Boosted Top Tagging through Flavour-violating interactions at the LHC

Shreecheta Chowdhury

Reference : https://arxiv.org/abs/2310.10763v1
Authors : Shreecheta Chowdhury(SRM University-AP), Dr. Amit Chakraborty (SRM University-AP), Dr. Saunak Dutta (ATLAS SkillTech University)

Phoenix 2023

Why Top-tagging ??

- Opportunity to study "bare" quark: Top quark is the heaviest known particle in SM,

$$\textbf{lifetime } (10^{-25}\ \text{s}) < \textbf{Hadronisation time } (10^{-24}\ \text{s})$$

- Top taggers address **leptonic or Hadronic** decay modes of top……

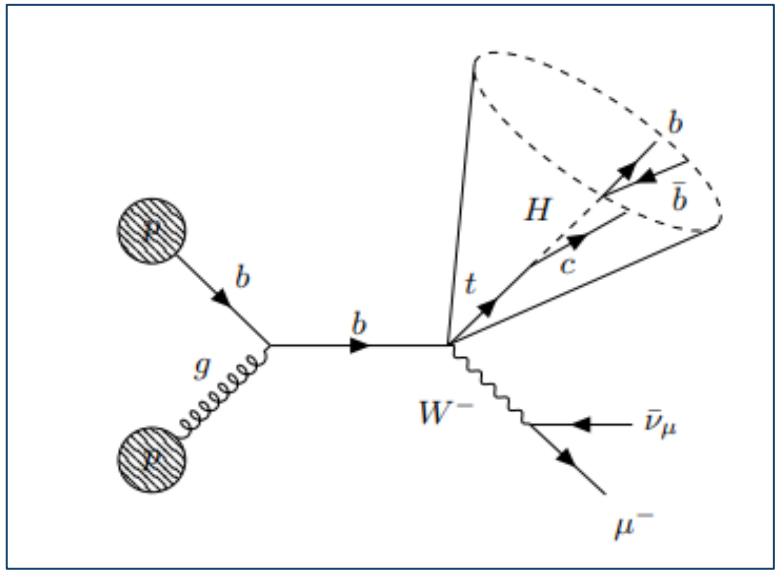
- leptonic decay tagging example: one of the recent works entitled, Tagging a boosted top quark with a τ final state [Phys.Rev.D 108 (2023) 3, 035011], tagging top in t→ b W, W → $\tau\vartheta_\tau$ mode.
- Another one Boosted top quark tagging and polarization measurement using machine learning[Phys. Rev. D, 105(4):042005, 2022].

- Hadronic decay tagging example: JH Top-Tagger [Phys. Rev. Lett. 101, 142001 (2008)] , HepTopTagger[*JHEP* 10 (2010) 078] … works for t→ b W, W having a hadronic decay mode.

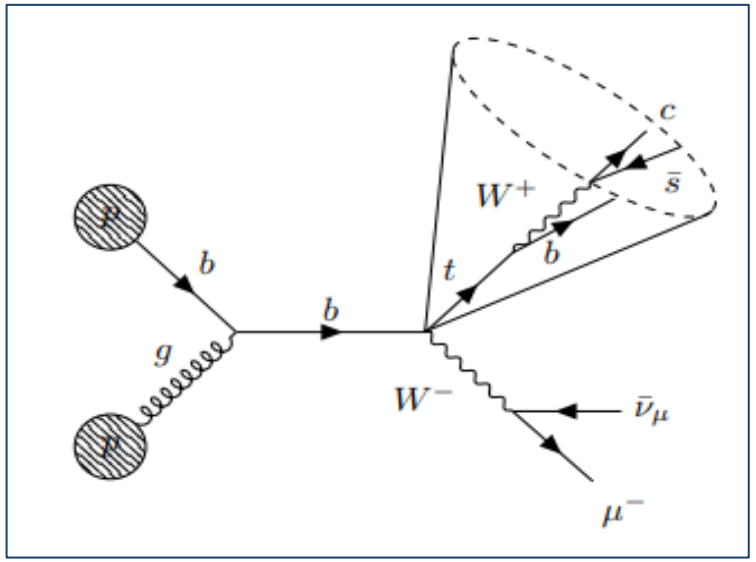
- Top can have **FCNC** decay modes too, like t→ c H, there these taggers may fail!!

Processes considered:

Sample type	Process simulated	Generation level cuts
Signal	$pp \to tW^-,\ t \to cH,\ W^- \to \mu^- \bar{\nu}_\mu, H \to b\bar{b}$	$p_{T,min}^{top} = 350$ GeV
Background 1	$pp \to tW^-,\ t \to bW^+,\ W^- \to \mu^- \bar{\nu}_\mu, W^+ \to c\bar{s}$	$p_{T,min}^{top} = 350$ GeV
Background 2	$pp \to jj$	$p_{T,min}^{j} = 350$ GeV
Background 3	$pp \to jj$	$p_{T,min}^{j} = 500$ GeV
Background 4	$pp \to gW^{\pm},\ W^{\pm} \to \mu^{\pm}\nu_\mu$	$p_{T,min}^{gluon} = 350$ GeV

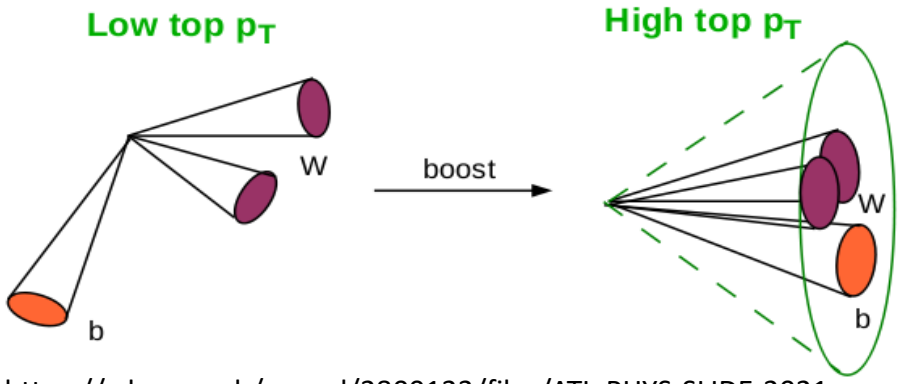


Signal

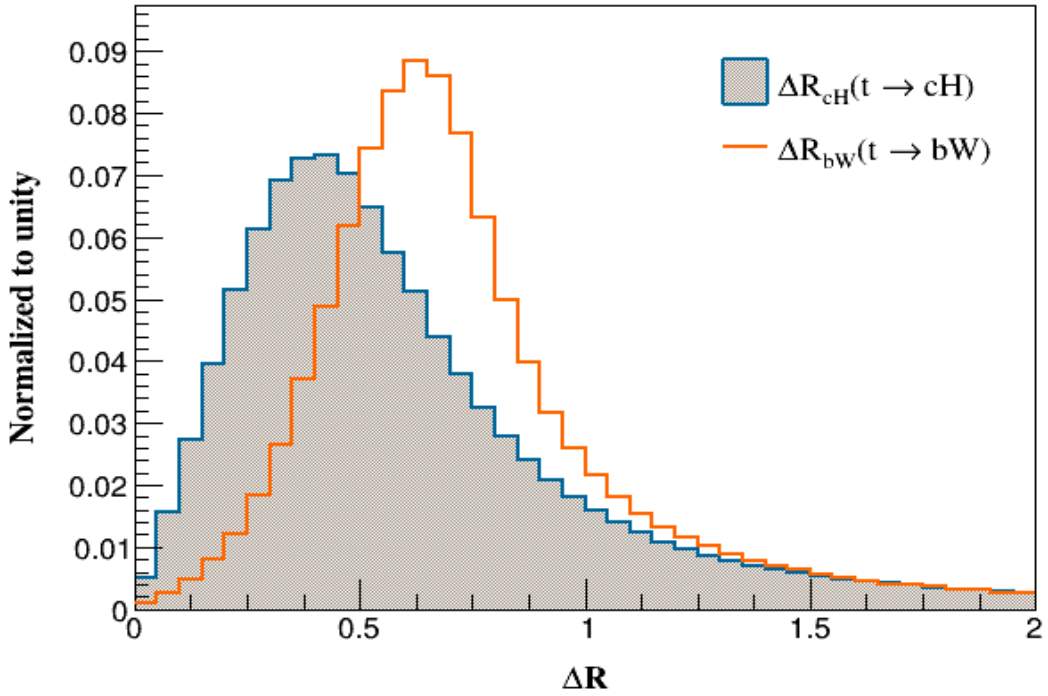


Background1

Jet Reconstruction :

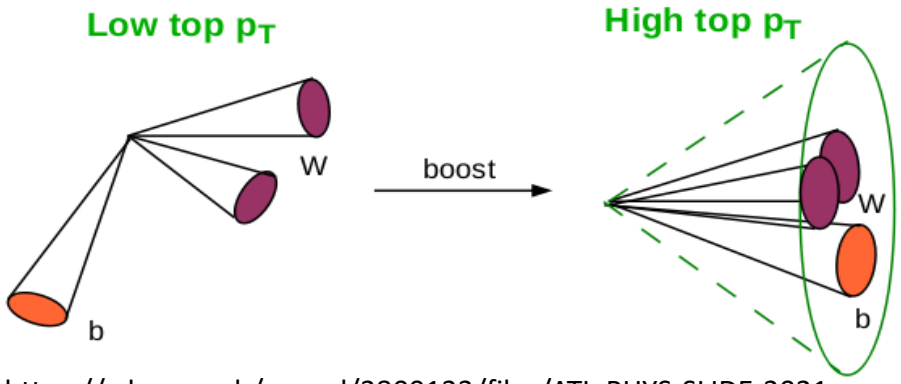


https://cds.cern.ch/record/2800123/files/ATL-PHYS-SLIDE-2021-758.pdf

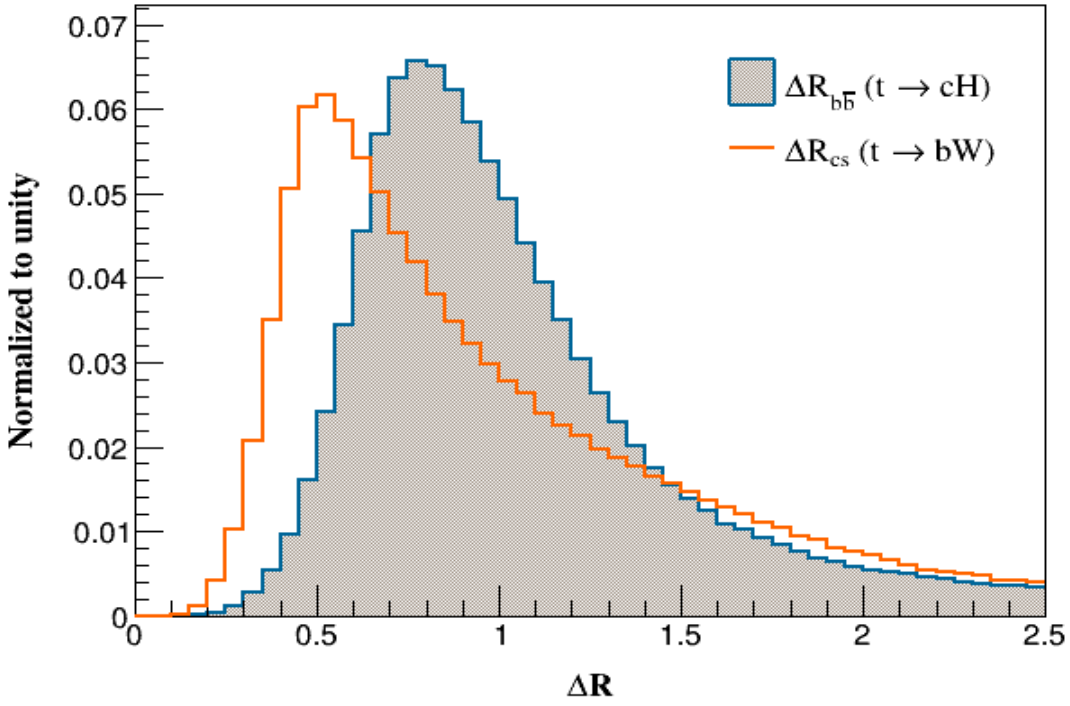


- High pT in the generation level ≈ collimated decay products.
- Depending on the separation between final state objects jet Radius determined..

