for the Run-2 $W' \rightarrow tb$ search in the hadronic final state is presented. QCD dijet events generated by Pythia are used to investigate how systematic uncertainty of background estimation is affected by the choice of trigger. Choosing more restrictive triggers will reduce background rates, but will also increase the lower bound on the mass range of W'particles the search is sensitive to. Single jet p_T triggers, H_T triggers, multijet p_T triggers, b-tag triggers, and proposed H_T + b-tag triggers are investigated in this study. For each trigger investigated, the lower bound of the acceptable W' invariant mass range as well as the tagging efficiency ratios are presented.

1 Introduction



Figure 1: $W' \rightarrow tb$ Feynman diagram in the hadronic decay channel.

Many "Beyond the Standard Model" theories, particularly those that involve SU(2) gauge symmetries, predict the existence of new heavy gauage bosons called W' bosons. LHC Run-1 analyses at $\sqrt{s} = 8$ TeV searched for W' decaying into a top quark and a bottom quark¹. These analyses excluded W' bosons in the top quark leptonic decay channel ($W' \rightarrow \ell \nu$) and hadronic decay channel ($W' \rightarrow tb \rightarrow qqbb$) below 1.7 TeV and 1.5 TeV, respectively, at a 95% confidence level[1][2]. $W' \rightarrow tb$ in this note will henceforth refer to the W' search in the all hadronic final state, $W' \rightarrow tb \rightarrow qqbb$. The Feynman diagram for this state is shown in Figure 1.

QCD dijet events generated by Pythia are used in this study to investigate the lower-bound on the acceptable range of background events' m_{tb} . $W' \rightarrow tb$ events are dijet events with a large-*R* jet top-tagged and a high- p_T small-*R* jet b-tagged. QCD multijet events provide the highest background contribution by far: about 99%. Contribution from $qq \rightarrow W^+ \rightarrow t\bar{b}$ can be reduced with a low-end energy cut, and contribution from other QCD multijet events can be reduced with flavor-tagging requirements. These requirements improve the ratio in cross-sections of background events to signal events from $O(10^6)$ to $O(10^1)$ [3]. The ATLAS trigger system will be used to apply the energy requirement, the W' top-tagger[2] is used to tag top jets, and a neural-network based algorithm is used to tag b-jets.

Figure 2 shows data[2] from the Run-1 hadronic state analysis fitted to the expected background distribution. A more restrictive event trigger would not reach its trigger plateau until a higher value of m_{tb} than the one here, resulting in a need to shift the left-edge of the plot to the right, meaning a loss of sensitivity to signals at lower m_{tb} . This study is a preliminary investigation of triggers proposed for use in the Run-2 $W' \rightarrow tb$ search, with the intent of maximizing the m_{tb} range the analysis is sensitive in.

¹ The decays $W'^+ \to t\bar{b}$ and $W'^- \to \bar{t}b$ are both equally considered. For simplicity this document refers to both as $W' \to tb$.



Figure 2: Data from the run 1 W' search fitted to a four-parameter function.

2 Study Approach

2.1 Trigger Turn-on Curves

Trigger efficiency curves, or "turn-on" curves, were made for the triggers being investigated in this study. A trigger's efficiency is determined by dividing a quantity such as leading jet p_T for events passing the trigger by the same quantity for all events. These efficiencies were plotted as a function of leading small-*R* jet p_T , leading large-*R* jet p_T , small-*R* jets' H_T , and large-*R* jets' H_T . These curves can be found in Appendix A. Sample turn-on curves for the three triggers HLT_ht700, HLT_ht850, and HLT_ht1000 are shown in figure 3. The "trigger plateau" region is taken as the region where 95% or more of events pass the trigger. Note that the trigger plateau region begins at higher values on the x-axis for more restrictive triggers. A trigger's behavior is stable in this region, so background contribution can be investigated without introducing bias from the trigger, For this reason we use the trigger plateau region for the efficiency curves to determine the minimum p_T cut for the event categorization routine in Section 2.2. For each trigger we choose the start of the trigger plateau from either the small-*R* jet p_T turn-on curve, whichever has the higher value.

