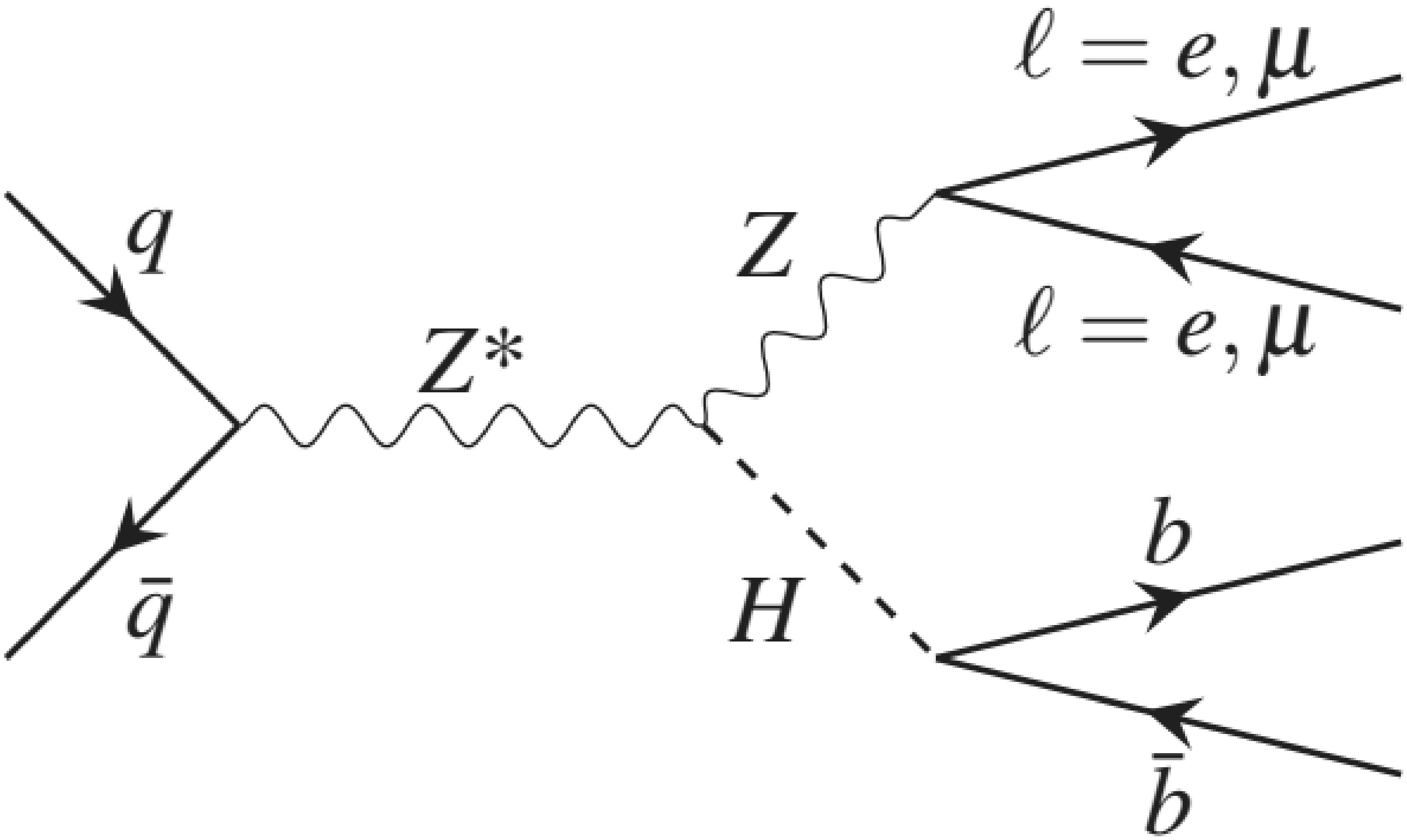
Search for the Higgs Boson decay to b quarks

Ana Patrícia Afonso & Joshua Winter | The ATLAS *Group*

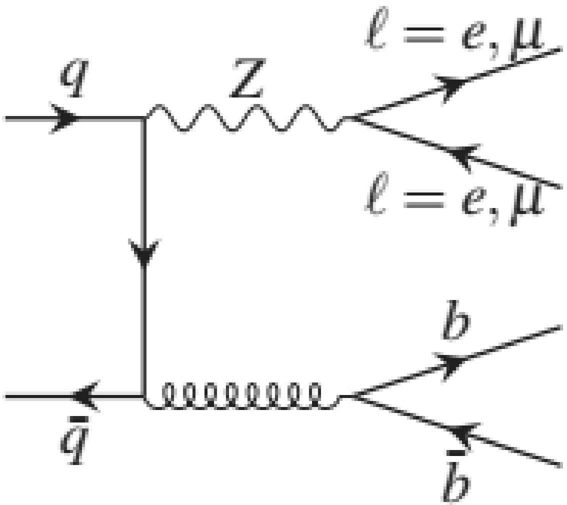
Why are we still searching?

- The existence of the Higgs boson is an essential part of the Standard Model.
- It was found in 2012, decaying into a pair of vector bosons (WW or ZZ) or into two photons.
- $H \to bb$ is the most probable decay channel (57%) according to the SM, but it has not been observed yet.
- This channel is key to studying the properties of the Higgs in greater detail and to test its correlation to the predictions of the Standard Model.