Jet Reconstruction :

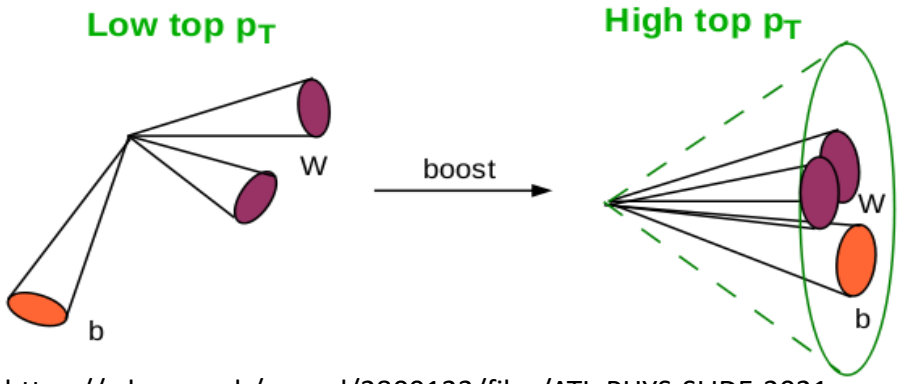


https://cds.cern.ch/record/2800123/files/ATL-PHYS-SLIDE-2021-758.pdf



- High pT in the generation level ≈ collimated decay products.
- Depending on the separation between final state objects jet Radius determined..

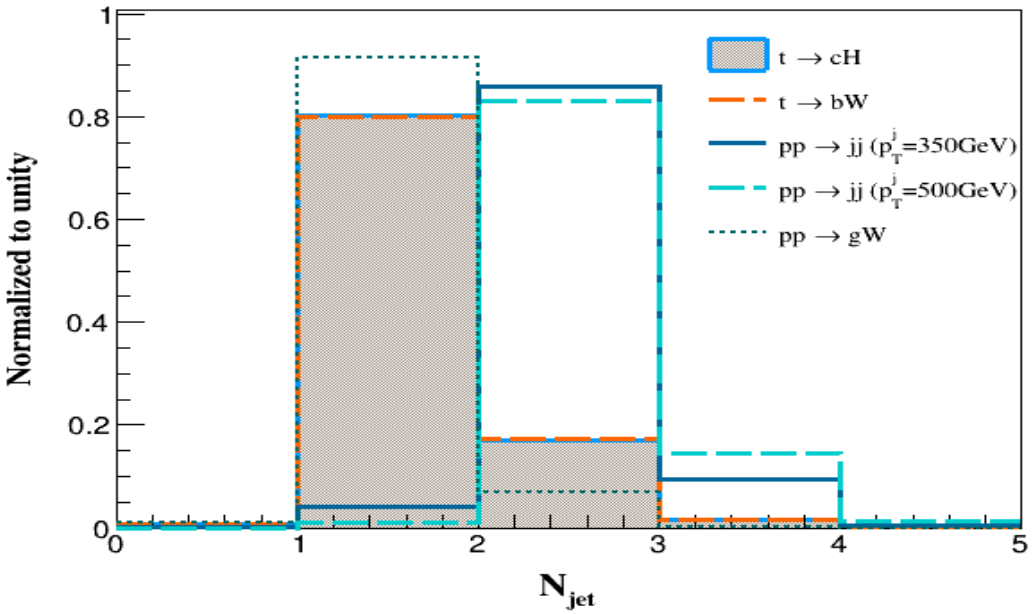
Jet Reconstruction :



https://cds.cern.ch/record/2800123/files/ATL-PHYS-SLIDE-2021-758.pdf

- High pT in the generation level ≈ collimated decay products.
- Depending on the separation between final state objects jet Radius determined..

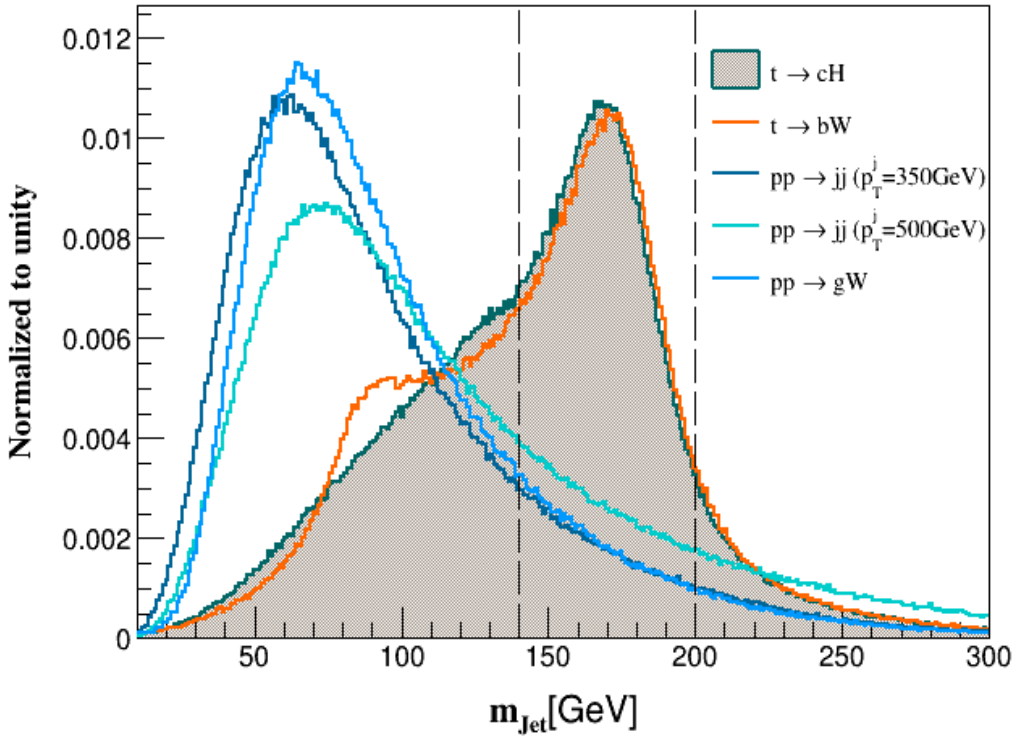
- Delphes card CMS, FastJet, anti-kT clustering algorithm…..
- Radius=1.0
- pT -jet(min)= 200GeV
- Signal and Background1 peaking at one…..



Top Candidate jet :

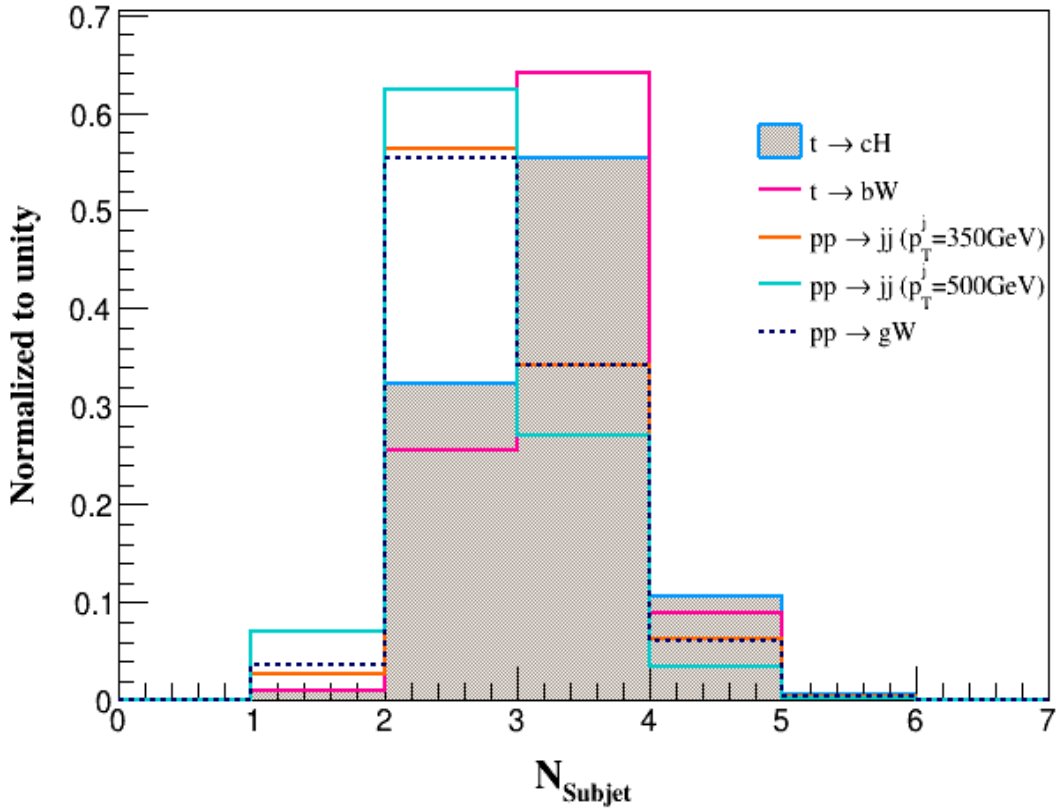
- Jet with, 140GeV < mass < 200GeV and pT > 400 GeV, considered as top candidate jet… eliminate almost 80% QCD background..

- Checked with matching of top quark with the jet, obtained only 10% discrepancy in result..
- To be more convenient with experiment , proceed with the first approach.



Features of interest:

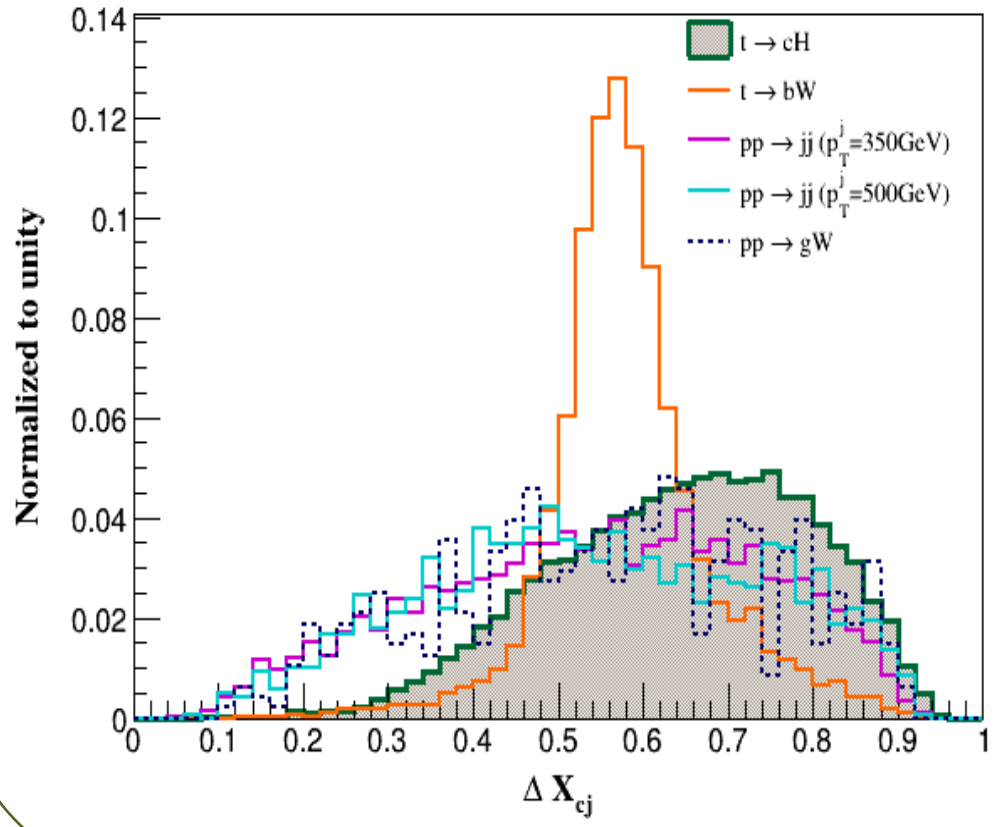
- Number of subjets



Features of interest:

- Number of subjets
- Fractional distribution of the invariant mass.