The MC data used in this study was binned into ten samples by p_T of the leading parton at truth level. The events in these samples are weighted on per-sample basis and for some samples a per-event basis also. To generate the trigger turn-on curves, the per-sample weights were applied, but not the per-event weights because the weighting scheme was found to be inappropriate for generating turn-on curves as a function of leading-jet p_T . As a result, the "stegosaurus" shape exhibited by some of the turn-on curves remains as an artifact of this weighting. The x-axis location of the spikes correspond to the p_T binning of the MC samples. Some triggers, in particular two of the b-tag triggers and a multijet trigger, had low statistics and it is probably not appropriate to draw conclusions about them based on the results of this study.



Figure 3: Some sample trigger turn-on curves, see appendix A for more. The trigger plateau starts later on the x-axis for more restrictive triggers.

2.2 Event Categorization and Jet Selection

A $W' \rightarrow tb$ event takes the form of a top-tagged R=1.0 anti-kt jet and a b-tagged high- $p_T R=0.4$ anti-kt jet. We require that jets pass jet cleaning level MediumBad. We also require for the large-R jet $|\eta| < 2.0$ and for the small R jet $|\eta| < 2.5$, and $\Delta R < 2.0$ between the two jets. There is also a minimum p_T requirement that is chosen as the low-end of the trigger plateau region for the trigger being studied, rounded up in multiples of 50 GeV. This p_T requirement is allowed to be in the range 350 GeV - 800 GeV. See Section 2.1 for details. Events without jets that meet these requirements are rejected and not considered in the background estimation.

A data-driven method called the ABCD method is used to estimate the background distribution in the signal region. Events failing the top-tagging requirement and/or the b-tagging requirement are classified as belonging to one of the control regions: A, B, or C. The orthogonality of the tagging requirements then allows the control regions to be used to estimate the contribution of background in the signal region, D. This estimation is calculated as $D = \frac{B \times C}{A}$. Figure 4 details the criteria for the event classification and the algorithm for classifying a dijet event into one of these regions while selecting the representative large-*R* and small-*R* jets.

The actual m_{tb} distribution of events classified in region D is compared to the ABCD method's prediction. Since we are using QCD dijet events, the signal region is dominated by background too, so the two distributions should agree well. These comparisons can be found in Appendix B. The left edge of the X-axis is determined by adding 1200 GeV to the start of the lower trigger leading p_T plateau threshold for that trigger.



Figure 4: Event classification with the ABCD method. All jets must pass jet cleaning, min. p_T and max. $|\eta|$ requirements before being considered for selection.

3 Trigger Study

3.1 Triggers Investigated

This study was performed to investigate the impact of trigger choice on the Run-2 $W' \rightarrow tb$ hadronic final state analysis. Table 1 lists the triggers we selected for this study. In addition to these, we investigated the possibility of an H_T + b-tag trigger. Events that pass an H_T trigger and a less restrictive b-tagging requirement are deemed to pass this H_T + b-tag trigger.

Table 1: List of triggers selected for study.

		X X	
Single jet	H_T	B-tag	Multijet
HLT_j15	HLT_ht400	HLT_j300_bloose	HLT_7j45
HLT_j85	HLT_ht550	HLT_j225_bloose	HLT_6j45
HLT_j400	HLT_ht700	HLT_j175_bmedium	HLT_5j85
HLT_j460	HLT_ht850	HLT_2j65_btight_j85	HLT_4j100
	HLT_ht1000		HLT_3j175

3.2 Lower *m*_{tb} Limit

We generated efficiency curves were produced for each trigger being studied, and then determined the start of the plateau for both of the leading jet p_T turn-on curves. From these, we determined the lowest acceptable m_{tb} for each trigger as (1200 GeV + AK4 plateau start). These values can be found in Table 2.

