1b2b Tagger: b-tagged Jet Identification

PIKIMO 2023

Beni Pazár

Indiana University

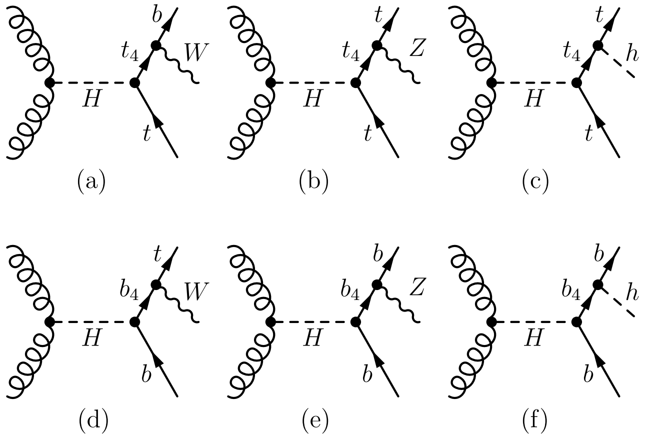
11/11/2023

In collaboration with E. Lunghi

- As everyone knows, the Standard Model is incredibly successful
- However, we know that it is still incomplete
- Extra Higgs doublets and vector-like fermions are two straight forward extensions currently being searched for at the LHC
- The current searches look at the two model extensions independently. If they exist simultaneously then cascade decays ($H \rightarrow VLQ$ or $VLQ \rightarrow H$) dominate.
- With this in mind, current searches may be looking for non-ideal signals.

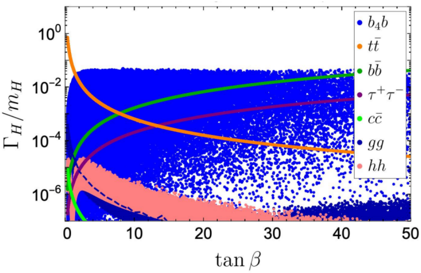
Cascade Decays

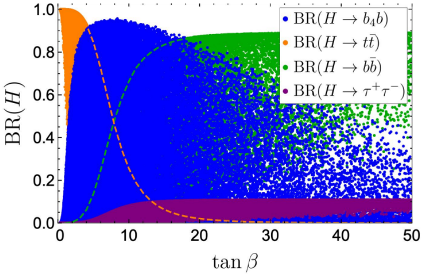
When the heavy Higgs' mass is greater than the vector-like quarks, these processes are common. I focused on $H \to b_4 b \to Zbb$



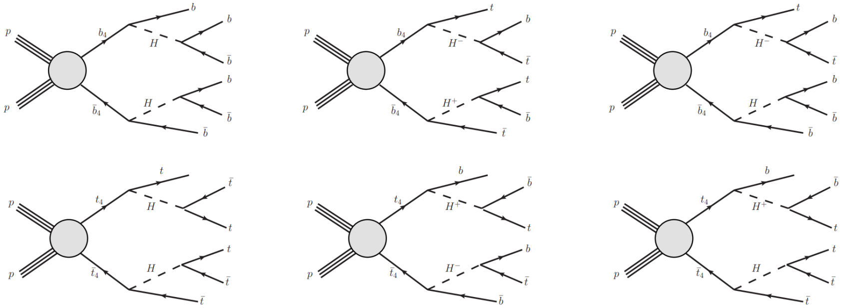
Why focus on b-quarks?

Decays into $b_4 b$ and $b\bar{b}$ dominate





However, if the vector-like quark instead decays into the heavy Higgs then processes such as these can occur. In both cases, proper identification of the final state b-quarks is paramount.



Standard b-tagging methodology

- ATLAS and CMS both have very robust tagging algorithms. They search for a *b*-quark in the jet (arXiv: 1907.05120, 1211.4462)
- This does not give any information about the particle initiating the jet
- For the process studied $g \to b\bar{b}$ is a dominant background
- So we need a tagger that distinguishes between jets which originated from a b-quark or gluon. We call this the 1b2b-tagger

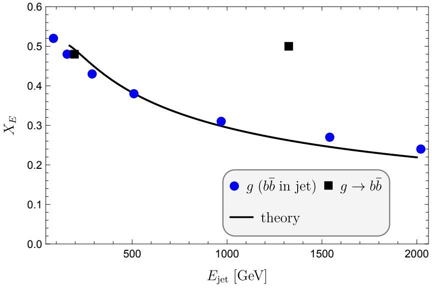
Variables of interest

Our work expanded upon that done by Goncalves, Krauss, and Linten (arXiv: 1512.05265)

- Charged Track Multiplicity (N_{ch})
  Total number of charged particles in the final state
  Tend to be fewer in b-quark compared to gluon initiated jets
- Girth (g)
  A measure of the spread of jet constituents
  Tends to be smaller for b-quark compared to gluon initiated jets
- Energy Fragmentation Fraction (X_E)
  Ratio of the energy of the tagged b-hadron to the total jet energy
  Tends to be closer to unity for b-quark compared to gluon initiated jets

Gluon Energy Fragmentation

- When a gluon is forced to split into a $b\bar{b}$ pair, each quark carries roughly 50% of the jet energy (Black points)
- However, when the gluon is allowed to radiate before splitting, the energy of each quark is significantly less than the jet energy (Blue points)
- The amount of collinear gluon radiation prior to splitting is governed by the initial energy of the gluon[1]