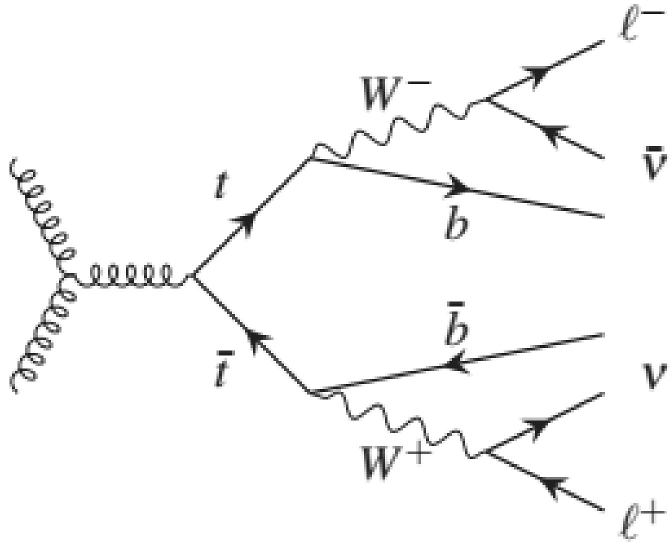
ZH signal



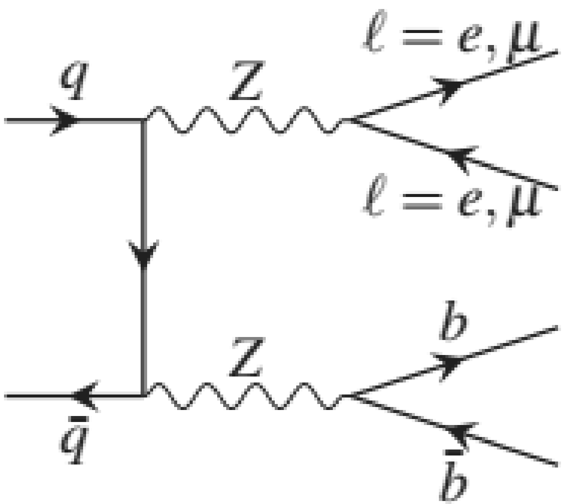
Main Backgrounds



Z+b jets, Z→LL

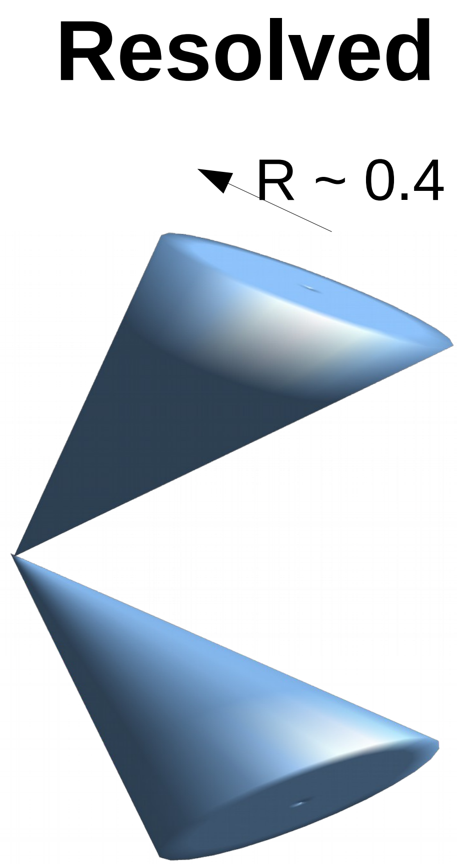


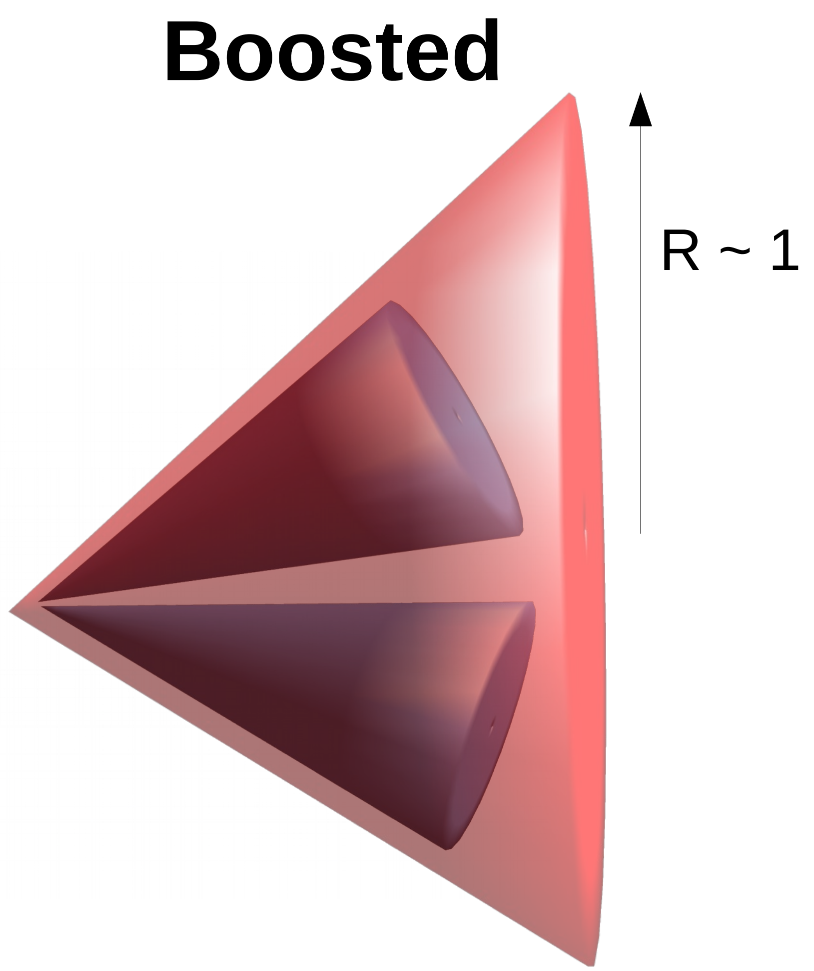
ttbar→WW+bb



ZZ→LL+bb

ZH signal





Objectives

- Develop our search for the Higgs boson by using simulations of signal and background events to:
  - Study the **Higgs mass resolution** for events reconstructed in the **resolved** and **boosted** regimes;
  - Compare these methods and attempt to improve the **significance** of our analysis by considering the **efficiency** in selection and the **resolution** of boosted reconstructed ZH events in the **high pT** range.

Event selection

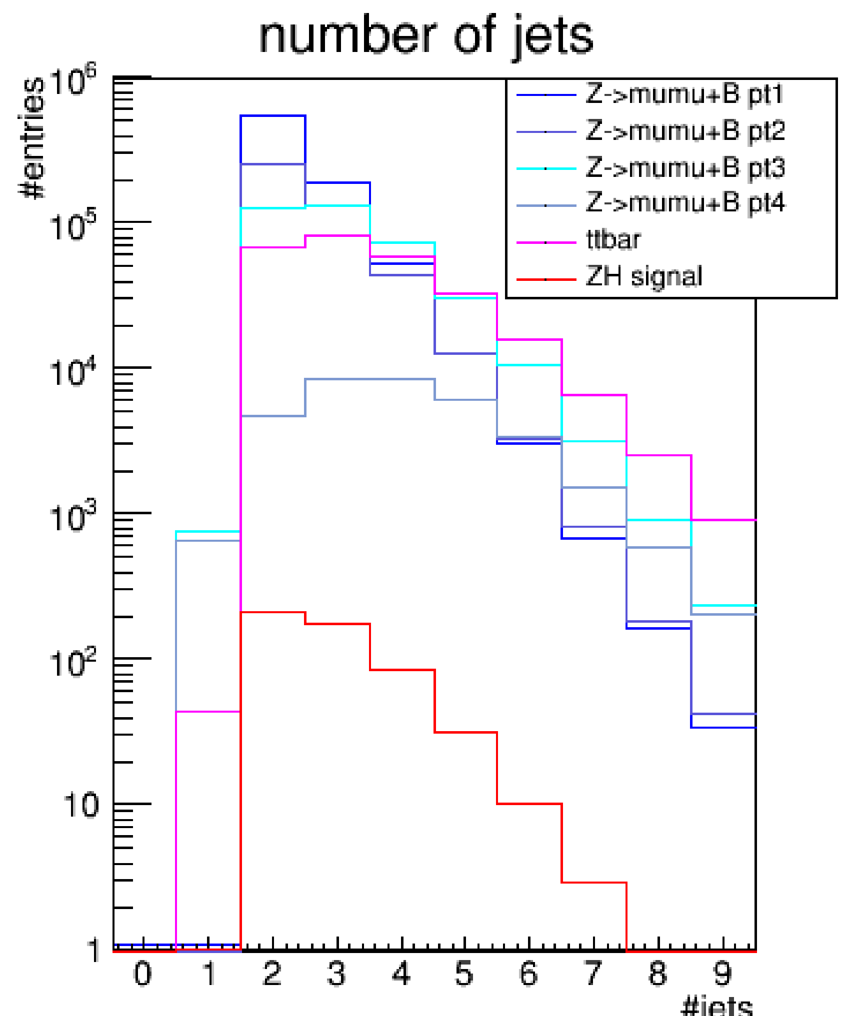
- Events with Z decaying to **two same flavour, opposite charge leptons** (electrons or muons)
- H decaying to **two b-tagged jets** or **one fat jet** with **two b-tagged subjets**
- Invariant **mass** of **lepton pair** corresponds to Z decay

Event selection Z→ll cuts

- Cut 1: Transverse momentum of leading lepton: **pTL1 > 27 GeV**
- Cut 2: Invariant mass of lepton pair: **81 < mLL < 101 (GeV)**
- Cut 3: Same flavour leptons; opposite charge