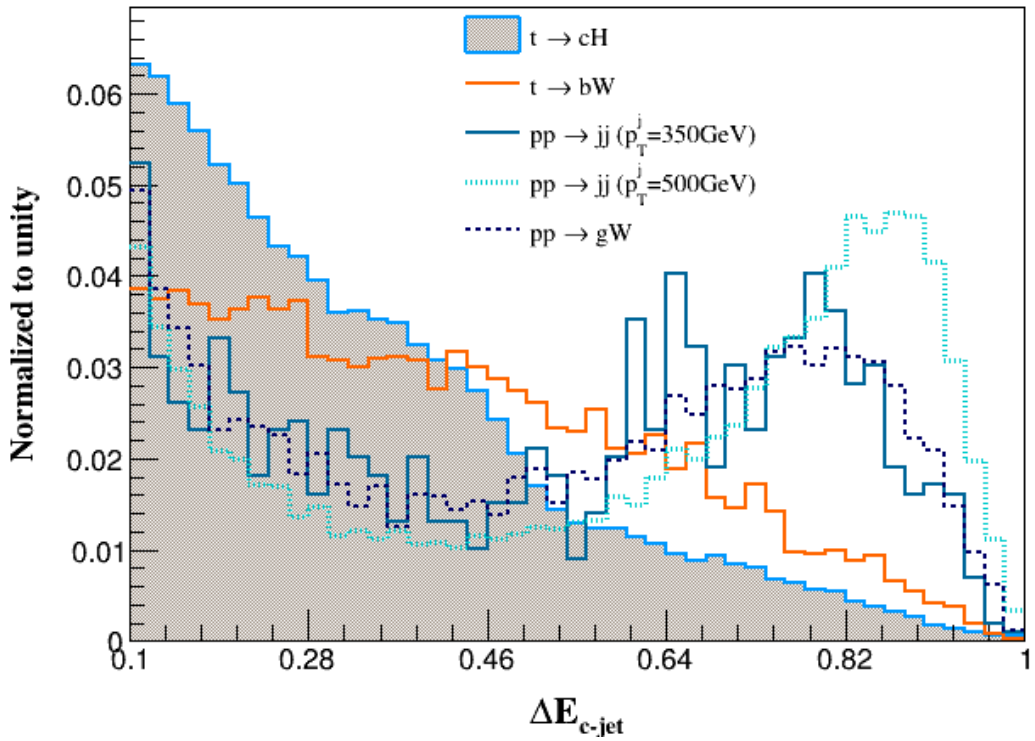
$$\Delta x_{cj} = 1 - \frac{invariant\ mass\ of\ c-jet\ and\ light-jet}{mass\ of\ the\ top-candidatejet}$$



Features of interest:

- Number of subjets.
- Fractional distribution of the invariant mass.
- Energy fraction of the subjets.

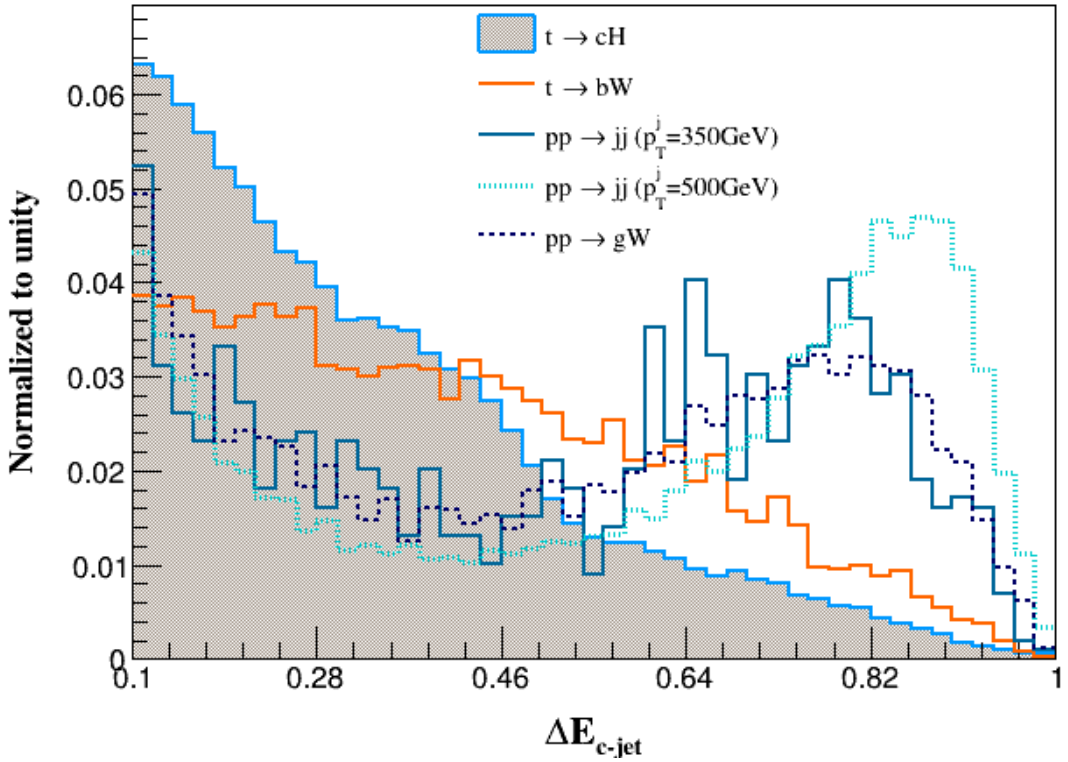
$$\Delta E_{C-jet} = \frac{energy\ of\ c-jet}{energy\ of the\ top-candidate jet}$$



Features of interest:

- Number of subjets.
- Fractional distribution of the invariant mass.
- Energy fraction of the subjets.
- Number of tracks
- Number of Displaced tracks
- Number of displaced vertex
- Invariant mass of the displaced vertices
- N-subjettiness(τ_N, τ_N / τ_{N-1})
- Number of b-jets, c-jets, light-jets
- pT and mass of leading b-jet and c-jet.

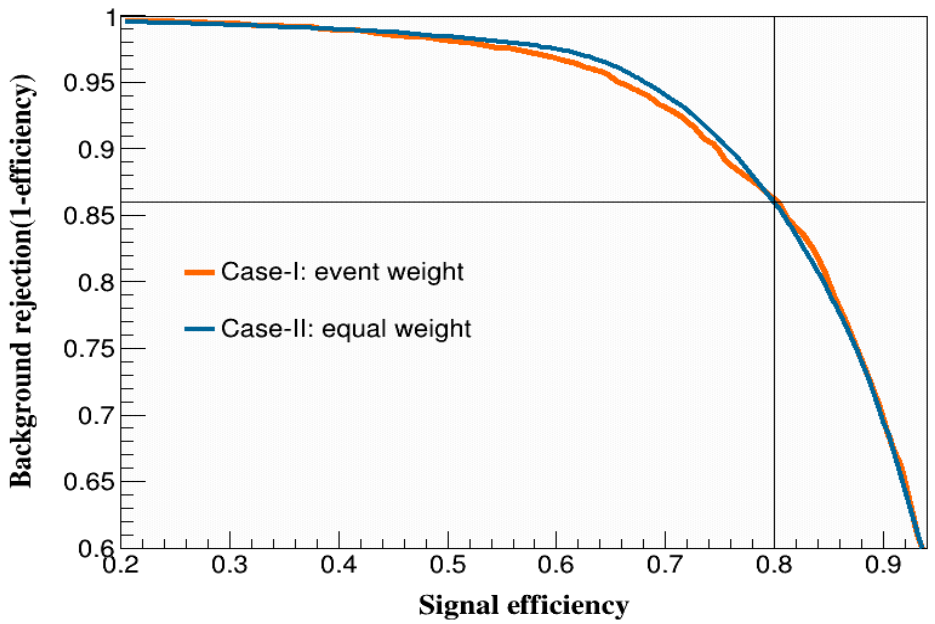
$$\Delta E_{C-jet} = \frac{energy\ of\ c-jet}{energy\ of the\ top-candidate jet}$$



Receiver Operating Characteristics(ROC):

Ranking	variable
1.	ΔX_{cj}
2.	ΔX_{bc}
3.	τ_{32}^{1}
4.	τ_{21}^{1}
5.	ΔE_{b-jet1}
6.	M_{c-jet}
7.	p_T^{c-jet}
8.	ΔE_{c-jet}
9.	p_T^{b-jet}
10.	ΔX_{bj}
11.	M_{b-jet}
12.	$N_{light-jet}$
13.	N_{Subjet}

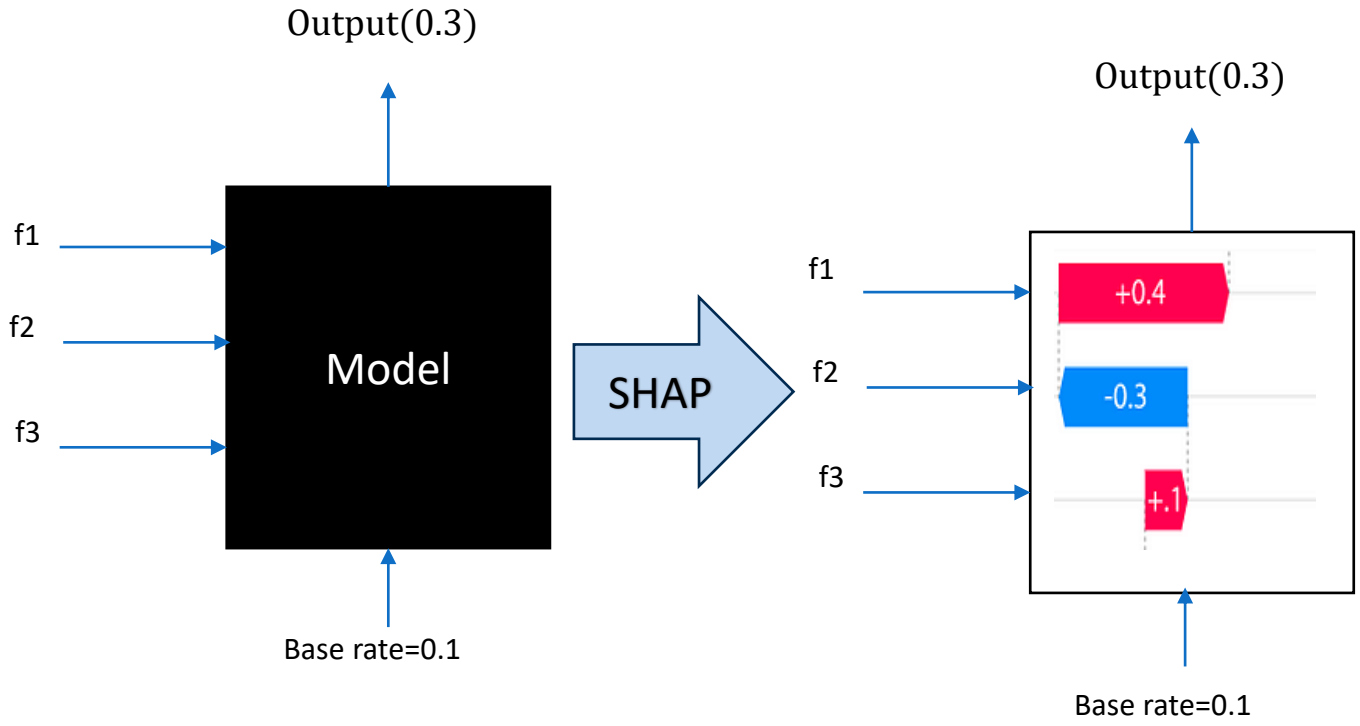
Algorithm: BDT
Classifier: AdaBoost
Cuts:1b-jet, 1c-jet



- $Weight = \frac{1}{N_{MC}} \times \sigma_{MC} \times BR \times \int \mathcal{L}\, dt$
- $N_{MC} = Number\ of\ monte\ carlo\ events\ generated$
- $\sigma_{MC} = Production\ cross-section\ for\ the\ process$
- $BR = Net\ branching\ fraction\ of\ the\ combined\ decay\ channel$
- $\int \mathcal{L}\, dt = Integrated\ luminosity$

Considering only the Background1 **XGBoost** proved to be better than AdaBoost….