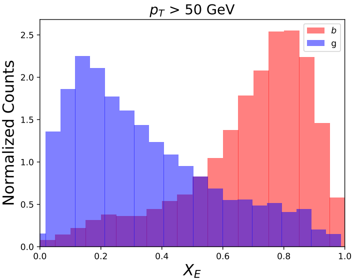


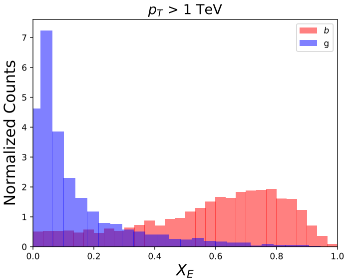
[1] The theory equation came from Ellis, Stirling, and Webber, *QCD and Collider Physics*

XE for different energy scales

As the jet p_T increases the X_E of the gluon initiated jets decrease significantly, while be b-quark initiated jets only shift slightly

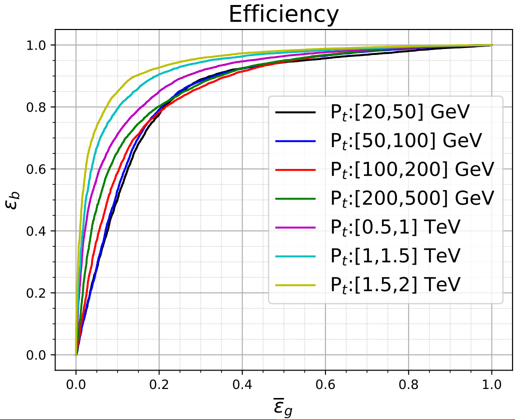
The background also includes charm and light jets which were misidentified. They further reduce X_E



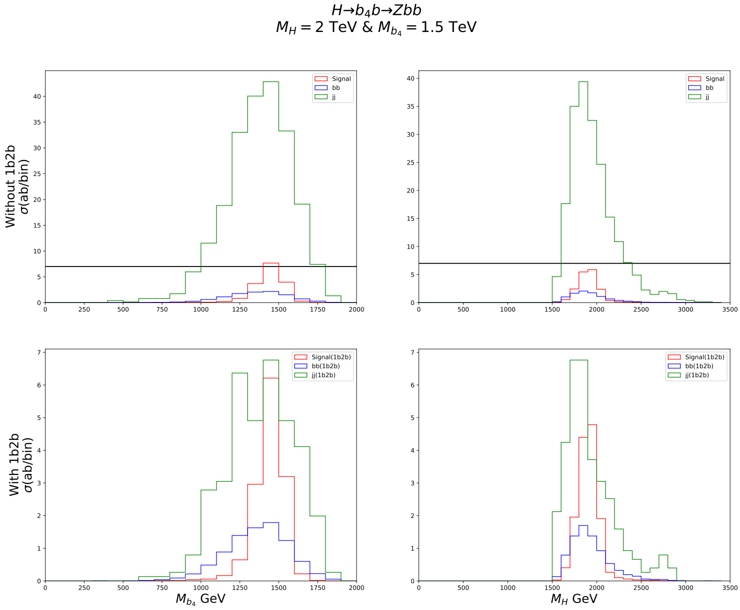


Tagger Efficiency

Combining all three parameters into a logistic regression yields the following efficiency curves for the different p_T ranges we investigated. ϵ_b represents the fraction of b-quark initiated jets which pass the 1b2b tagger, while $\bar{\epsilon}_g$ is the fraction of gluon initiated jets which are mistagged.