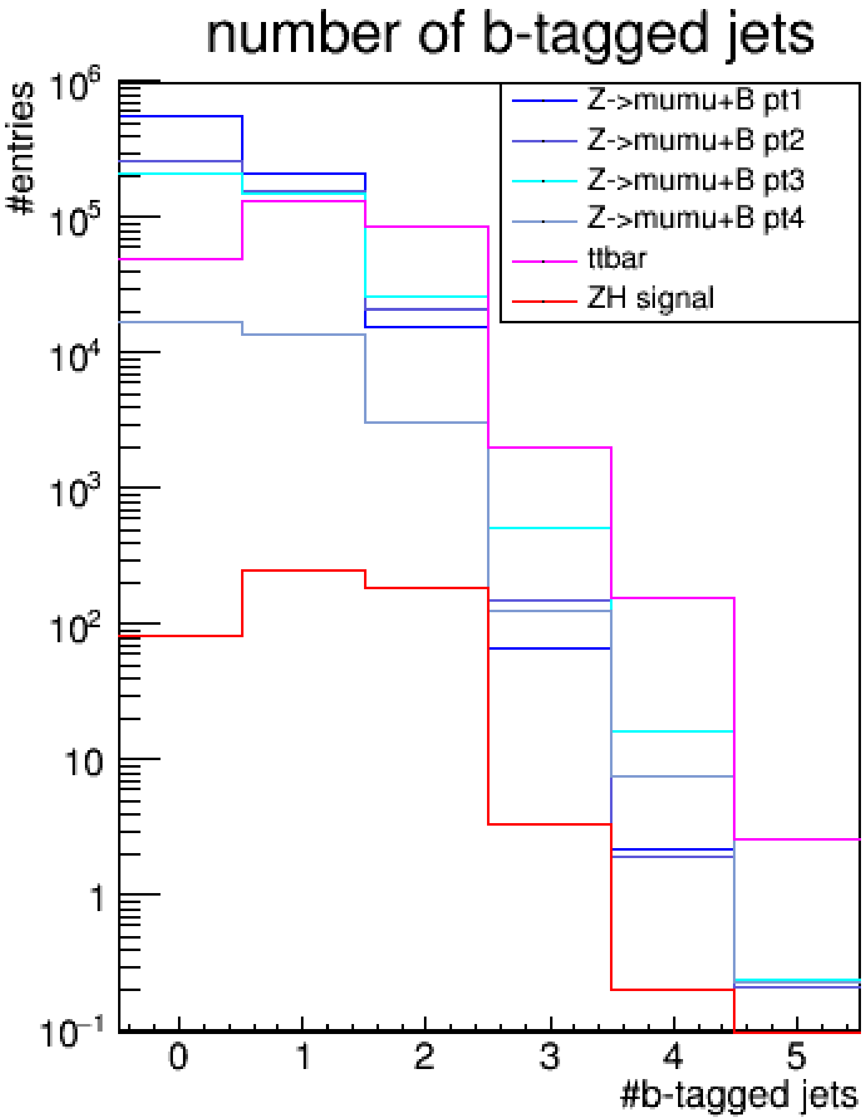
Event selection (Resolved)

- Cut 4: Number of Jets: **nJets >= 2**



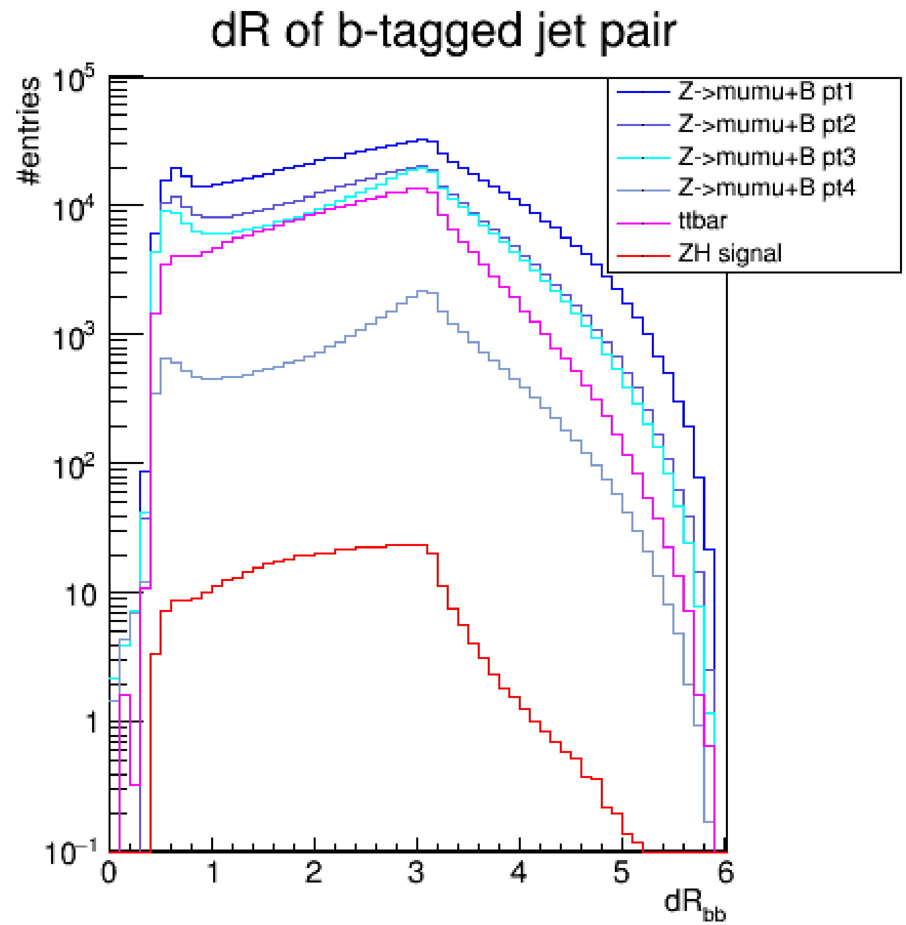
Event selection (Resolved)

- Cut 4: Number of Jets: **nJets >= 2**
- Cut 5: Number of b-tagged Jets: **nbJets = 2**

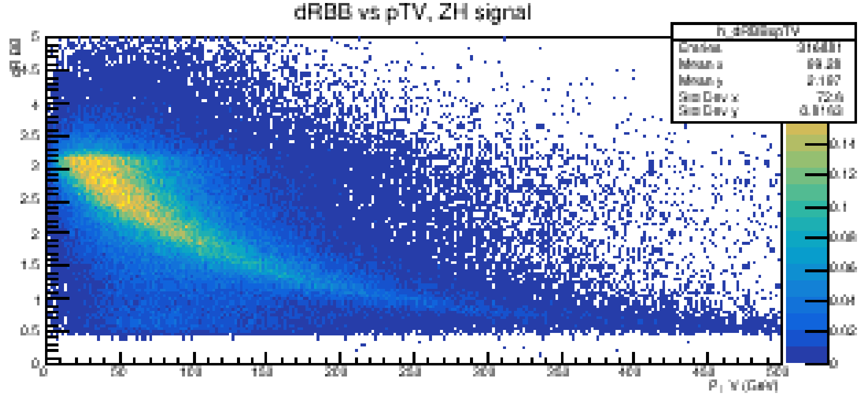


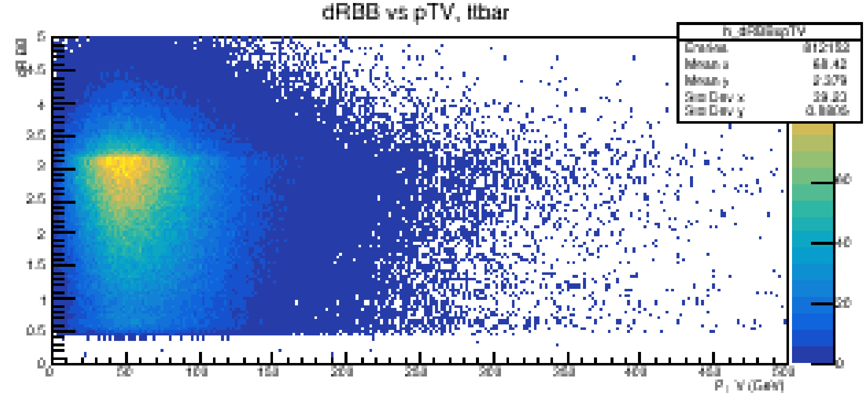
Event selection (Resolved)