Trigger	Ak4 plateau start [GeV]	Ak10 plateau start [GeV]	Lowest m_{tb}
No trigger	350	350	1550
HLT_j400	450	650	1650
HLT_j460	500	700	1700
HLT_j85	350	350	1550
HLT_j15	350	350	1550
HLT_ht400	350	350	1550
HLT_ht550	350	350	1550
HLT_ht700	400	450	1600
HLT_ht850	450	500	1650
HLT_ht1000	500	550	1700
HLT_j300_bloose	600	550	1800
HLT_j225_bloose	800	800	2000
HLT_j175_bmedium	400	450	1600
HLT_2j65_btight_j85	350	350	1550
HLT_7j45	350	500	1550
HLT_6j45	350	500	1550
HLT_5j85	550	500	1750
HLT_4j100	550	500	1750
HLT_3j175	650	550	1850
"B-tag"	350	350	1550
"HLT_ht400 + B-tag"	350	350	1550
"HLT_ht550 + B-tag"	350	350	1550
"HLT_ht700 + B-tag"	400	450	1600
"HLT_ht850 + B-tag"	450	500	1650
"HLT_ht1000 + B-tag"	500	500	1700

Table 2: Start of trigger plateaus for leading- p_T turn-on curves, and lowest acceptable m_{tb} .

3.3 Ratio of Tagging Efficiencies

An important part of this study is to identify whether the triggers introduce bias into the top-tagging and b-tagging requirements. If the tagging requirements are not orthogonal, then the ABCD method cannot be used to estimate background contribution in the signal region. In order to investigate how orthogonal the tagging requirements are to each other, we calculated and compared tagging efficiencies across the borders of the different ABCD regions.

We define tagging efficiencies as:

$$\varepsilon_b = \frac{B}{A+B}; \ \varepsilon_t = \frac{C}{A+C}; \ \varepsilon_{bt} = \frac{D}{A+B+C+D}$$

where ε_b is the efficiency to b-tag, ε_t is the efficiency to top-tag, and ε_{bt} is the efficiency to both b-tag and top-tag. *A*, *B*, *C*, and *D* are the total number of events categorized in that region by the ABCD method. These efficiencies are related by:

$$\varepsilon_b \varepsilon_t = f_c \varepsilon_{bt}$$

where f_c is a coefficient that describes the correlation of the tagging efficiencies. It should be close to 1 for uncorrelated tagging requirements.

We also compared the values of efficiency to b-tag across the borders of regions A and B as well as regions C and D. We define:

$$\varepsilon_{b1} = \frac{B}{A+B}; \ \varepsilon_{b2} = \frac{D}{C+D}$$

where ε_{b1} is the efficiency to b-tag across the region A-B border and ε_{b2} is the efficiency to b-tag across the region C-D border. These efficiencies are related by:

$$\varepsilon_{b1} = f_c \varepsilon_{b2}$$

where f_c is a coefficient that describes how correlated the b-tag requirement is to whether or not there was a top-tag. This value should be close to 1 if there is no correlation.

Similarly, we define:

$$\varepsilon_{t1} = \frac{C}{A+C}; \ \varepsilon_{t2} = \frac{D}{B+D}$$

where ε_{t1} the efficiency to top-tag across the region A-C border and ε_{t2} is the efficiency to top-tag across the region B-D border. These efficiencies are related by:

$$\varepsilon_{t1} = f_c \varepsilon_{t2}$$

where again, the correlation coefficient f_c should be close to 1 to indicate no correlation.

Table 3 lists these f_c values when the different triggers are applied. Values that deviate from 1 more than 10% are suffixed with a "!". To get more of a sense for the statistical uncertainty, we also calculate $\frac{\sqrt{N_D}}{N_D}$ where N_D is the number of events in region D, for each trigger and scale to the value for when no trigger is applied.

To ensure that these ratios' closeness to 1 comes from bias in the tagging requirements, we repeated the ABCD method but with both tagging requirements replaced with a simple random 50% chance for each jet to pass or fail. These tagging requirements should definitely be uncorrelated and should result in an f_c value close to 1, so any deviation from 1 can be taken as the fractional uncertainty due to Monte Carlo statistics. These coefficients are given in Table 4.