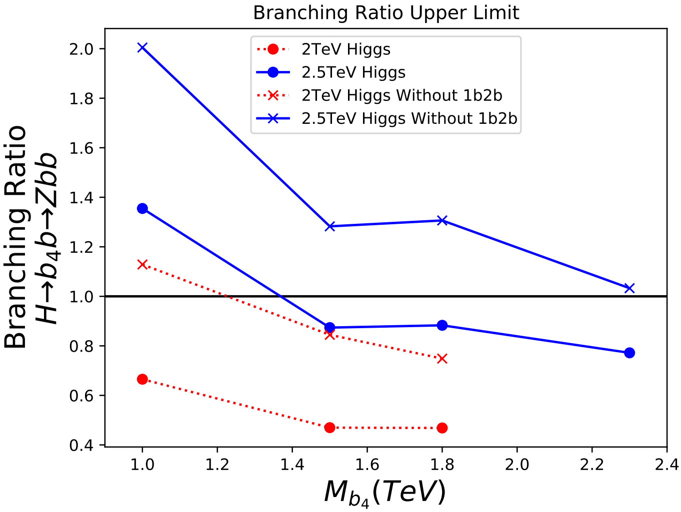


1b2b Tagger Impact



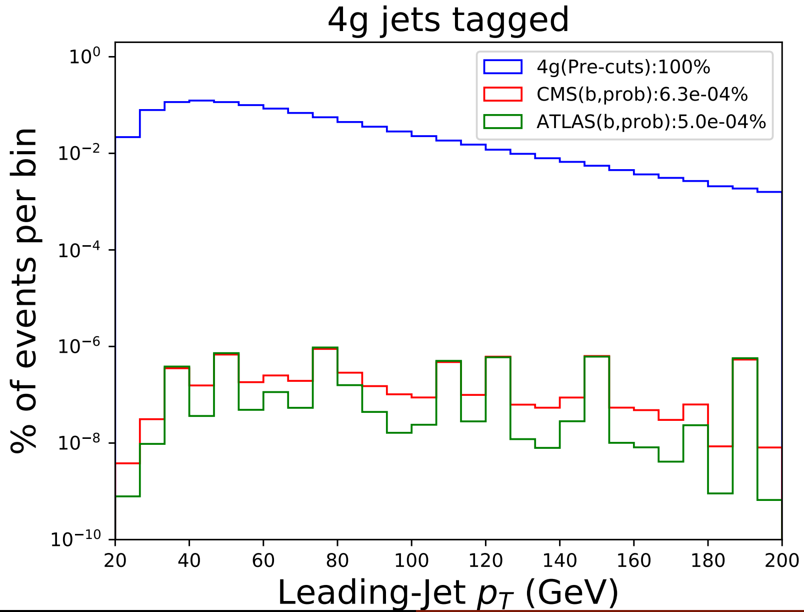
Branching Ratio Impact

Using the 1b2b tagger, we were able to reduce the upper limit of the detectable branching ratio for the cascade decay.

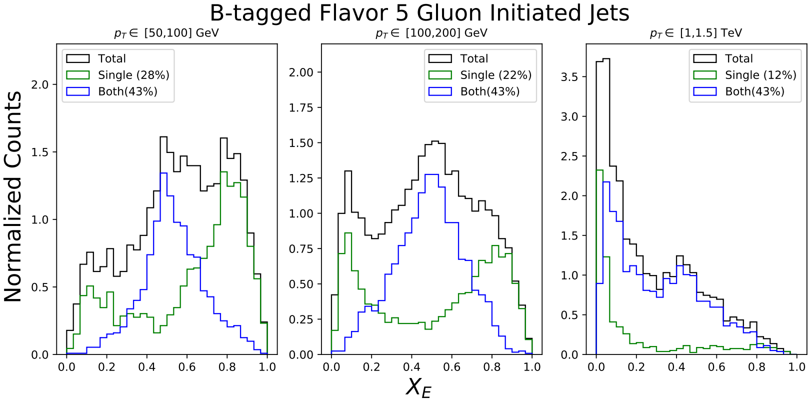


Conclusion

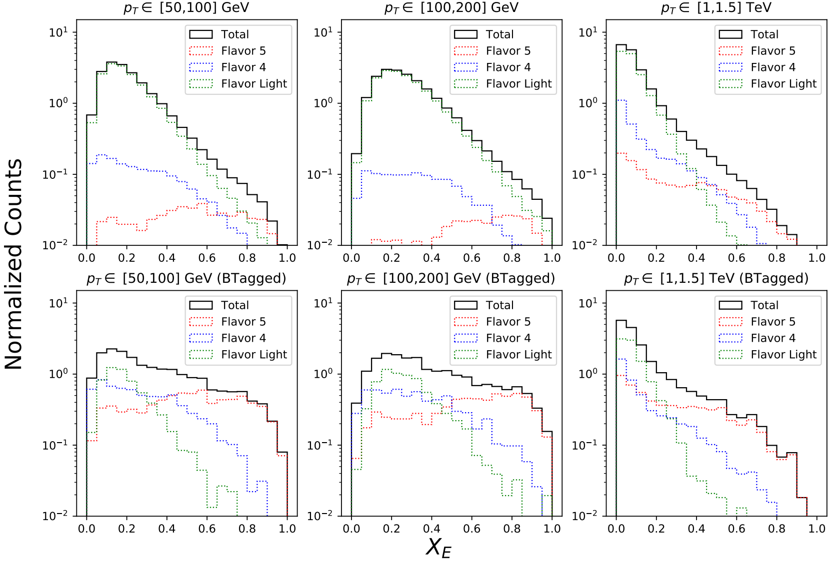
- Distinguishing between b-quark and gluon initiated jets using jet substructure parameters is import for both QCD and new physics searches.
- The 1b2b-tagger presented is able to retain approximately 80% of b-quark initiated jets while rejecting 79-93% of gluon jets
- In the cascade decay process presented, viable searches would be impossible without the 1b2b-tagger
- This approach to jet identification should be further improved and applied in future experimental analysis


4g jets tagged
4g(Pre-cuts):100%
CMS(b,prob):6.3e-04%
ATLAS(b,prob):5.0e-04%
% of events per bin
Leading-Jet p_T (GeV)
10^{0}
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
20
40
60
80
100
120
140
160
180
200

p_T	Btag	f	$Btag_f$	$Btag_f/Btag$	prob jet_f	Btag prob of jet_f
20-50 GeV	0.023	b	0.0062	0.27	0.012	0.52
		c	0.0062	0.27	0.048	0.13
		l	0.010	0.46	0.94	0.011
50-100 GeV	0.037	b	0.013	0.34	0.019	0.68
		c	0.013	0.34	0.061	0.21
		l	0.012	0.32	0.92	0.013
100-200 GeV	0.054	b	0.020	0.37	0.035	0.57
		c	0.018	0.34	0.085	0.21
		l	0.016	0.29	0.88	0.018
200-500 GeV	0.062	b	0.025	0.40	0.04	0.63
		c	0.019	0.31	0.10	0.19
		l	0.018	0.29	0.86	0.021
0.5-1 TeV	0.072	b	0.025	0.35	0.052	0.48
		c	0.018	0.25	0.13	0.14
		l	0.029	0.40	0.82	0.035
1-1.5 TeV	0.081	b	0.019	0.24	0.07	0.27
		c	0.017	0.21	0.16	0.11
		l	0.045	0.55	0.77	0.058
1.5-2 TeV	0.089	b	0.015	0.19	0.077	0.19
		c	0.014	0.20	0.18	0.078
		l	0.059	0.61	0.75	0.079