Shapley Additive explanation (SHAP):

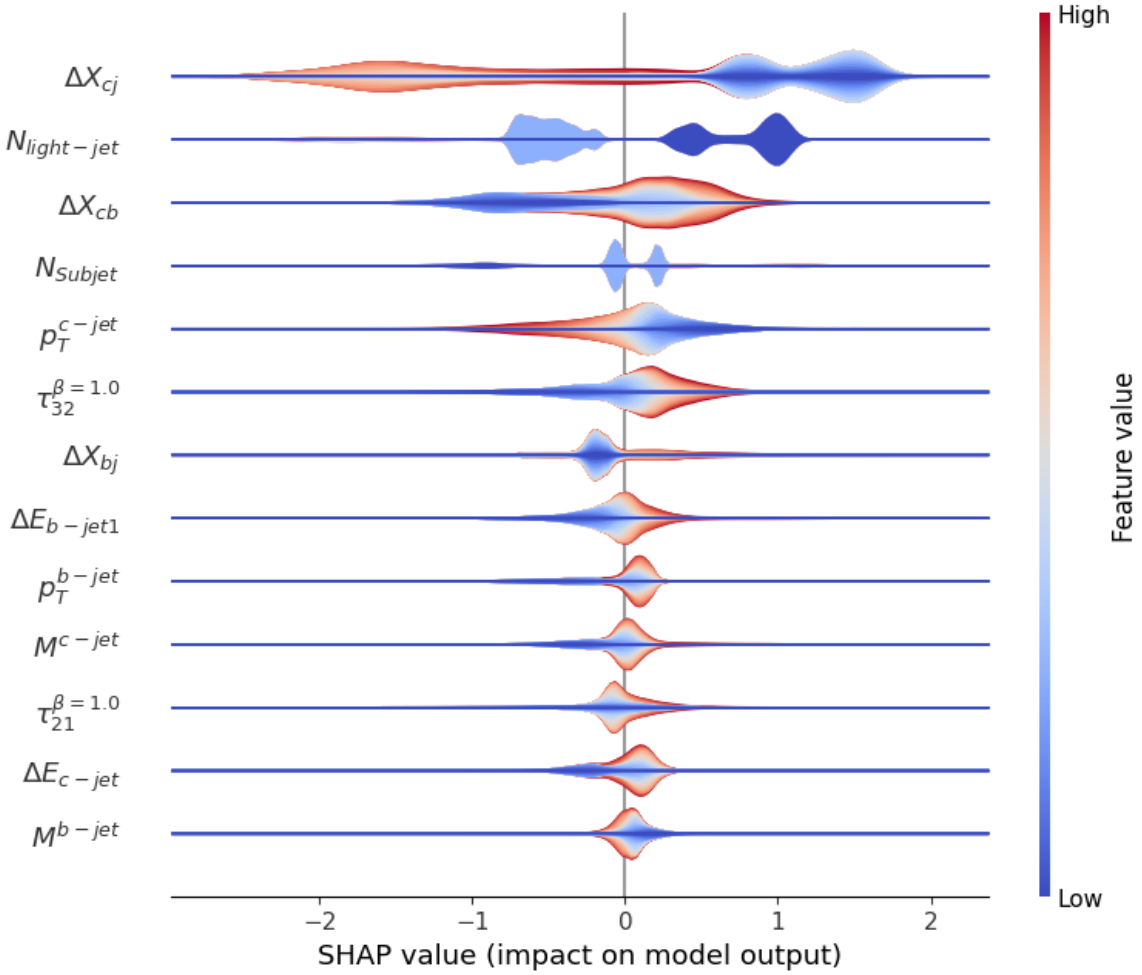


- Proposed by Lloyd Shapley in 1954..

$$\varphi_i(v) = \frac{|S|!\,(n-|S|-1)!}{n!} \sum_{S \subseteq N \setminus \{i\}} (v(S \cup \{i\}) - v(S))$$

- S : subset of the total number of features n, removing i th one
- |S| : cardinality of the set S
- v : a function which maps the subset of features to an integer number

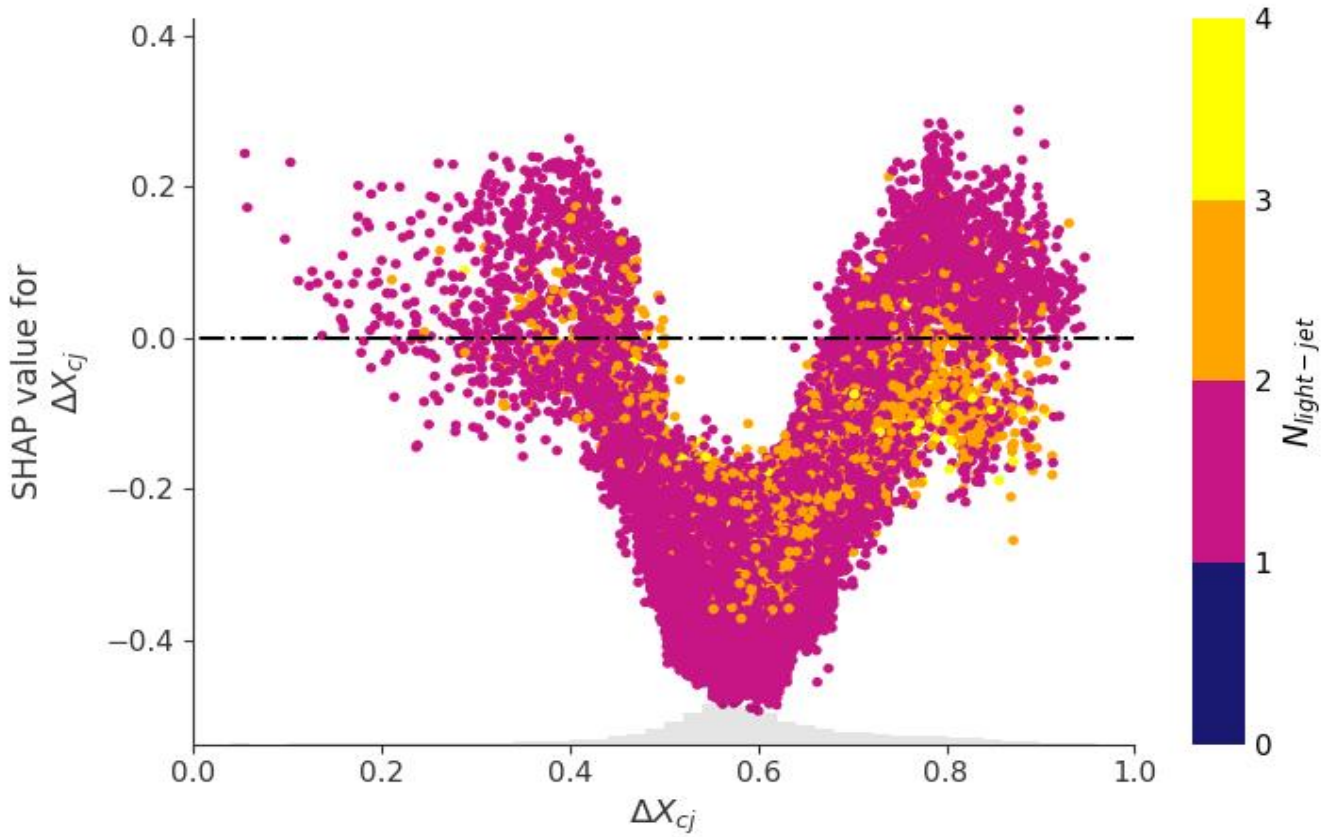
Summary plot:



- Decreasing order of importance .....
- Positive value ∈ Signal like
- Negative value ∈ Background like
- Thick regions represent high density of data points around that Shapley value.
- Ex:
  - Most important feature Δx_{cj}, Higher value of the feature ≅ background-like events, smaller value of the feature ≅ signal-like events.
  - Lower value of number of light-jets ≅ signal-like events, higher value of the feature ≅ background-like events.

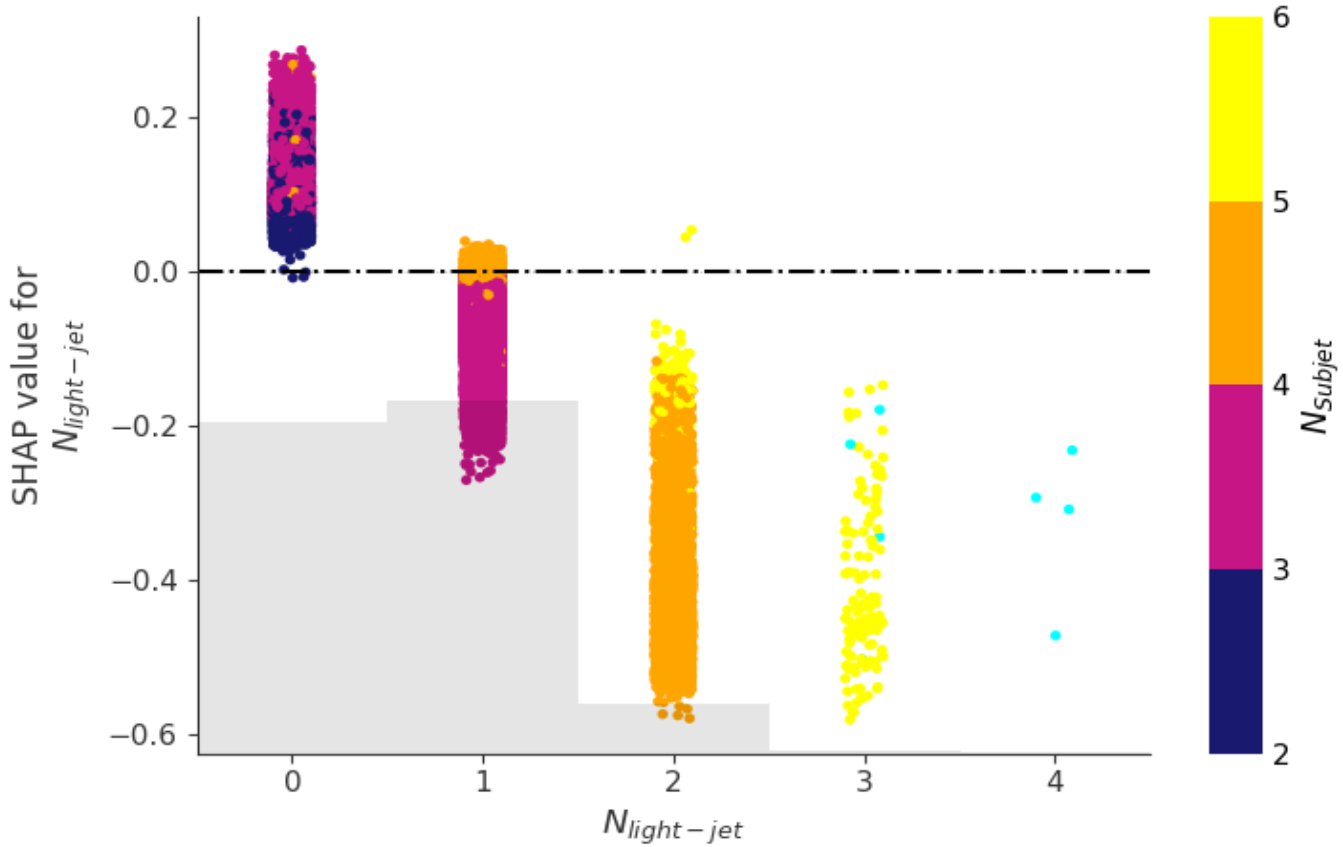
Scatter dependence plot:

- Global interpretation of data ….
- Shows distribution of Shapley values of a feature with respect to that feature..
- Shows dependence of particular feature on other feature ..
- Ex:
  - Δx_{cj} around 0.6 have negative Shapley value ≅ background-like
  - 2 or more than 2 number of light-jets mostly in background region.