	Table 5: tagging enciency contention constants for real tagging.					
Trigger	f_c	B-tagging f_c	Top-tagging f_c	Scaled $\frac{\sqrt{N_D}}{N_D}$		
No Trigger	0.975355	0.977014	0.975966	1		
HLT_j400	0.996143	0.996557	0.996231	1.04422		
HLT_j460	0.945653	0.951759	0.946896	1.04977		
HLT_j85	0.975355	0.977014	0.975966	1		
HLT_j15	0.975355	0.977014	0.975966	1		
HLT_ht400	0.975355	0.977014	0.975966	1		
HLT_ht550	0.975355	0.977014	0.975966	1		
HLT_ht700	1.06669	1.06139	1.06508	1.01651		
HLT_ht850	0.996872	0.997148	0.996943	1.0236		
HLT_ht1000	1.05709	1.05172	1.05576	1.0297		
HLT_j300_bloose	1.06917	1.06301	1.06204	1.46776		
HLT_j225_bloose	0.786452!	0.807185!	0.807068!	1.50283		
HLT_j175_bmedium	1.14166!	1.13142!	1.11264!	2.0892		
HLT_2j65_btight_j85	4.06371!	3.20732!	3.10619!	15.6666		
HLT_7j45	0.747129!	0.813747!	0.751114!	4.22229		
HLT_6j45	1.26336!	1.19898!	1.2565!	2.51391		
HLT_5j85	1.33767!	1.23807!	1.33028!	2.54692		
HLT_4j100	1.11289!	1.08839	1.11076!	1.73526		
HLT_3j175	1.10218!	1.08426	1.0999	1.44365		
"B-tagging"	1.41157!	1.37876!	1.33684!	1		
HLT_ht400 + "B-tagging"	1.41157!	1.37876!	1.33684!	1		
HLT_ht550 + "B-tagging"	1.41157!	1.37876!	1.33684!	1		
HLT_ht700 + "B-tagging"	1.51886!	1.47003!	1.43221!	1.01651		
HLT_ht850 + "B-tagging"	1.46616!	1.41568!	1.39414!	1.0236		
HLT_ht1000 + "B-tagging"	1.47468!	1.42333!	1.40097!	1.0236		

Table 3: tagging efficiency correlation constants for real tagging.

	Table 4. tagging enciency correlation constants for random tagging.					
Trigger	f_c	B-tagging f_c	Top-tagging f_c	Scaled $\frac{\sqrt{N_D}}{N_D}$		
No trigger	1.01805	1.00905	1.01349	1		
HLT_j400	1.00086	1.00041	1.00065	1.05469		
HLT_j460	0.988465	0.994564	0.991161	1.06679		
HLT_j85	1.01805	1.00905	1.01349	1		
HLT_j15	1.01805	1.00905	1.01349	1		
HLT_ht400	1.01805	1.00905	1.01349	1		
HLT_ht550	1.01805	1.00905	1.01349	1		
HLT_ht700	1.02997	1.01512	1.02221	1.01638		
HLT_ht850	1.00227	1.00113	1.0017	1.02529		
HLT_ht1000	0.97973	0.989931	0.984729	1.03214		
HLT_j300_bloose	0.967832	0.98426	0.975731	2.56099		
HLT_j225_bloose	0.999675	0.999843	0.999753	2.61512		
HLT_j175_bmedium	1.07208	1.0365	1.05283	4.60289		
HLT_2j65_btight_j85	0.567637!	0.73624!	0.669779!	31.4556		
HLT_7j45	0.86202!	0.927897!	0.89685!	5.73062		
HLT_6j45	1.00967	1.00474	1.00721	3.17457		
HLT_5j85	0.970144	0.985646	0.977206	3.42981		
HLT_4j100	1.01543	1.00726	1.01181	2.14315		
HLT_3j175	1.02132	1.00972	1.01635	1.57934		
"B-tagging"	1.1138!	1.05751	1.083	2.35134		
HLT_ht400 + "B-tagging"	1.1138!	1.05751	1.083	2.35134		
HLT_ht550 + "B-tagging"	1.1138!	1.05751	1.083	2.35134		
HLT_ht700 + "B-tagging"	0.987586	0.993821	0.99065	2.3796		
HLT_ht850 + "B-tagging"	0.944994	0.973187	0.957684	2.39551		
HLT_ht1000 + "B-tagging"	0.94873	0.975094	0.960507	2.39551		

Table 4: tagging efficiency correlation constants for random tagging.