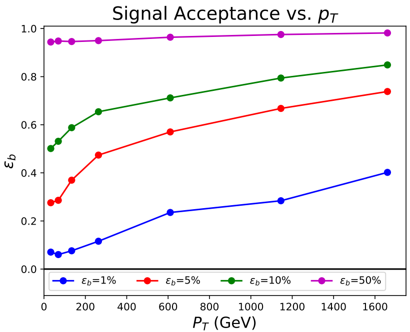

B-tagged Flavor 5 Gluon Initiated Jets
$p_T \in$ [50,100] GeV
$p_T \in$ [100,200] GeV
$p_T \in$ [1,1.5] TeV
Total
Single (28%)
Both(43%)
Single (22%)
Both(43%)
Single (12%)
Both(43%)
Normalized Counts
X_E

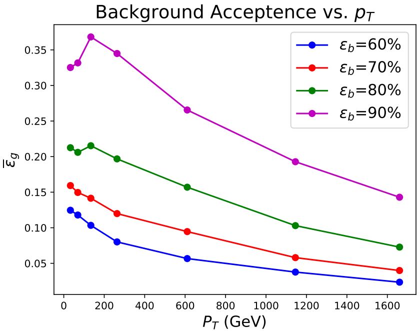
Flavor Breakdown of Gluon Initiated Jets

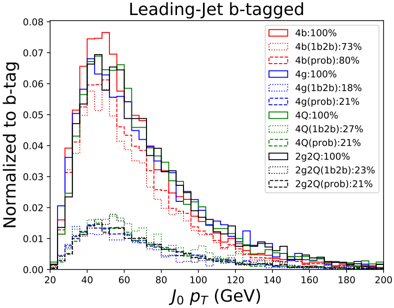


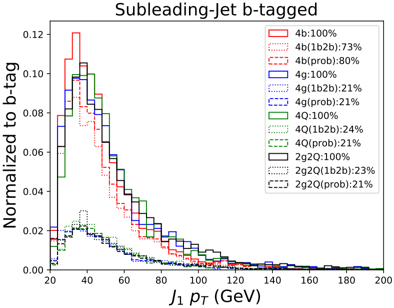
p_T range	p_T^{avg}	$\varepsilon_b = 0.6$	$\varepsilon_b = 0.7$	$\varepsilon_b = 0.8$	$\varepsilon_b = 0.9$	α_{ch}	α_g
20–50 GeV	33 GeV	0.817	0.766	0.699	0.584	0.00581	0.416
50–100 GeV	69 GeV	0.747	0.692	0.615	0.488	0.00309	-0.038
100–200 GeV	134 GeV	0.652	0.581	0.483	0.329	-0.00061	-0.463
200–500 GeV	263 GeV	0.551	0.469	0.356	0.182	-0.00466	-0.750
0.5–1 TeV	0.61 TeV	0.482	0.391	0.269	0.114	-0.0059	-0.084
1–1.5 TeV	1.15 TeV	0.490	0.404	0.287	0.126	-0.00397	0.309
1.5–2 TeV	1.66 TeV	0.481	0.389	0.275	0.121	-0.00363	0.609

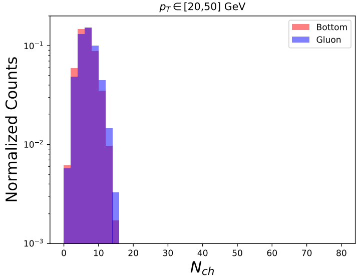
p_T range	$\varepsilon_b = 0.6$	$\varepsilon_b = 0.7$	$\varepsilon_b = 0.8$	$\varepsilon_b = 0.9$
20–50 GeV	0.123	0.16	0.212	0.321
50–100 GeV	0.117	0.15	0.206	0.332
100–200 GeV	0.104	0.142	0.216	0.365
200–500 GeV	0.08	0.119	0.195	0.348
0.5–1 TeV	0.057	0.094	0.157	0.265
1–1.5 TeV	0.038	0.058	0.102	0.193
1.5–2 TeV	0.024	0.041	0.072	0.143

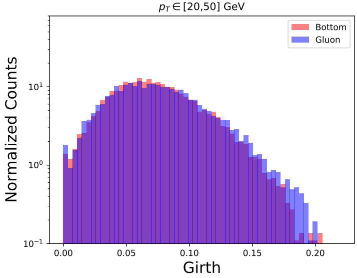

Signal Acceptance vs. p_T
ε_b
P_T (GeV)
ε_b=1%
ε_b=5%
ε_b=10%
ε_b=50%

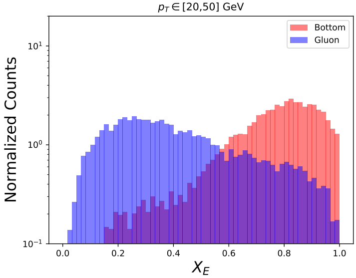

Background Acceptence vs. p_T
$\overline{\varepsilon}_g$
P_T (GeV)
ε_b=60%
ε_b=70%
ε_b=80%
ε_b=90%