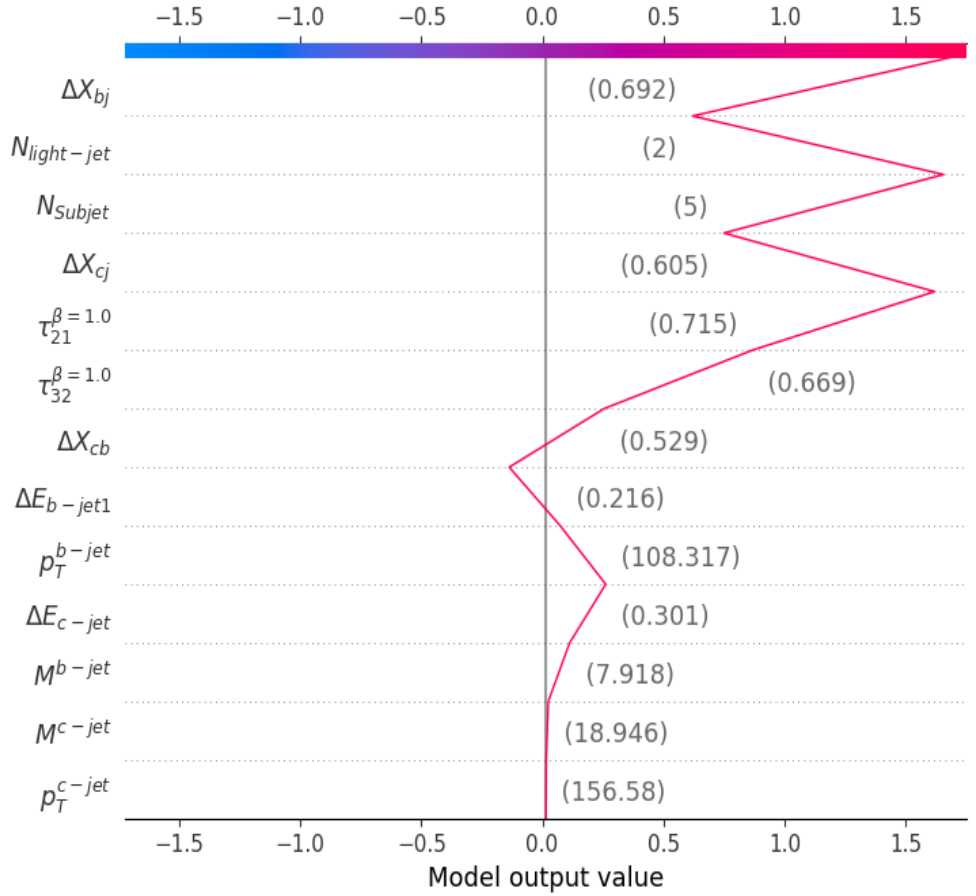
Scatter dependence plot :

- Global interpretation of data ….
- Shows distribution of Shapley values of a feature with respect to that feature..
- Shows dependence of particular feature on other feature ..
- Ex:
  - 3 subjets and zero light-jet events identified as signal..
  - 3 subjets with one light-jet events as background…
  - More than one light-jet events understands the event as background-like..



Decision plot :

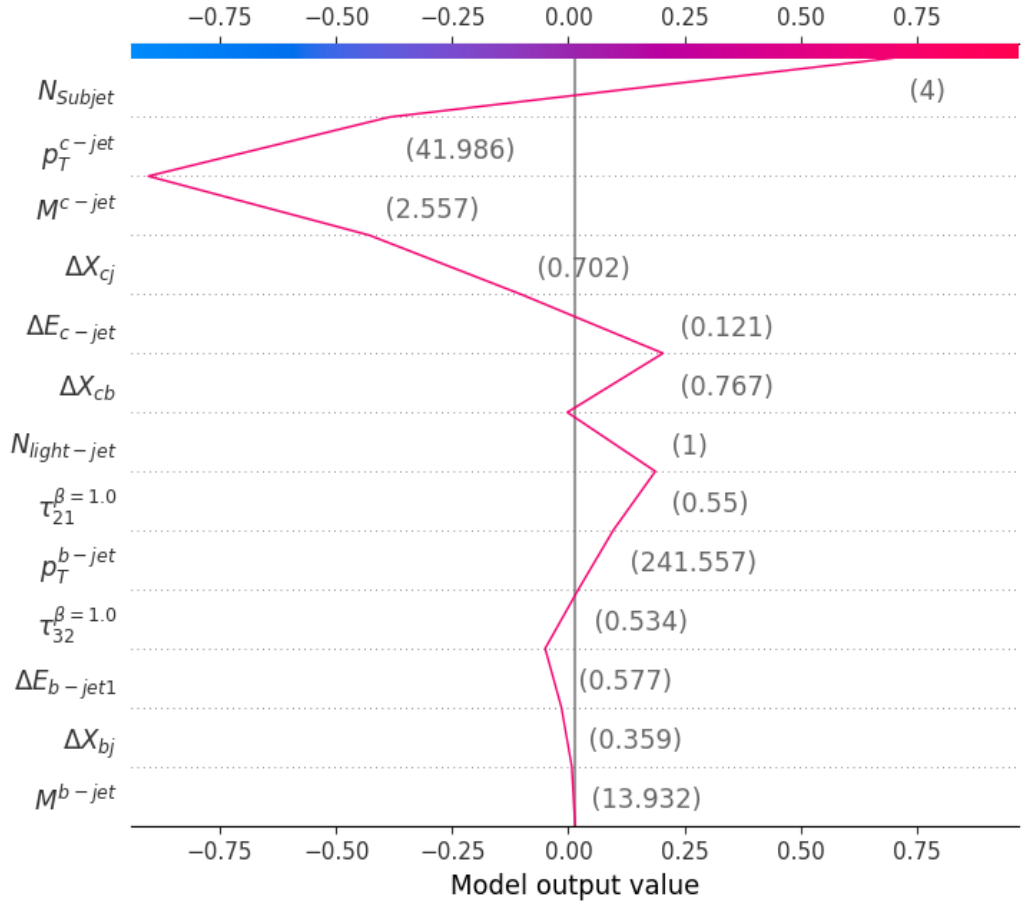
- Event by event representation of feature contribution..
- Starts from base value, ends at the predicted output value..
- Feature at bottom is the least important, and at the top is most important….
- Order of importance different from the summary plot as portrays local trait..
- Ex: The features chiefly responsible for proper identification of this event
  - **the residual mass fraction of the c-tagged jet b-tagged jet pair** carrying value 0.529 for this event, $\boldsymbol{\tau_{32}}$ **and** $\boldsymbol{\tau_{21}}$ with values 0.669 and 0.715 respectively, **number of subjets** 5.



Correctly classified signal

Decision plot :

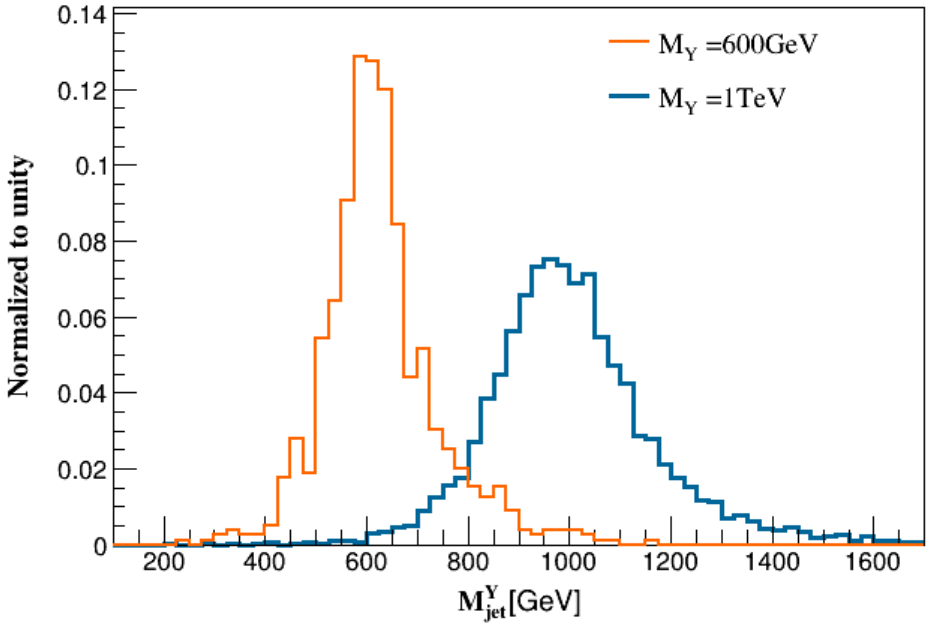
- Event by event representation of feature contribution..
- Starts from base value, ends at the predicted output value..
- Feature at bottom is the least important, and at the top is most important….
- Order of importance different from the summary plot as portrays local trait..
- Ex: The features chiefly responsible for improper identification of this event
  - the **pT of c-jet** 41.99 for this event, τ_{32} **and** τ_{21} with values > 0.5, **number of subjets** 4, **pT of b-jet** 241.56 .



Misclassified Signal

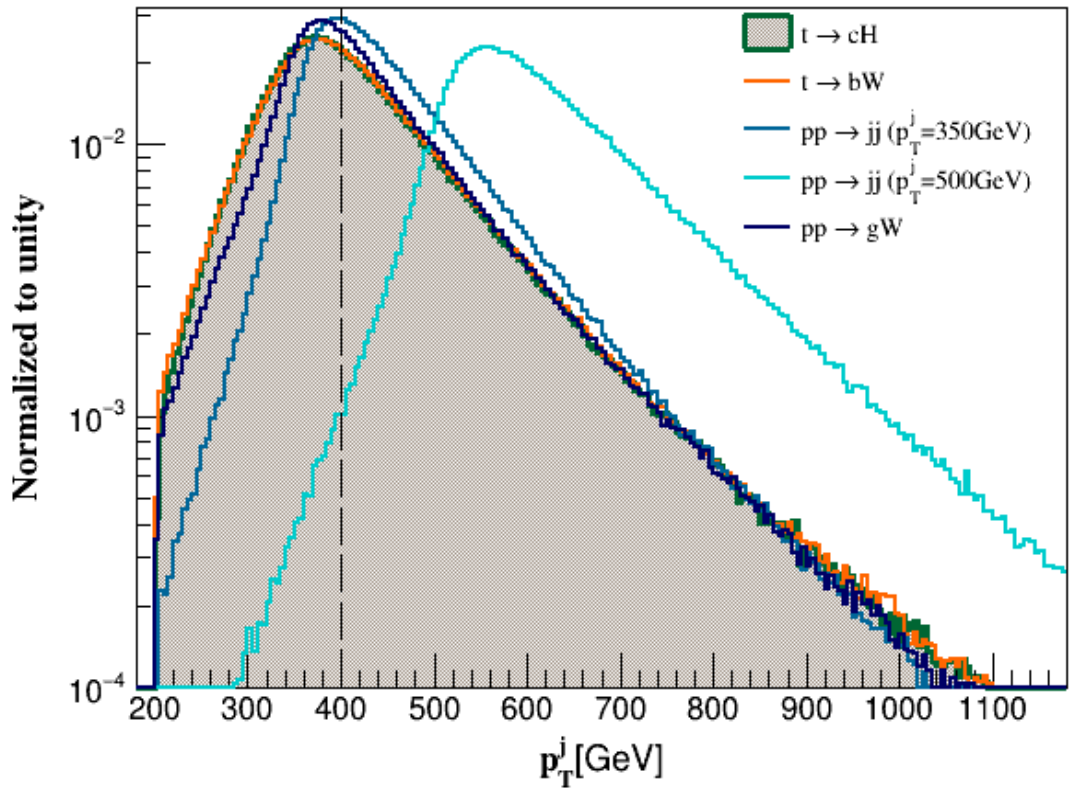
Conclusion:

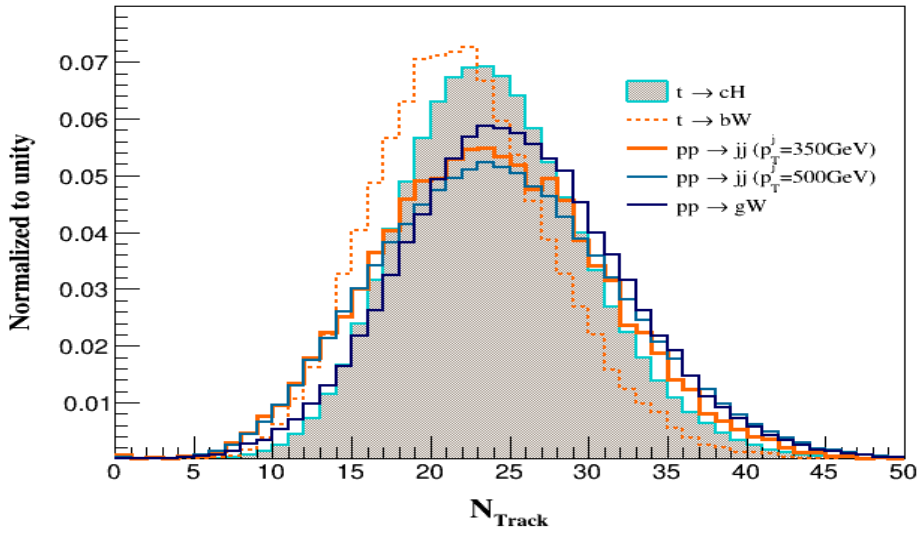
- Model independent generic approach.
- Efficiency of Signal identification with conventional top taggers found to be around 30%.
- Efficiency of Signal identification with the tagger described in this work is 80%, over the background efficiency of 15%.
- Can be used in other BSM particles with similar substructure.
- Checked with Vector-like quark Y having at least 1b 1c in final state, successfully reconstructed jet mass at mass of Y.