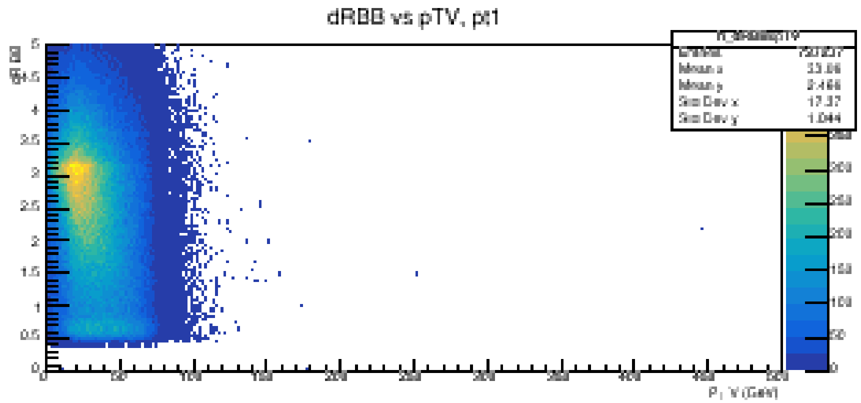
- Cut 4: Number of Jets: **nJets >= 2**
- Cut 5: Number of b-tagged Jets: **nbJets = 2**
- Cut 6: **dRBB** for **pTV** regions
  - pTV ≤ 75 GeV:
    - 1.7 < dRBB < 3.2;
  - 75 ≤ pTV ≤ 150 (GeV):
    - 1.1 < dRBB < 3.0;
  - 150 ≤ pTV ≤ 200 (GeV):
    - 0.9 < dRBB < 1.8;
  - 200 ≤ pTV (GeV):
    - dRBB < 1.2.

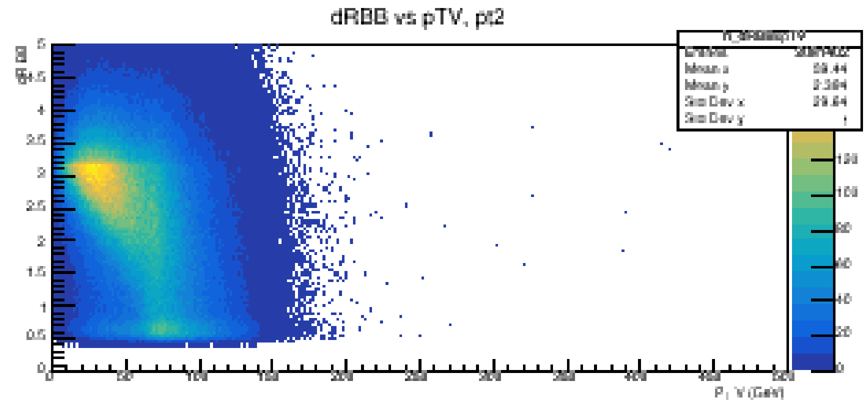


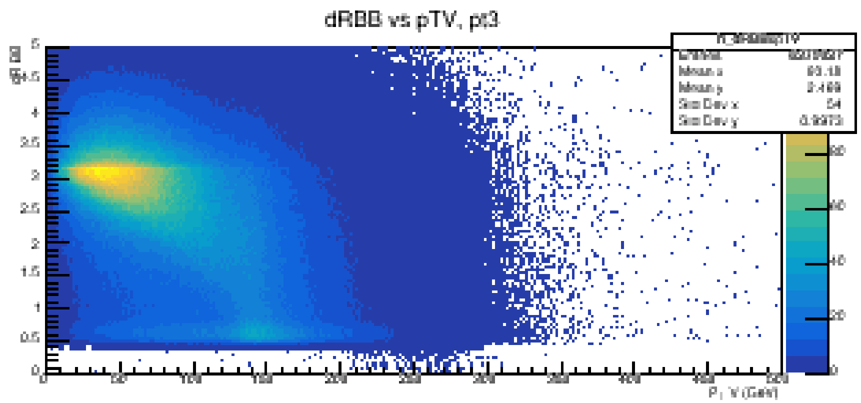
$$\Delta R = \sqrt{(\Delta\eta^2 + \Delta\varphi^2)}$$