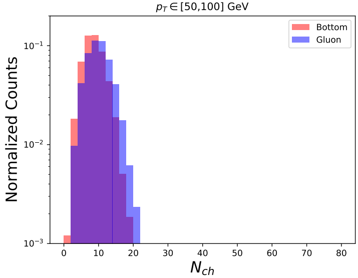

Leading-Jet b-tagged
Normalized to b-tag
J_0 p_T (GeV)
4b:100%
4b(1b2b):73%
4b(prob):80%
4g:100%
4g(1b2b):18%
4g(prob):21%
4Q:100%
4Q(1b2b):27%
4Q(prob):21%
2g2Q:100%
2g2Q(1b2b):23%
2g2Q(prob):21%


Subleading-Jet b-tagged
Normalized to b-tag
J_1 p_T (GeV)
4b:100%
4b(1b2b):73%
4b(prob):80%
4g:100%
4g(1b2b):21%
4g(prob):21%
4Q:100%
4Q(1b2b):24%
4Q(prob):21%
2g2Q:100%
2g2Q(1b2b):23%
2g2Q(prob):21%

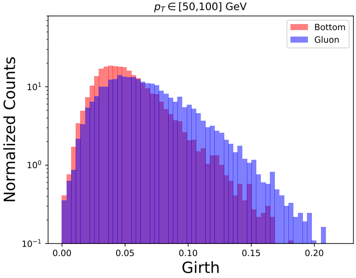

$p_T \in [20,50]$ GeV
Bottom
Gluon
Normalized Counts
N_{ch}


$p_T \in [20,50]$ GeV
Bottom
Gluon
Normalized Counts
Girth

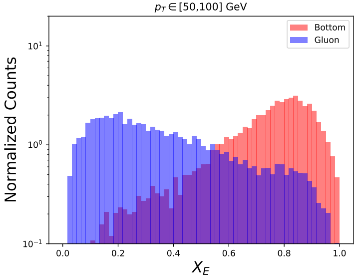

$p_T \in [20,50]$ GeV
Bottom
Gluon
Normalized Counts
X_E



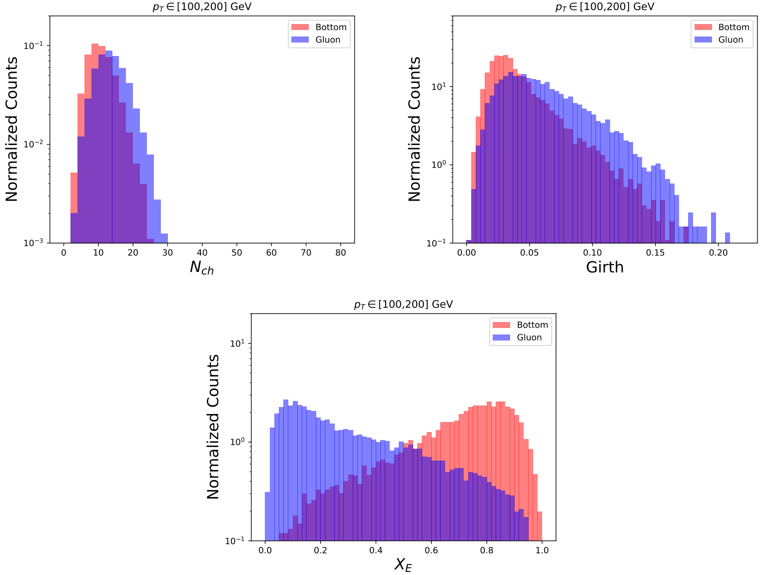
$p_T \in [50,100]$ GeV
Bottom
Gluon
Normalized Counts
N_{ch}