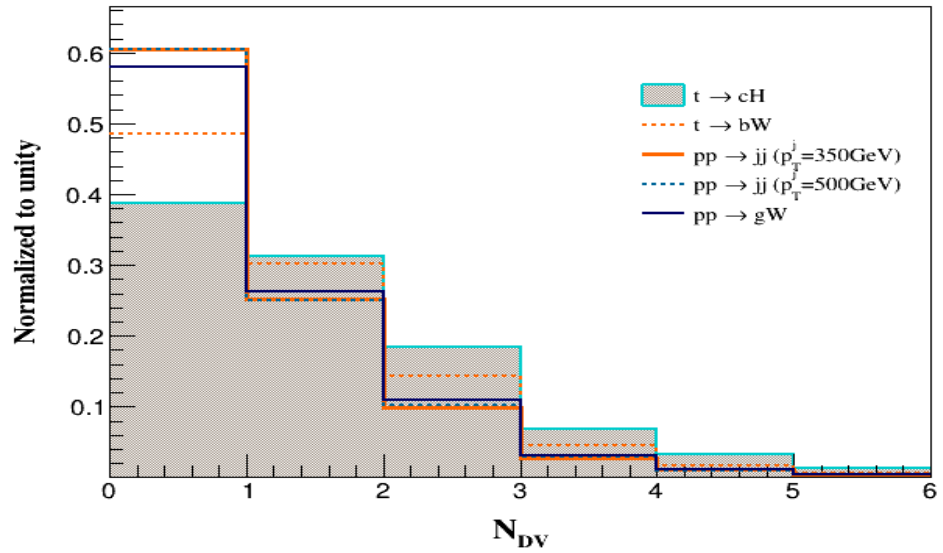


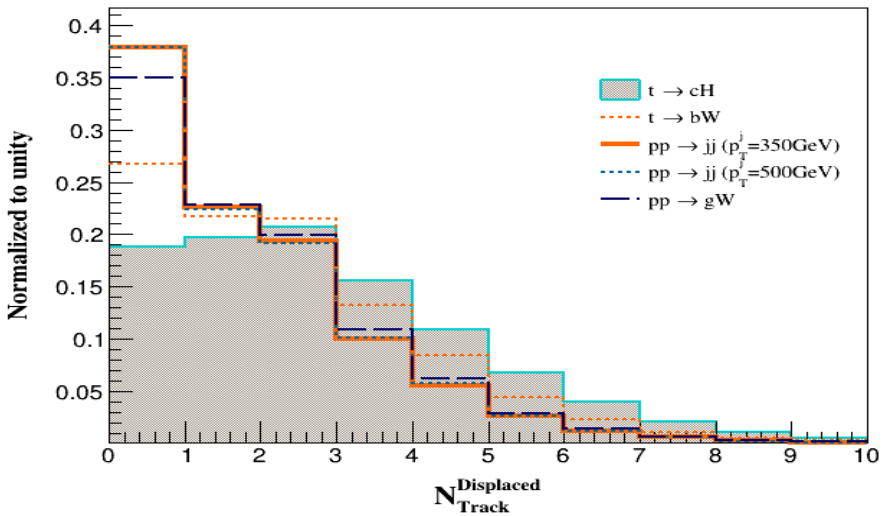
Thank You.....

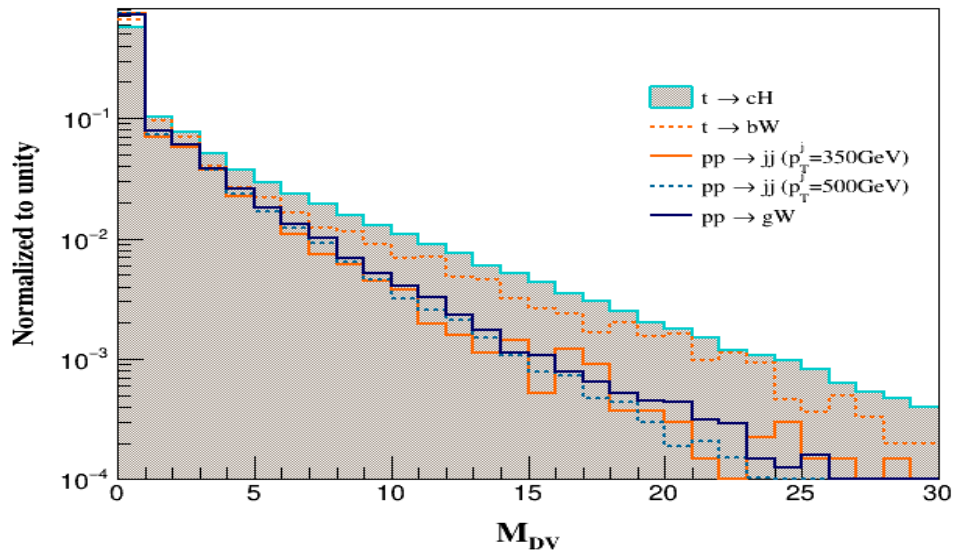
Backup Slides….

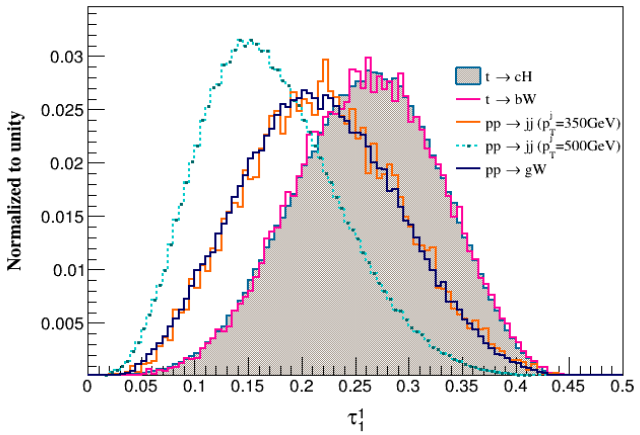

t → cH
t → bW
pp → jj (p_T^j=350GeV)
pp → jj (p_T^j=500GeV)
pp → gW
Normalized to unity
10^{-2}
10^{-3}
10^{-4}
200
300
400
500
600
700
800
900
1000
1100
p_T^j[GeV]

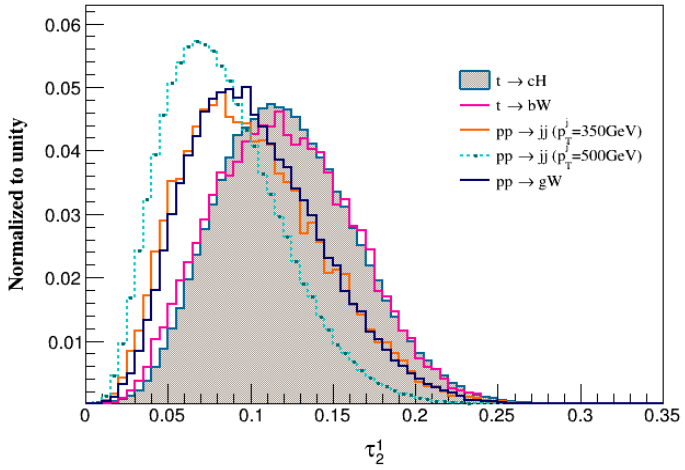

Normalized to unity
N_{Track}
t → cH
t → bW
pp → jj (p_T^j=350GeV)
pp → jj (p_T^j=500GeV)
pp → gW

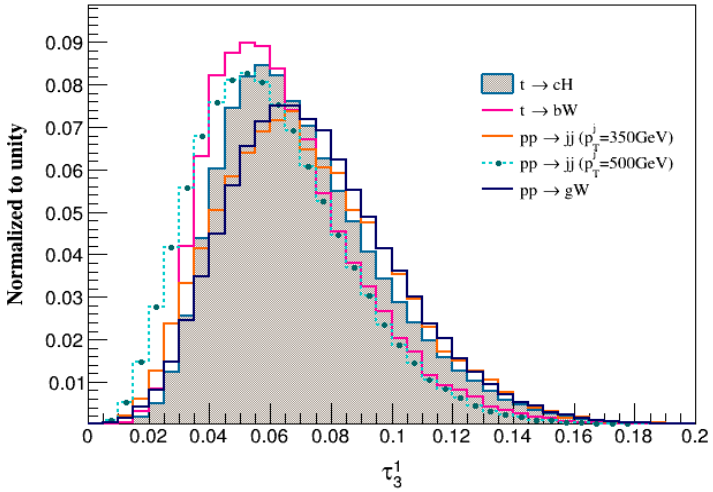

Normalized to unity
N_{DV}
t → cH
t → bW
pp → jj (p_T^j=350GeV)
pp → jj (p_T^j=500GeV)
pp → gW

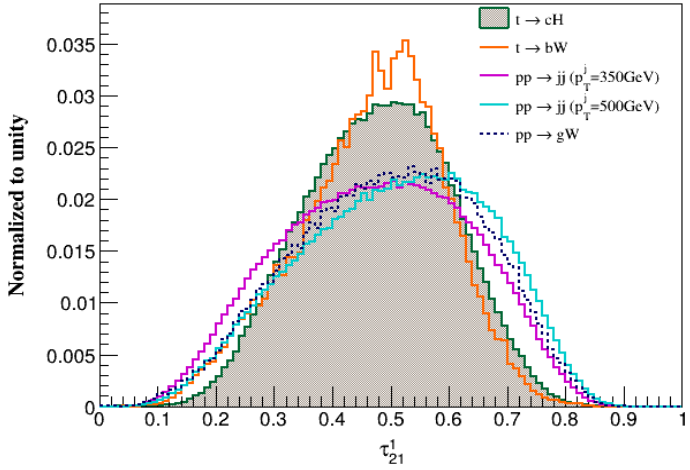

Normalized to unity
$N_{Track}^{Displaced}$
t → cH
t → bW
pp → jj (p_T^j=350GeV)
pp → jj (p_T^j=500GeV)
pp → gW

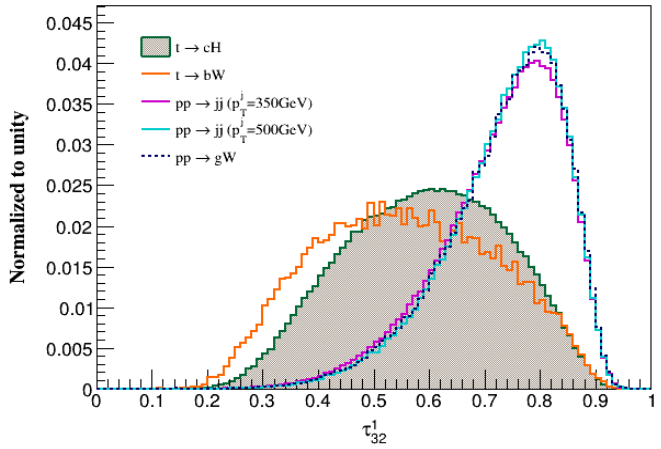

Normalized to unity
M_{DV}
t → cH
t → bW
pp → jj (p_T^j=350GeV)
pp → jj (p_T^j=500GeV)
pp → gW

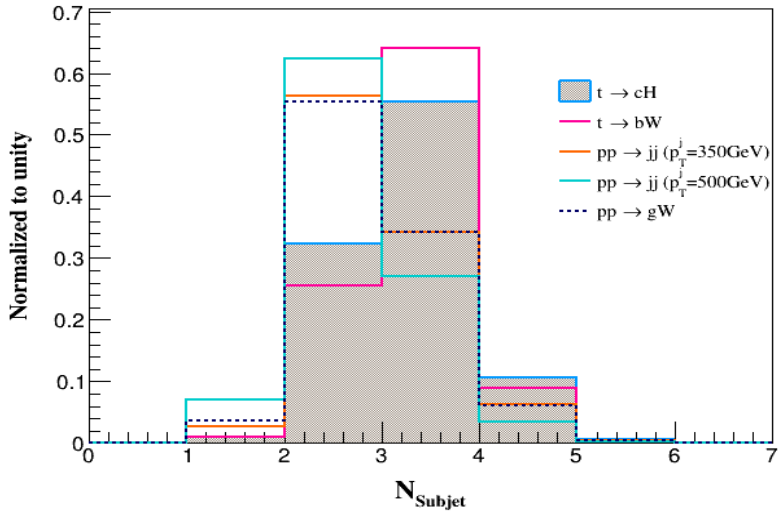

Normalized to unity
t → cH
t → bW
pp → jj (p_T^j=350GeV)
pp → jj (p_T^j=500GeV)
pp → gW
τ_1^1