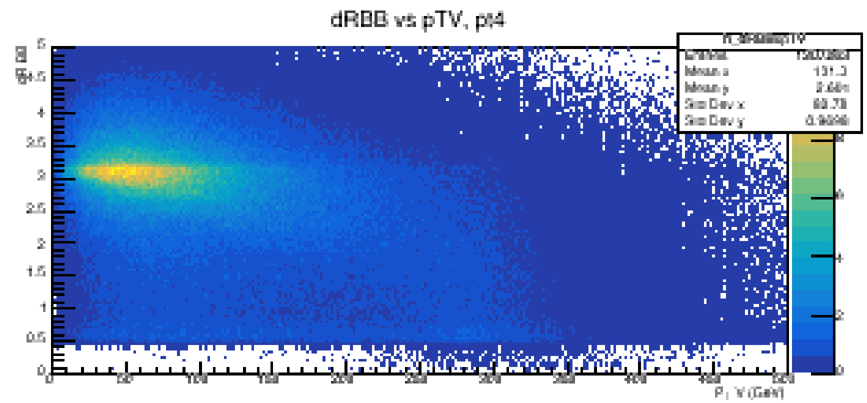

dRBB vs pTV, ZH signal


dRBB vs pTV, ttbar


dRBB vs pTV, pt1

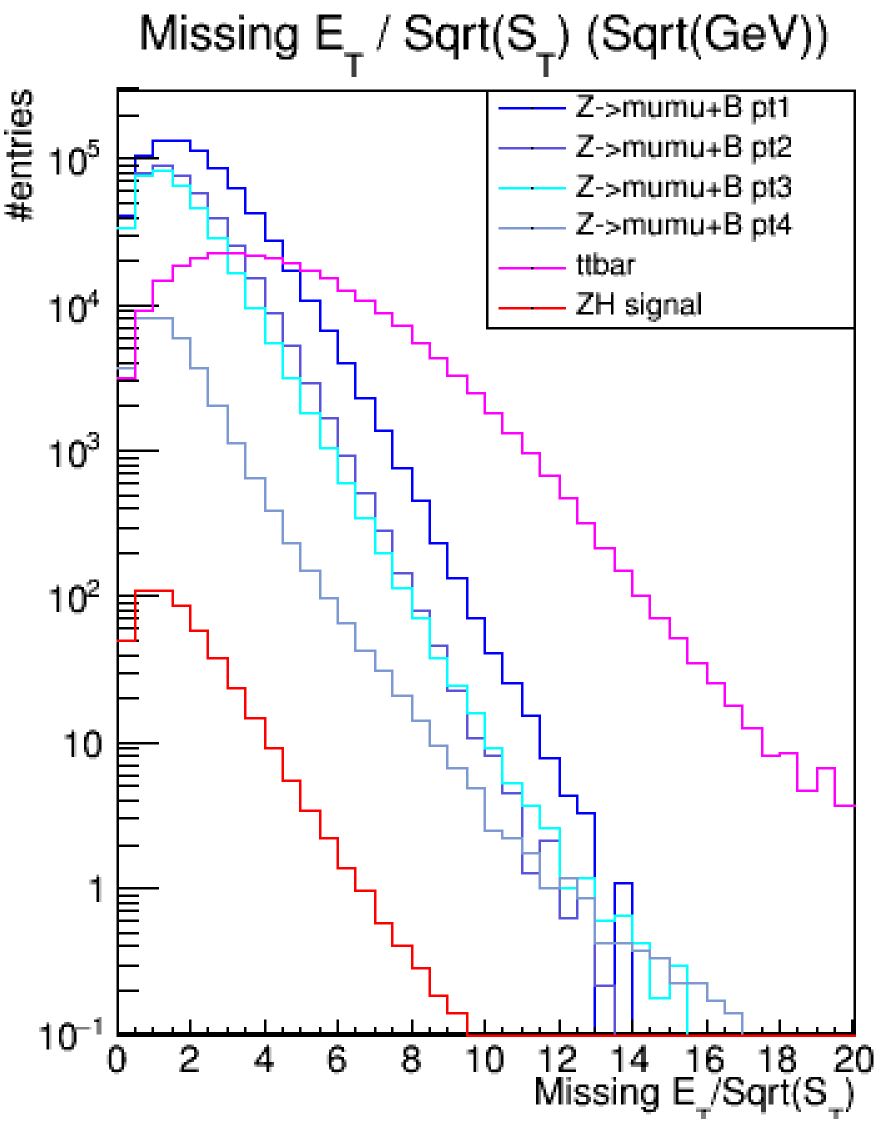

dRBB vs pTV, pt2


dRBB vs pTV, pt3


dRBB vs pTV, pt4

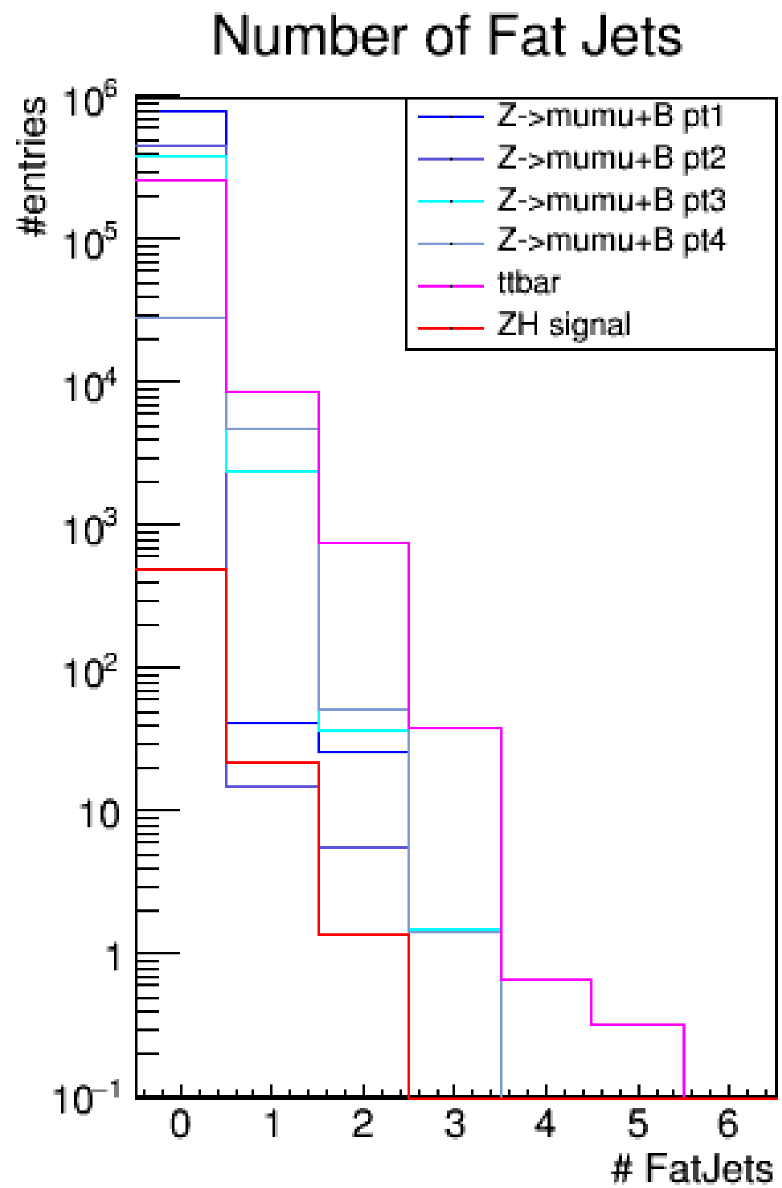
Event selection (Resolved)

- Cut 4: Number of Jets: **nJets >= 2**
- Cut 5: Number of b-tagged Jets: **nbJets = 2**
- Cut 6: **dRBB** for **pTV** regions
- Cut 7: “Significance” of Missing Transverse Energy relative to total transverse momentum: **METHT < 3.5 √GeV**



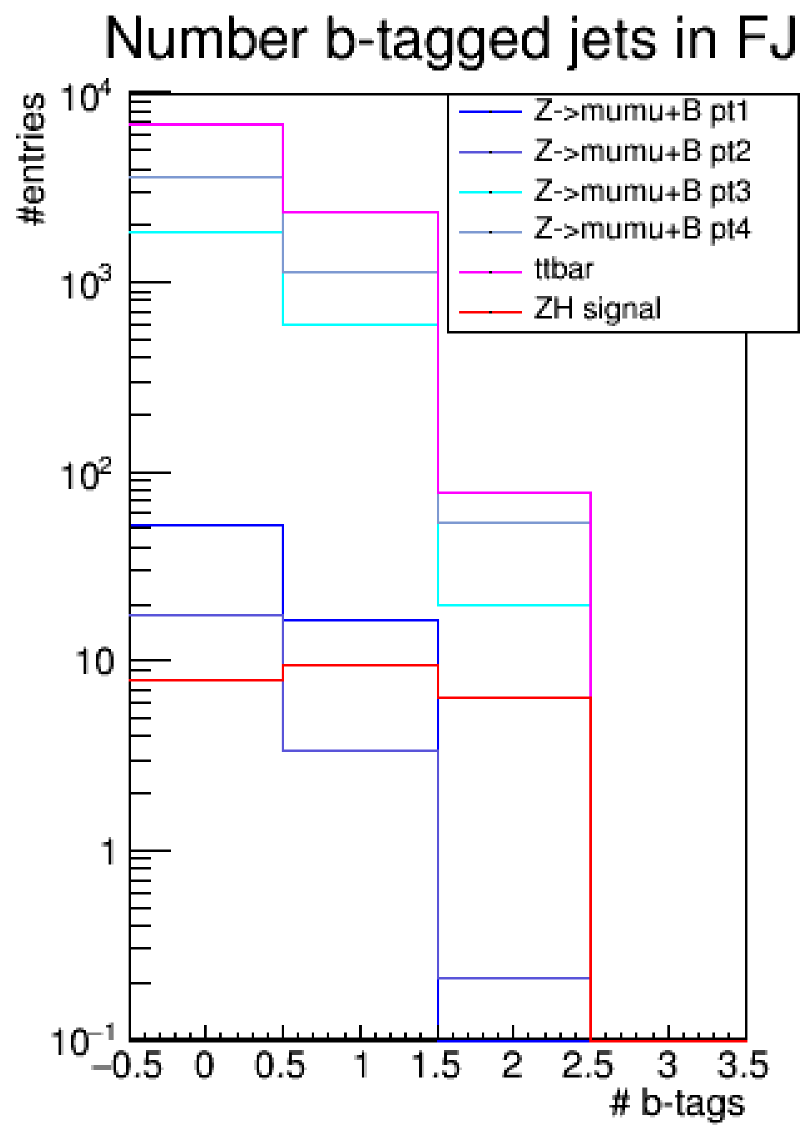
Event selection (Boosted)