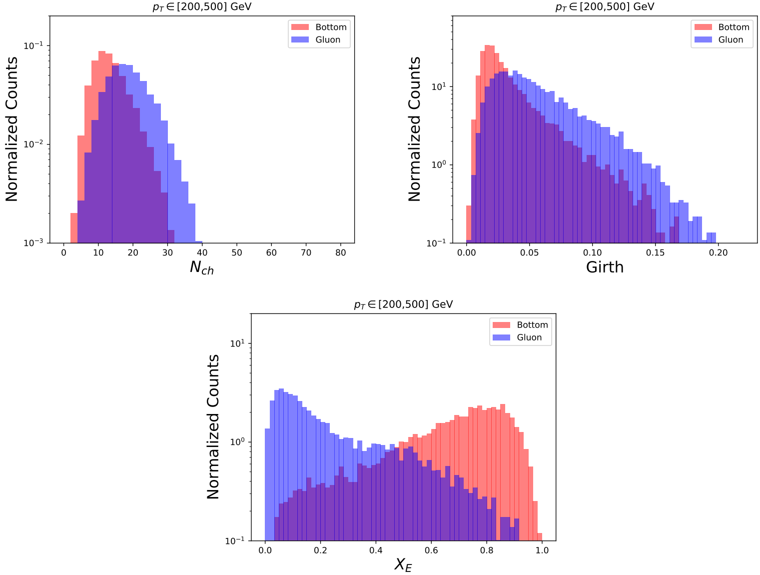


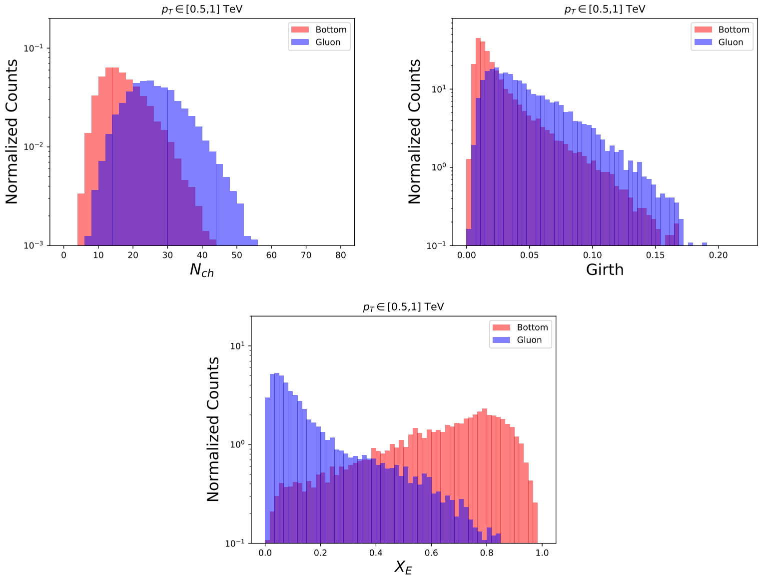
$p_T \in [50,100]$ GeV
Bottom
Gluon
Normalized Counts
Girth



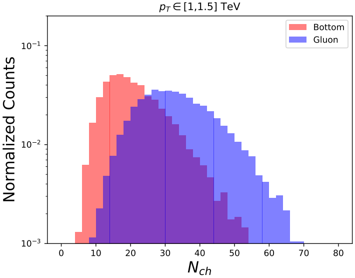
$p_T \in [50,100]$ GeV
Bottom
Gluon
Normalized Counts
X_E

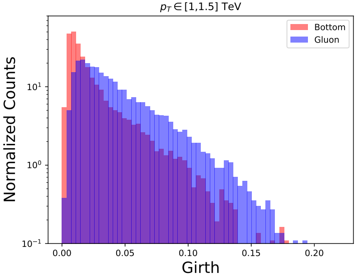

$p_T \in [100,200]$ GeV
Normalized Counts
N_{ch}
Bottom
Gluon
$p_T \in [100,200]$ GeV
Normalized Counts
Girth
Bottom
Gluon
$p_T \in [100,200]$ GeV
Normalized Counts
X_E
Bottom
Gluon

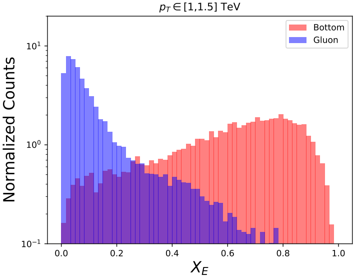

$p_T \in [200,500]$ GeV
Bottom
Gluon
Normalized Counts
N_{ch}
$p_T \in [200,500]$ GeV
Bottom
Gluon
Normalized Counts
Girth
$p_T \in [200,500]$ GeV
Bottom
Gluon
Normalized Counts
X_E

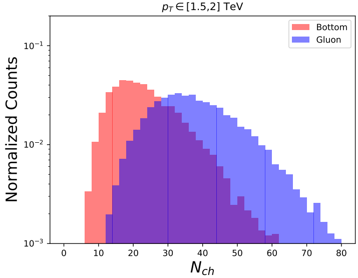


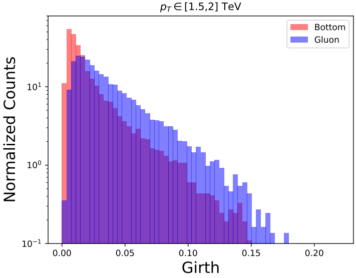
$p_T \in [0.5,1]$ TeV
Bottom
Gluon
Normalized Counts
N_{ch}
$p_T \in [0.5,1]$ TeV
Bottom
Gluon
Normalized Counts
Girth
$p_T \in [0.5,1]$ TeV
Bottom
Gluon
Normalized Counts
X_E

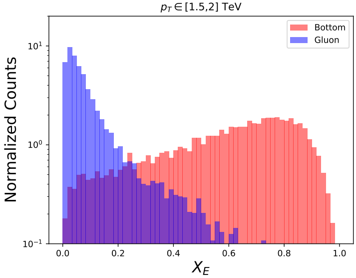

$p_T \in [1,1.5]$ TeV
Bottom
Gluon
Normalized Counts
10^{-1}
10^{-2}
10^{-3}
0
10
20
30
40
50
60
70
80
N_{ch}