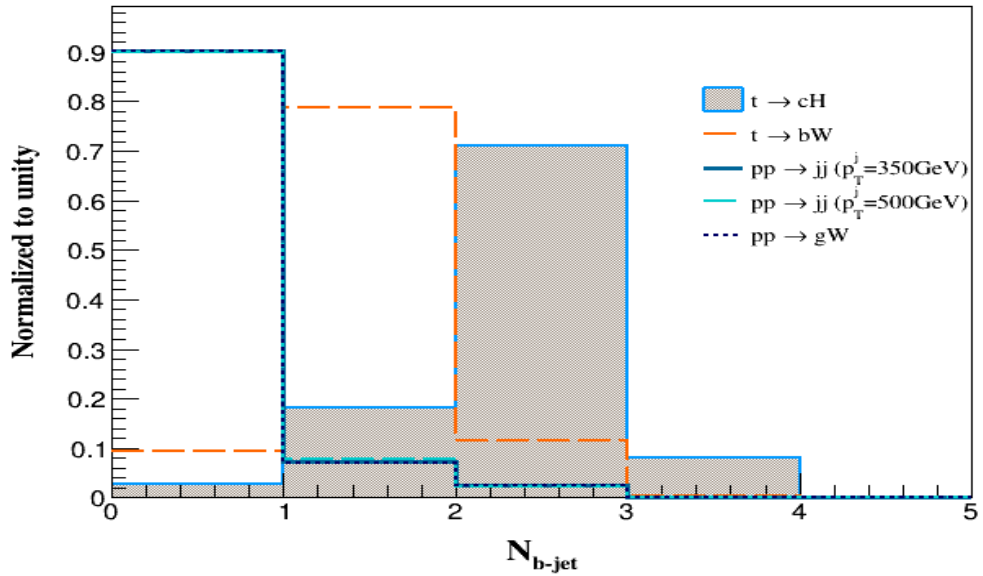

Normalized to unity
t → cH
t → bW
pp → jj (p_T^j=350GeV)
pp → jj (p_T^j=500GeV)
pp → gW
τ_2^1

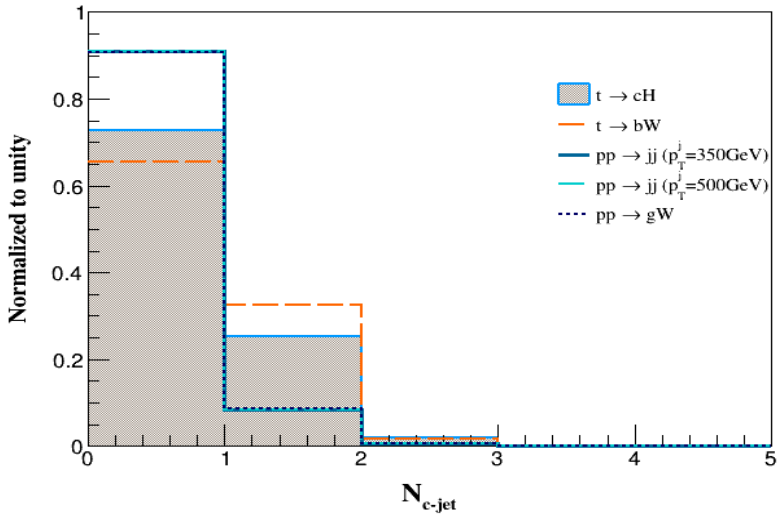

Normalized to unity
t → cH
t → bW
pp → jj (p_T^j=350GeV)
pp → jj (p_T^j=500GeV)
pp → gW
τ_3^1

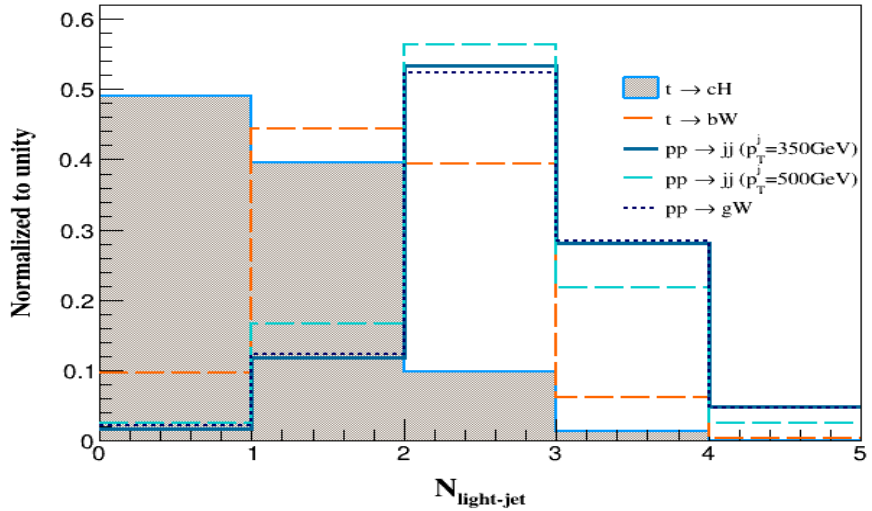

Normalized to unity
t → cH
t → bW
pp → jj (p_T^j=350GeV)
pp → jj (p_T^j=500GeV)
pp → gW
τ_{21}^1

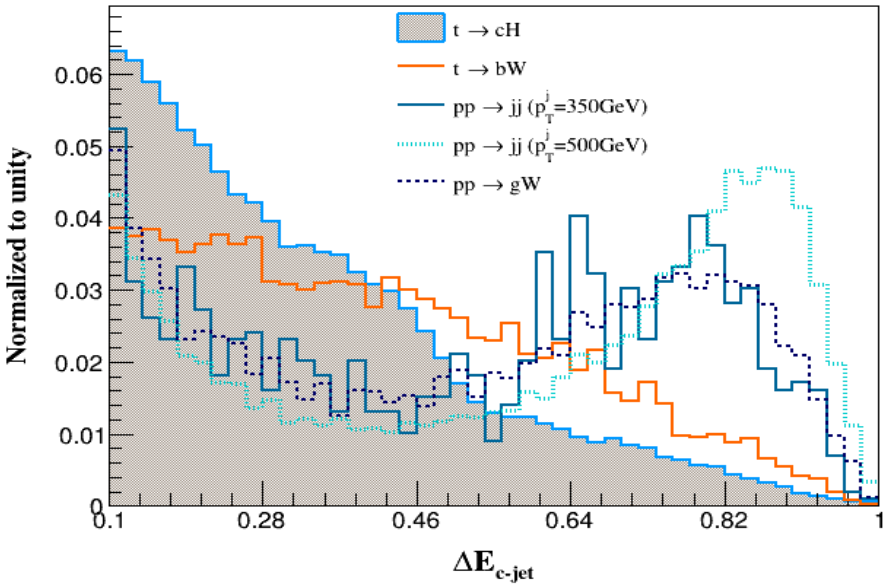

Normalized to unity
t → cH
t → bW
pp → jj (p_T^j=350GeV)
pp → jj (p_T^j=500GeV)
pp → gW
τ_{32}^1

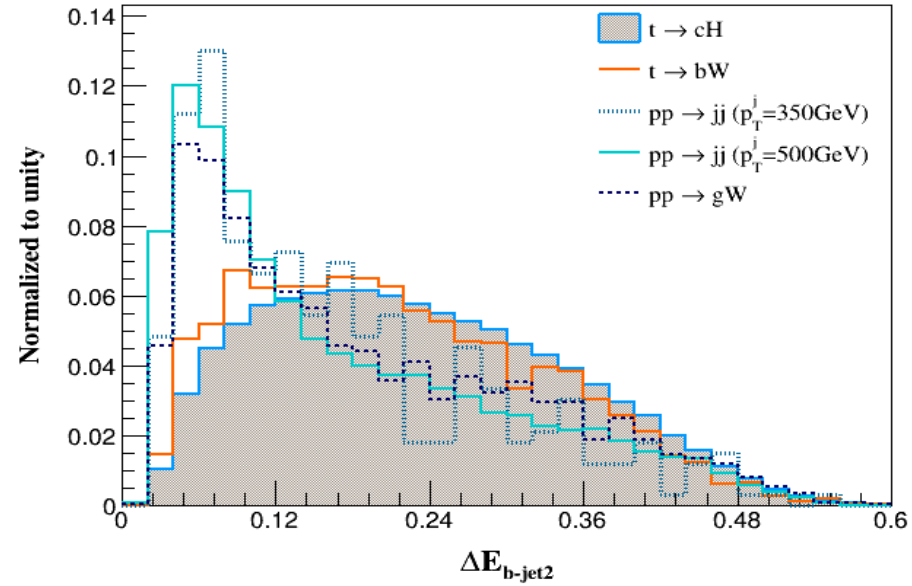

Normalized to unity
N_{Subjet}
$t \to cH$
$t \to bW$
$pp \to jj\ (p_T^j=350GeV)$
$pp \to jj\ (p_T^j=500GeV)$
$pp \to gW$

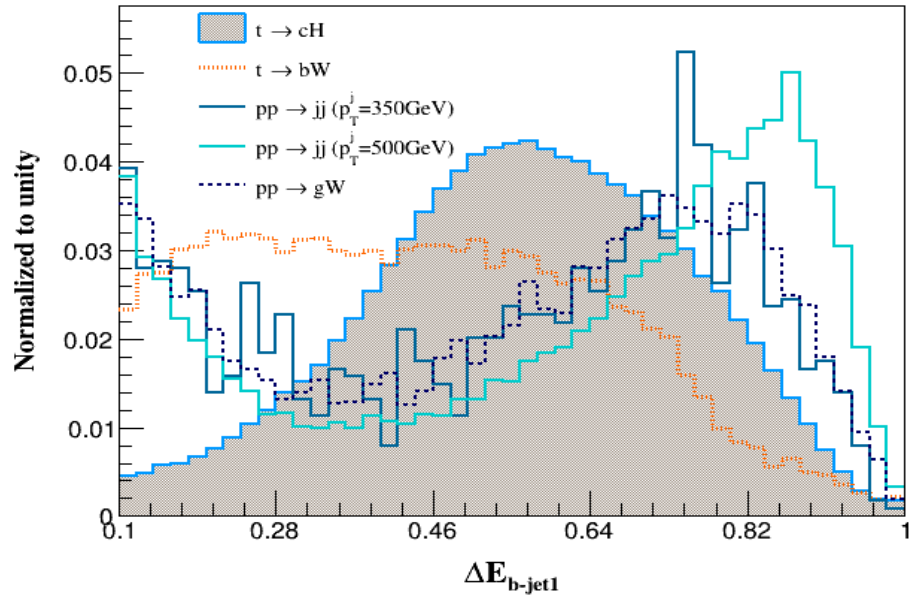

Normalized to unity
$N_{b\text{-}jet}$
$t \to cH$
$t \to bW$
$pp \to jj\ (p_T^j=350GeV)$
$pp \to jj\ (p_T^j=500GeV)$
$pp \to gW$


Normalized to unity
$N_{c\text{-}jet}$
$t \to cH$
$t \to bW$
$pp \to jj\ (p_T^j=350GeV)$
$pp \to jj\ (p_T^j=500GeV)$
$pp \to gW$

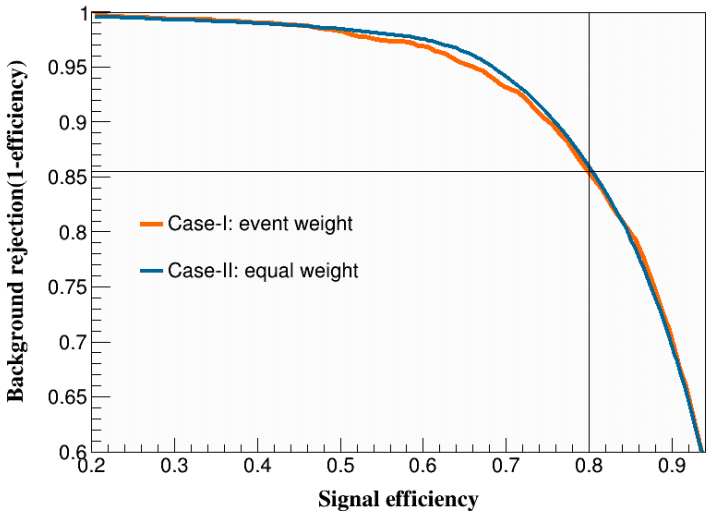

Normalized to unity
$N_{light\text{-}jet}$
$t \to cH$
$t \to bW$
$pp \to jj\ (p_T^j=350GeV)$
$pp \to jj\ (p_T^j=500GeV)$
$pp \to gW$