- Cut 8: Number of Fat jets: **nFatJets = 1**



Event selection (Boosted)

- Cut 8: Number of Fat jets: **nFatJets = 1**
- Cut 9: Number of b-tagged subjets: **nbTagsInFJ = 2**



Significance

Significance = signal ⁄ √(backgrounds)

	ZH signal	ttbar	Z→μμ (+jets)	Z→ee (+jets)	ZZ→LL+bb	Signi.
Initial	520.29	266895	1676338.5	992280	34685.9	**0.30**
Resolved	127.31	1560.84	27211.79	17666	1196.87	**0.58**
Boosted	5.10	3.96	59.51	46.19	25.48	**0.44**

In the mass range [105, 145] (GeV)

	ZH signal	ttbar	Z→μμ (+jets)	Z→ee (+jets)	ZZ→LL+bb	Signi.
Initial	244.53	53964.9	258217.54	179965	4651.47	**0.35**
Resolved	89.87	395.22	7353.51	4610.71	159.17	**0.80**
Boosted	3.73	0.66	9.45	6.78	2.58	**0.85**

NOTE: Only main backgrounds considered, this is useful as a comparison between the boosted and resolved methods

Efficiency

- #of HZ events selected / total # of HZ events
- For a fair comparison of the efficiencies in Resolved and Boosted selections, cuts were restructured.

Z→ll cuts

Cut 1: pTL1 > 27 GeV

Cut 2: 81 < mLL < 101 (GeV)

Cut 3: same flav, diff. charge

Cut 4: MET<35 GeV

Resolved cuts

Cut 5: nJets>=2

Cut 6: nbJets=2

Cut 7: dRBB vs pTV

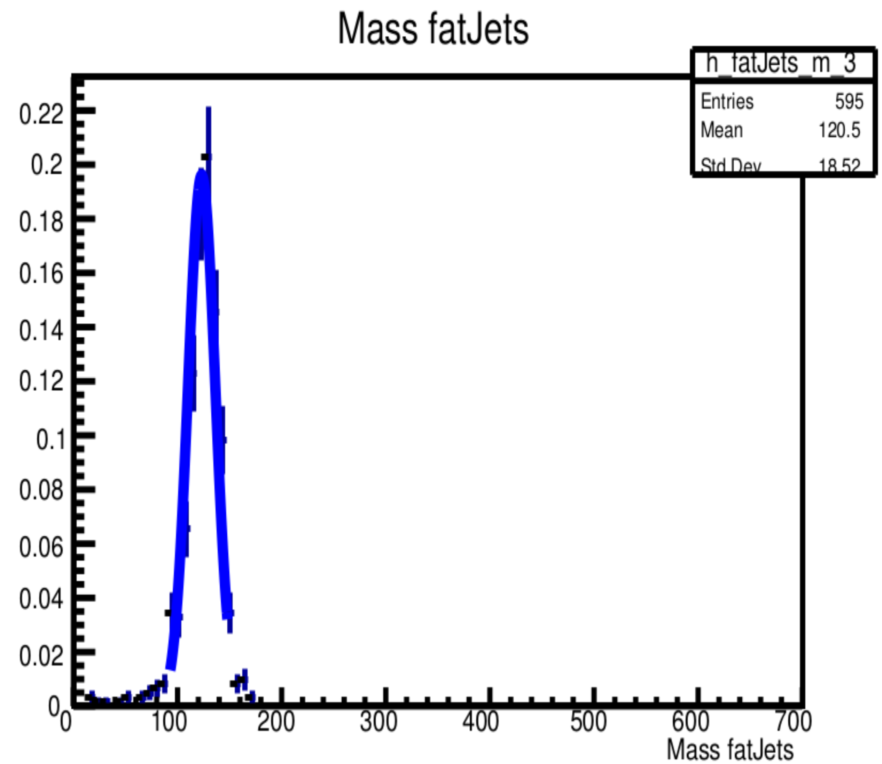
Boosted cuts

Cut 8: nFatJets = 1

Cut 9: nbTagsInFJ = 2

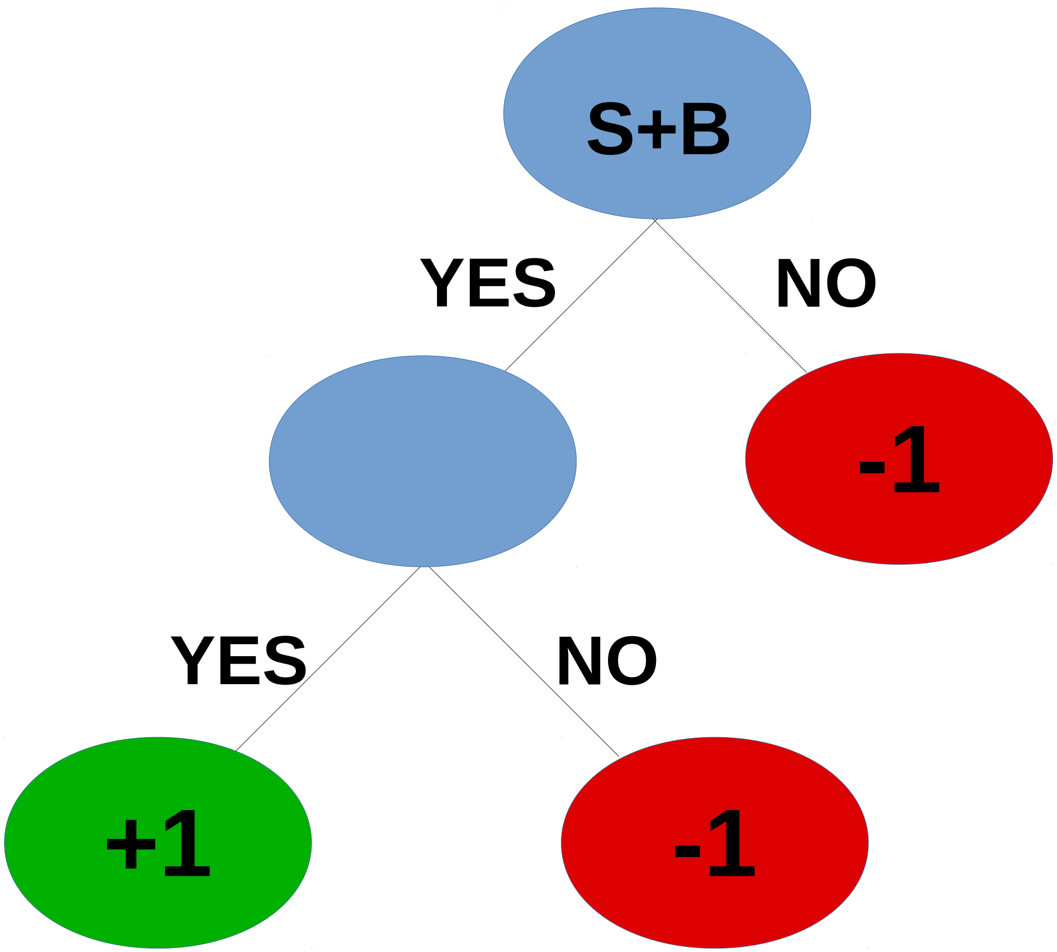
Scale and Resolution

- The **pTV** was divided into intervals and gaussian fits were made in order to calculate the resolution and the mean.
- Fitted variables:
  - **Resolved**: Invariant Mass b-jet pair (mbb)
  - **Boosted**: fatJets_m, fatJets_caloMass, fatJets_TAmass.