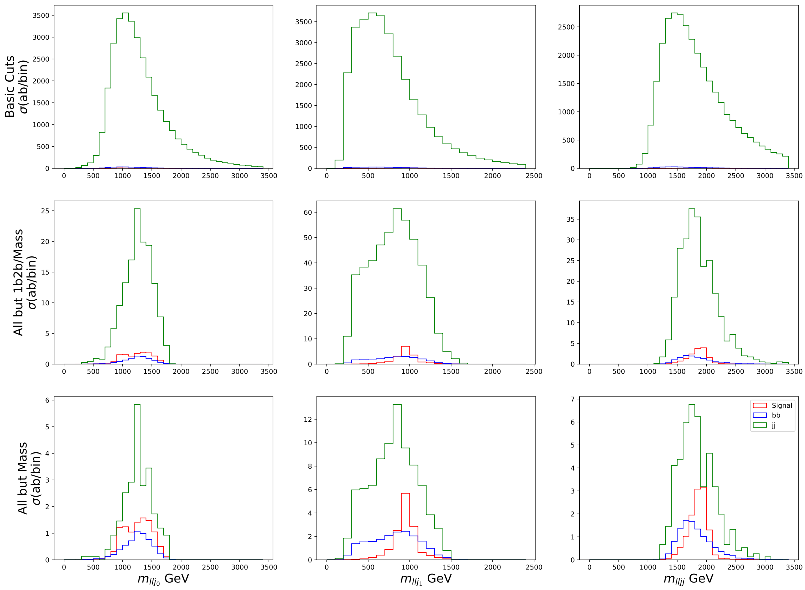

$p_T \in [1,1.5]$ TeV
Bottom
Gluon
Normalized Counts
10^{1}
10^{0}
10^{-1}
0.00
0.05
0.10
0.15
0.20
Girth


$p_T \in [1,1.5]$ TeV
Bottom
Gluon
Normalized Counts
10^{1}
10^{0}
10^{-1}
0.0
0.2
0.4
0.6
0.8
1.0
X_E

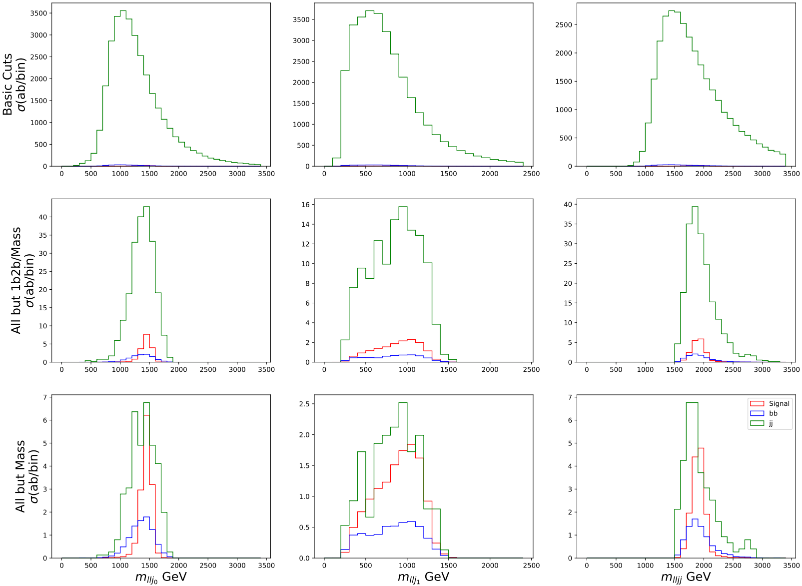

$p_T \in [1.5,2]$ TeV
Bottom
Gluon
Normalized Counts
10^{-1}
10^{-2}
10^{-3}
0
10
20
30
40
50
60
70
80
N_{ch}

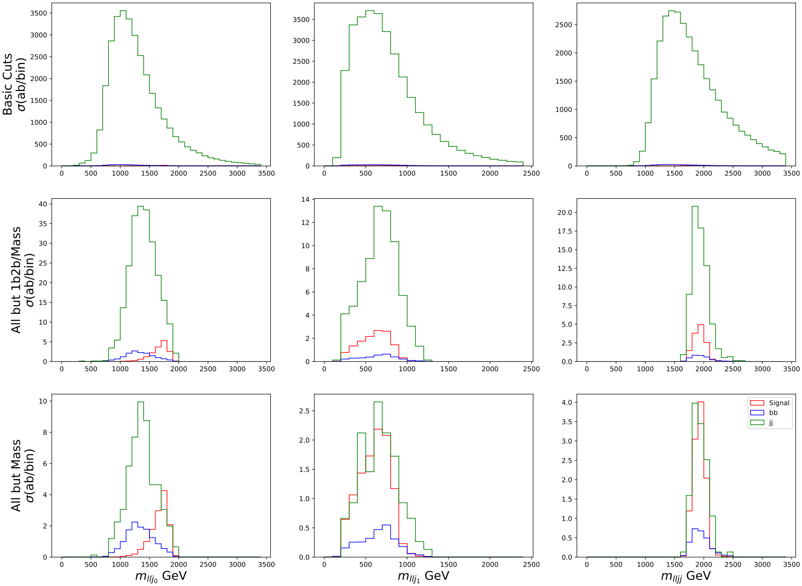

$p_T \in [1.5,2]$ TeV
Bottom
Gluon
Normalized Counts
10^{1}
10^{0}
10^{-1}
0.00
0.05
0.10
0.15
0.20
Girth


$p_T \in [1.5,2]$ TeV
Bottom
Gluon
Normalized Counts
10^{1}
10^{0}
10^{-1}
0.0
0.2
0.4
0.6
0.8
1.0
X_E


Basic Cuts
σ(ab/bin)
All but 1b2b/Mass
σ(ab/bin)
All but Mass
σ(ab/bin)
m_{llj_0} GeV
m_{llj_1} GeV
m_{lljj} GeV
Signal
bb
jj