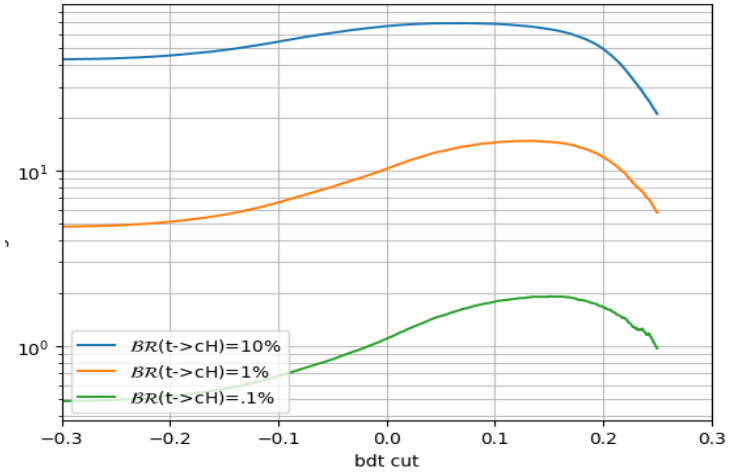

t → cH
t → bW
pp → jj (p_T^j=350GeV)
pp → jj (p_T^j=500GeV)
pp → gW
Normalized to unity
$\Delta E_{c\text{-}jet}$

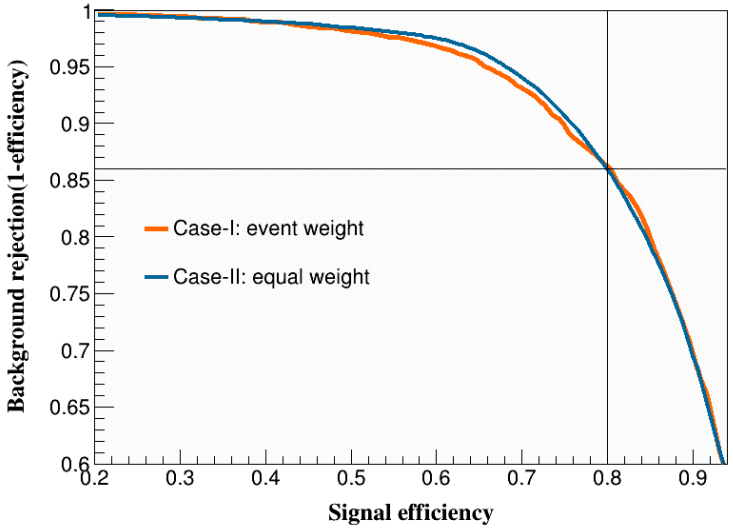

t → cH
t → bW
pp → jj (p_T^j=350GeV)
pp → jj (p_T^j=500GeV)
pp → gW
Normalized to unity
$\Delta E_{b\text{-}jet2}$


t → cH
t → bW
pp → jj (p_T^j=350GeV)
pp → jj (p_T^j=500GeV)
pp → gW
Normalized to unity
$\Delta E_{b\text{-}jet1}$

Ranking	variable
1.	ΔX_{cj}
2.	ΔX_{bc}
3.	τ_{32}^{1}
4.	τ_{21}^{1}
5.	ΔE_{b-jet1}
6.	M_{c-jet}
7.	p_T^{c-jet}
8.	ΔE_{c-jet}
9.	p_T^{b-jet}
10.	ΔX_{bj}
11.	M_{b-jet}
12.	$N_{light-jet}$
13.	N_{Subjet}
14.	$N_{Track}^{Displaced}$
15.	M_{DV}
16.	N_{DV}

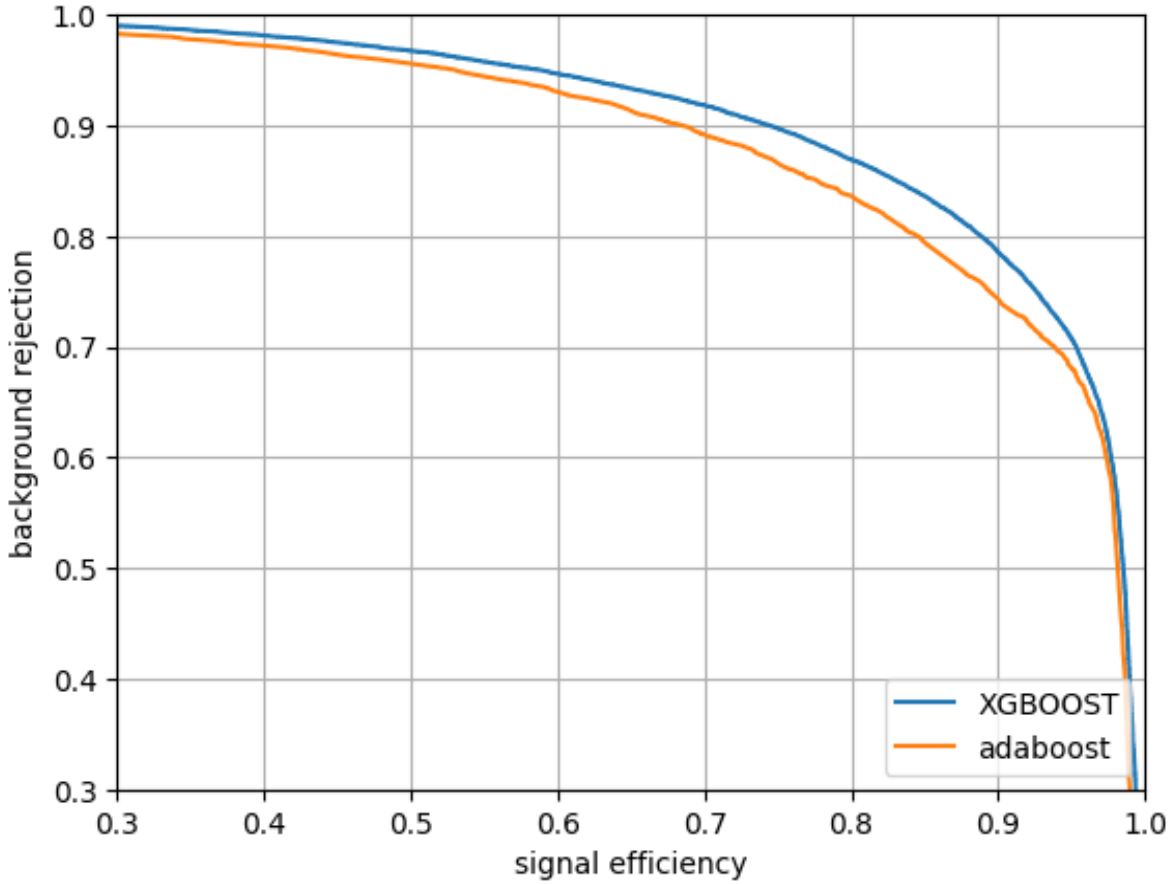


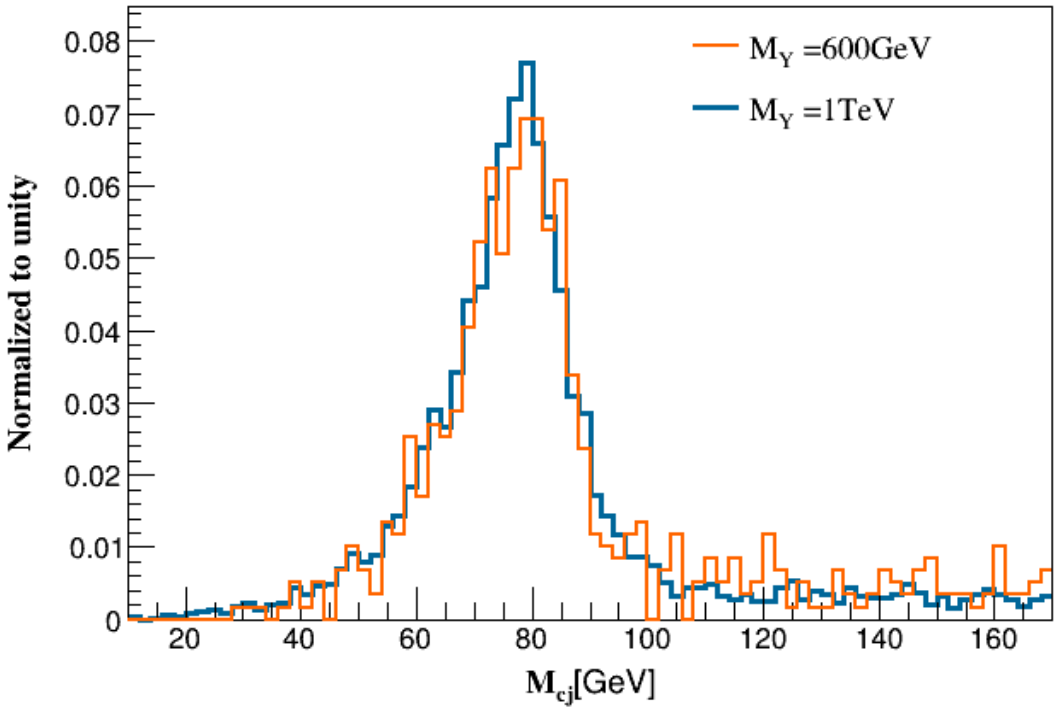


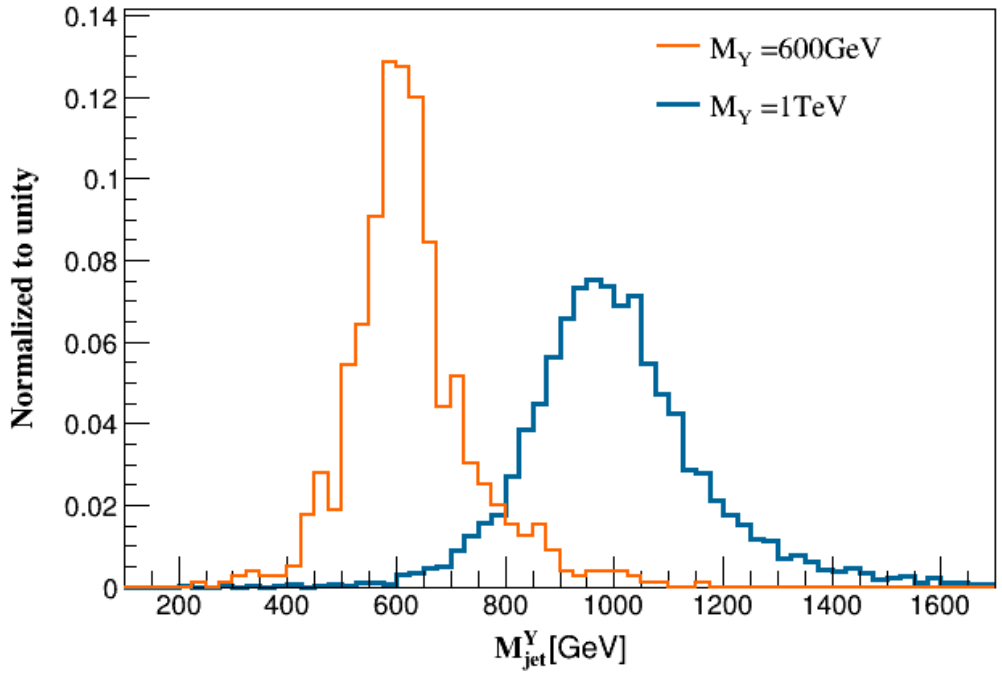


Branching ratio($t \to cH$)	BDT Cut	max(Significance)
100%	-0.08	270
10%	0.08	69
1%	0.12	14.6
0.1%	0.15	1.9

Sample type	*Weight*
• *Background1*	4.6×10^{-2}
• *Background2*	2.1×10^{3}
• *Background3*	1.8×10^{3}
• *Background4*	6.7×10^{-1}


1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
background rejection
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
signal efficiency
XGBOOST
adaboost


Normalized to unity
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
20
40
60
80
100
120
140
160
M_{cj}[GeV]
M_Y =600GeV
M_Y =1TeV


Normalized to unity
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
200
400
600
800
1000
1200
1400
1600
M^Y_{jet}[GeV]
M_Y =600GeV
M_Y =1TeV