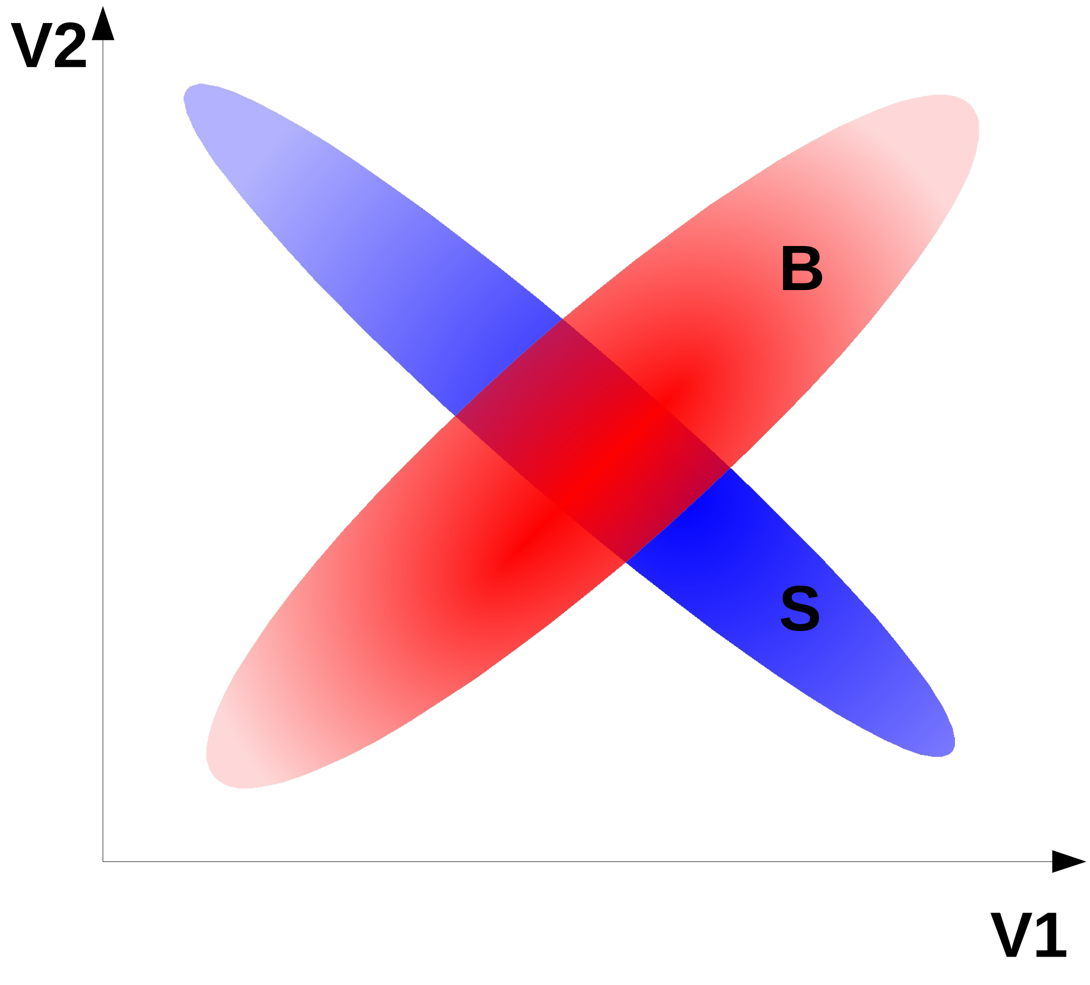


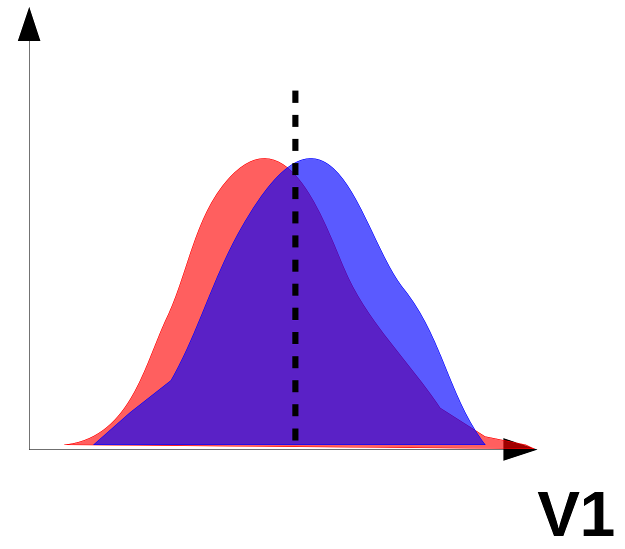
Multivariate Analysis

- A **Boosted Decision Tree (BDT)** is a tool that uses a set of input variables (and the correlation between them) to make selections on a sample of signal + background, with the goal of separating signal events from the background.