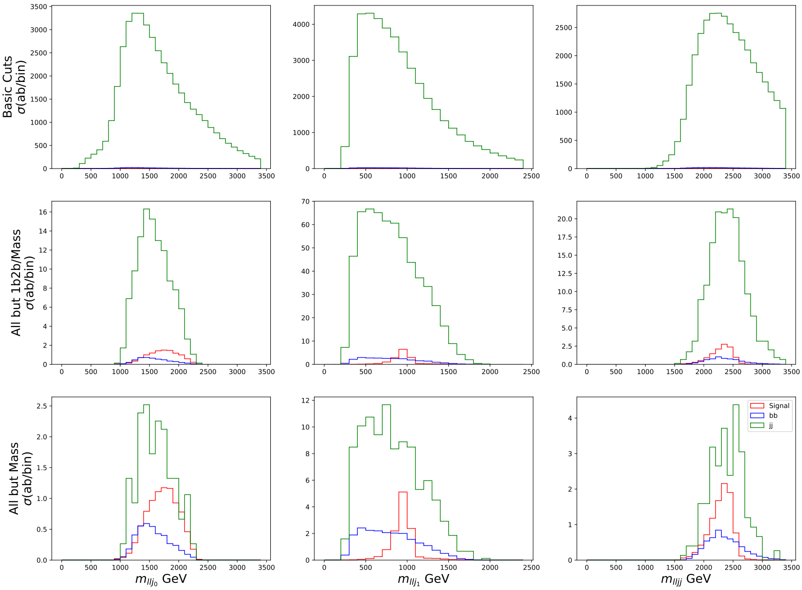
2TeVHiggs-1.5VLQ||Invariant Mass Plots



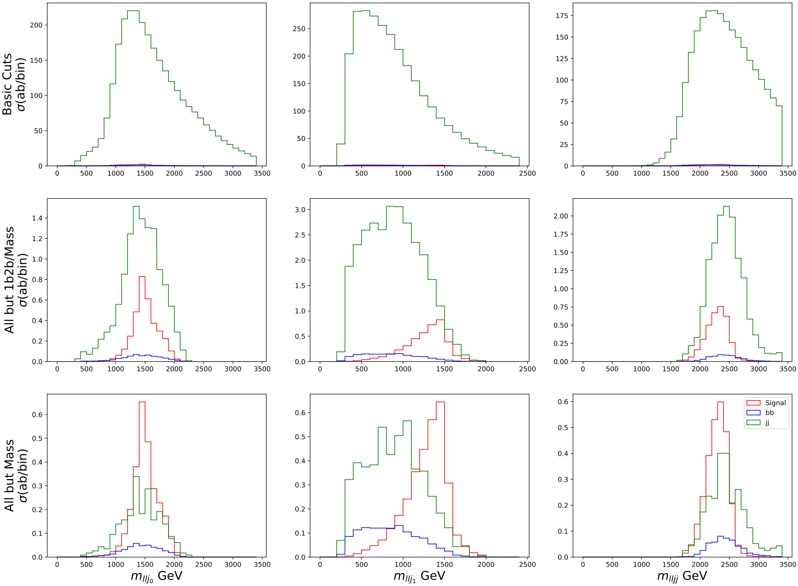


Basic Cuts
σ(ab/bin)
All but 1b2b/Mass
σ(ab/bin)
All but Mass
σ(ab/bin)
m_{llj_0} GeV
m_{llj_1} GeV
m_{lljj} GeV
Signal
bb
jj

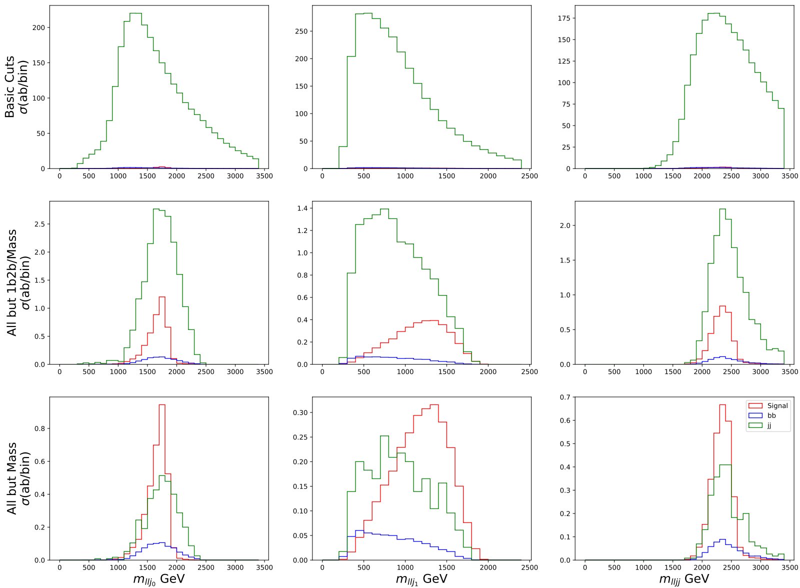
2TeVHiggs-1.0VLQ||Invariant Mass Plots



2.5TeVHiggs-1.5VLQ||Invariant Mass Plots



2.5TeVHiggs-1.8VLQ||Invariant Mass Plots



2.5TeVHiggs-2.3VLQ||Invariant Mass Plots