4 Summary and Interpretation

This study is a preliminary investigation of some triggers for use the in the Run-2 $W' \rightarrow tb$ hadronic final state analysis. We are interested in how some different triggers affect the acceptable $W' m_{tb}$ range, and how they introduce bias into the ABCD method. We made a list of some single jet p_T triggers, H_T triggers, b-tag triggers, and multijet p_T triggers that can be found in Table 1. In addition, we investigated the possibility of an H_T + b-tag trigger by requiring an event pass a less restrictive b-tag requirement along with an H_T trigger.

We generated efficiency curves for each trigger as a function of small-*R* jet p_T , large-*R* jet p_T , small-*R* jet H_T , and large-*R* jet H_T . These can be found in Appendix A. We determined the start of the trigger plateau region to be where 99% of events pass the trigger, and used this value to set a p_T requirement on jets in the next part of the study, the ABCD method categorization. Events are categorized into four regions based on whether their large-*R* jet passes the top-tagging requirement and/or their small-*R* jet passes the b-tagging requirement. Events are categorized into one of the control regions (A, B, and C) if they one or both of these requirements, or into the signal region (D) if they pass both requirements. The control regions can then be used to estimate background contribution in the signal region. Comparisons of the m_{tb} spectrum for events in region D to the ABCD method prediction can be found in Appendix B. Table 2 shows the values of p_T where the trigger plateau region begins for both leading small-*R* jet p_T and leading large-*R* jet p_T . Also listed is the determined minimum m_{tb} value where this analysis is sensitive.

Finally, ratios between tagging efficiencies were calculated for events passing each trigger: $\frac{\varepsilon_{b1}}{\varepsilon_b \varepsilon_t}$, $\frac{\varepsilon_{b1}}{\varepsilon_{b2}}$, and $\frac{\varepsilon_{c1}}{\varepsilon_t}$. These efficiency ratios were computed for the tagging requirements used in the $W' \rightarrow tb$ analysis and for randomized tagging requirements where both requirements are simply a 50% chance to tag. The coefficients for the real taggers are presented in Table 3 and for the random taggers in Table 4. These efficiency ratios provide insight into how biased the top-tag and b-tag requirements are when each trigger is applied. The ratios should be very close to 1 if the tagging requirements are unbiased.

As expected, the random tagging requirements show ratios close to 1 for every trigger except for one with low statistics. Since a randomized 50% chance to top-tag and b-tag should be completely unbiased, any deviation from 1 can be taken as the fractional uncertainty due to Monte Carlo statistics for that trigger. Under the actual top-tagging and b-tagging requirements:

Single jet p_T triggers and H_T triggers

The ratios are all only a few percent off from 1, indicating that they do not introduce bias into the tagging requirements and are suitable for use in the $W' \rightarrow tb$ Run-2 analysis.

B-tag triggers

HLT_j300_bloose has reasonably high statistics and shows ratios only a few percent off from 1, making it worthy of further investigation. HLT_j300_bloose's ratios indicate that its application introduces a bias of about 20% into the tagging methods, making it unsuitable for use with the ABCD method. The HLT_j175_bmedium and HLT_2j65_btight_j85 triggers both have high statistical uncertainty, so conclusions can't be drawn about them from this study.

Multijet triggers

HLT_7j45 has too high statistical uncertainty to draw conclusions about it from this study. The rest of the multijet triggers in general show efficiency ratios far departed from 1, indicating that they introduce bias into the tagging methods.

Proposed H_T + loose B-tag triggers

Efficiency ratios far from 1 indicate bias in the tagging ratios when these H_T triggers and a loose b-tagging requirement are imposed on events.

It should be noted that the turn-on curves for the multijet triggers and H_T + loose b-tagging proposed triggers are disfigured by the weighting scheme of the leading-parton- p_T binned Monte Carlo samples. The start of the trigger plateau region may have therefore been incorrectly determined for these triggers, resulting in bias being introduced from the instability of the trigger outside of the plateau region. Further study is needed to determine the effect of the weighting on the trigger turn-on curve.

This study only examined QCD dijet background MC samples. In the near future, these triggers will be studied with W' signal MC samples as well.

Acknowledgements

This work was supported in part by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the Science Undergraduate Laboratory Internships Program (SULI).

Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357.

A Trigger Efficiency Curves



Figure A.1: HLT_j400 efficiency curves.



Figure A.2: HLT_j460 efficiency curves.



Figure A.3: HLT_j85 efficiency curves.



Figure A.4: HLT_j15 efficiency curves.



Figure A.5: HLT_ht400 efficiency curves.



Figure A.6: HLT_ht550 efficiency curves.



Figure A.7: HLT_ht700 efficiency curves.



Figure A.8: HLT_ht850 efficiency curves.



Figure A.9: HLT_ht1000 efficiency curves.



Figure A.10: HLT_j300_bloose efficiency curves.



Figure A.11: HLT_j225_bloose efficiency curves.



Figure A.12: HLT_j175_bmedium efficiency curves.



Figure A.13: HLT_7j45 efficiency curves.



Figure A.14: HLT_6j45 efficiency curves.



Figure A.15: HLT_5j85 efficiency curves.



Figure A.16: HLT_4j100 efficiency curves.



Figure A.17: HLT_3j175 efficiency curves.



Figure A.18: "B-tag" efficiency curves.



Figure A.19: HLT_ht400 + "B-tagging" efficiency curves.



Figure A.20: HLT_ht550 + "B-tagging" efficiency curves.



Figure A.21: HLT_ht700 + "B-tagging" efficiency curves.



Figure A.22: HLT_ht850 + "B-tagging" efficiency curves.



Figure A.23: HLT_ht1000 + "B-tagging" efficiency curves.

B *m*_{tb} **Distributions, Actual vs. ABCD Method Prediction**



Figure B.1: HLT_j400 *m*_{tb} distributions.



Figure B.2: HLT_j460 m_{tb} distributions.



Figure B.3: HLT_j85 m_{tb} distributions.



Figure B.4: HLT_j15 m_{tb} distributions.



Figure B.5: HLT_ht400 *m*_{tb} distributions.



Figure B.6: HLT_ht550 m_{tb} distributions.



Figure B.7: HLT_ht700 m_{tb} distributions.



Figure B.8: HLT_ht850 m_{tb} distributions.



Figure B.9: HLT_ht1000 *m*_{tb} distributions.



Figure B.10: HLT_j300_bloose m_{tb} distributions.



Figure B.11: HLT_j225_bloose m_{tb} distributions.



Figure B.12: HLT_j175_bmedium m_{tb} distributions.



Figure B.13: HLT_2j65_btight_j85 m_{tb} distributions.



Figure B.14: HLT_7j45 m_{tb} distributions.



Figure B.15: HLT_6j45 m_{tb} distributions.



Figure B.16: HLT_5j85 m_{tb} distributions.



Figure B.17: HLT_4j100 m_{tb} distributions.



Figure B.18: HLT_3j175 m_{tb} distributions.



Figure B.19: "B-tag" m_{tb} distributions.



Figure B.20: "HLT_ht400 + B-tag" m_{tb} distributions.



Figure B.21: "HLT_ht550 + B-tag" m_{tb} distributions.



Figure B.22: "HLT_ht700 + B-tag" m_{tb} distributions.



Figure B.23: "HLT_ht850 + B-tag" m_{tb} distributions.



Figure B.24: "HLT_ht1000 + B-tag" m_{tb} distributions.

References

- [1] ATLAS Collaboration, G. Aad et al., Search for W' → tb̄ in the lepton plus jets final state in proton-proton collisions at a centre-of-mass energy of √s = 8 TeV with the ATLAS Detector, Physics Letters B 743 (2015) 235 255. http://dx.doi.org/10.1016/j.physletb.2015.02.051.
- [2] ATLAS Collaboration, G. Aad et al., Search for $W' \rightarrow tb \rightarrow qqbb$ Decays in pp Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector, The European Physical Journal C **75** (2015) no. 4, 1434. http://dx.doi.org/10.1140/epjc/s10052-015-3372-2.
- [3] J. Love, Searching for Physics Beyond the Standard Model in the 3rd Generation with the ATLAS Detector, Presentation, Mar, 2015. https://indico.hep.anl.gov/indico/conferenceDisplay.py?confId=586.