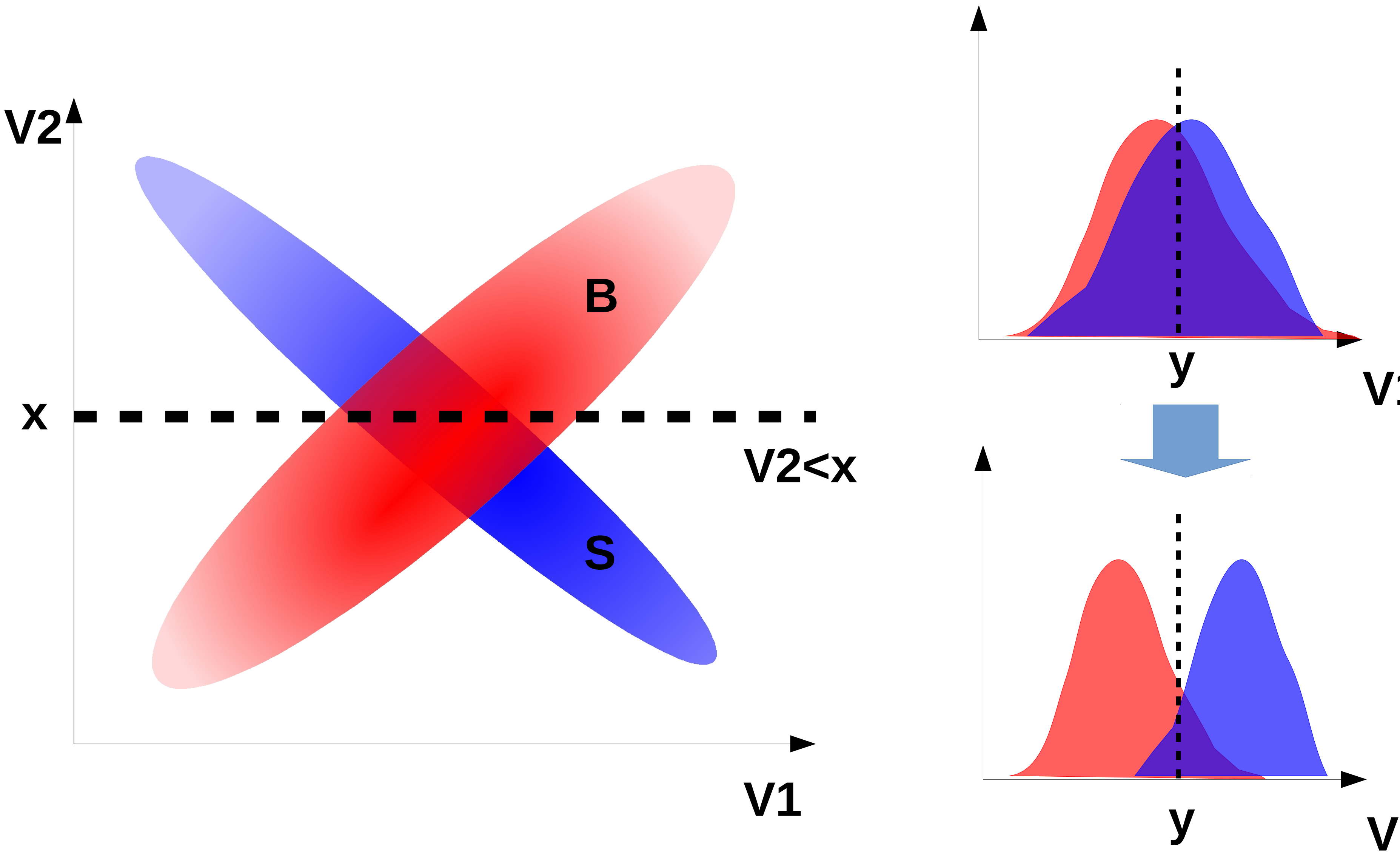


Multivariate Analysis





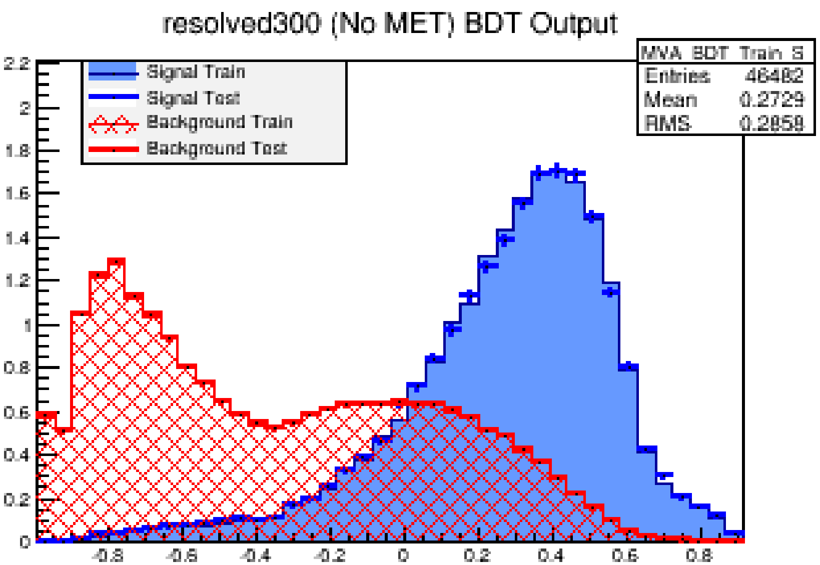
Multivariate Analysis

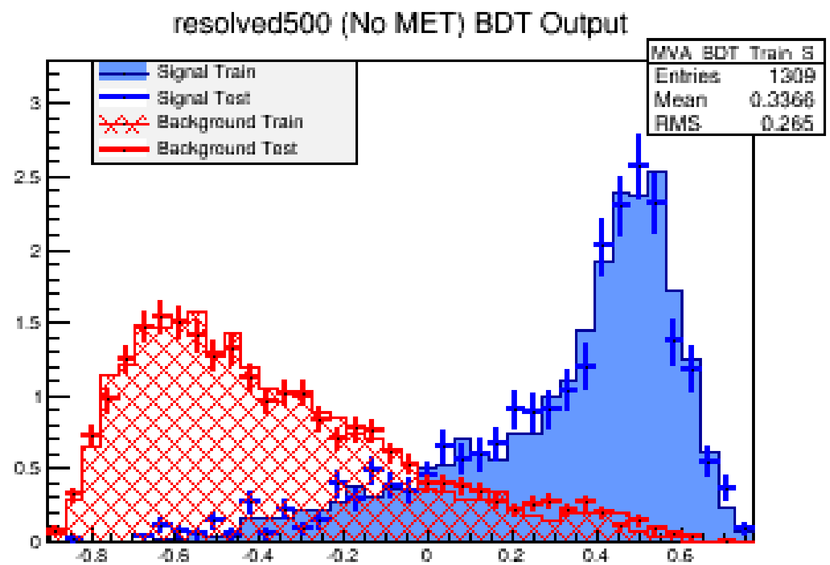


Multivariate Analysis

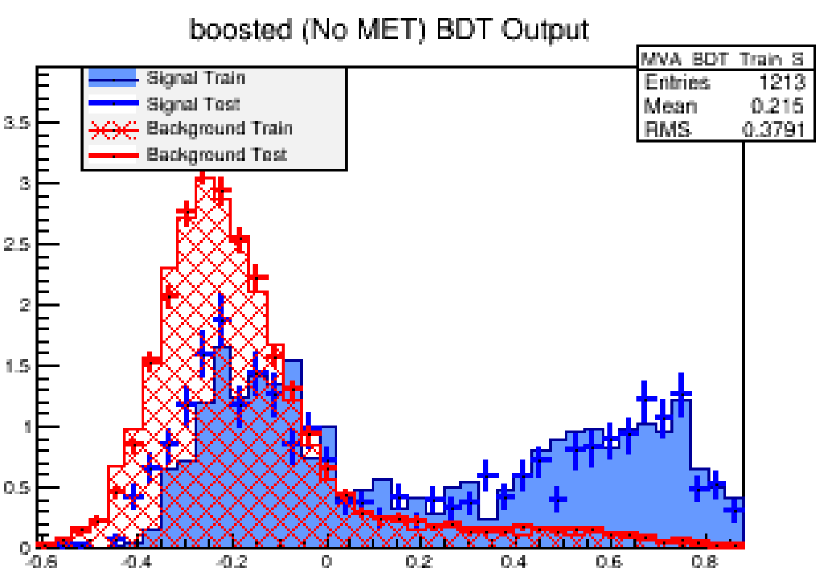
- Resolved, boosted and combined samples were given to the BDT with the cuts previously mentioned (except the **dRBB** and **MET** cuts). The following variables were used to train the BDT.
- The input variables for each regime follow:
  - **Resolved** *(pTV<300 e pTV ϵ [300,500] GeV)*
    - pTV, mLL, MET, dRBB, mBB, dPhiVBB, dEtaVBB, pTB1, pTB2;
  - **Boosted** *(pTV>300 GeV)*
    - pTV, mLL, MET, fatJets_C2, fatJets_D2, fatJets_Tau21, fatJets_m;
  - **Combined (***pTV ϵ [300,500] GeV)*
    - All variables.

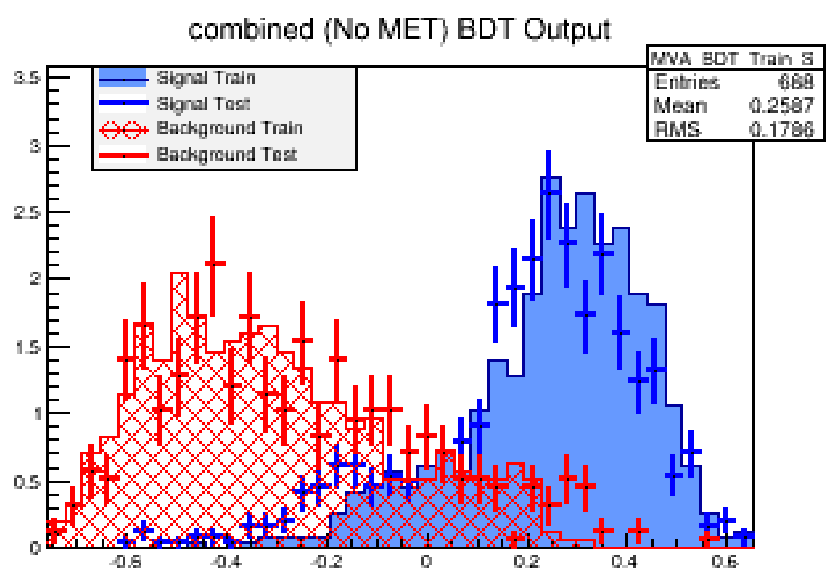
BDT Results



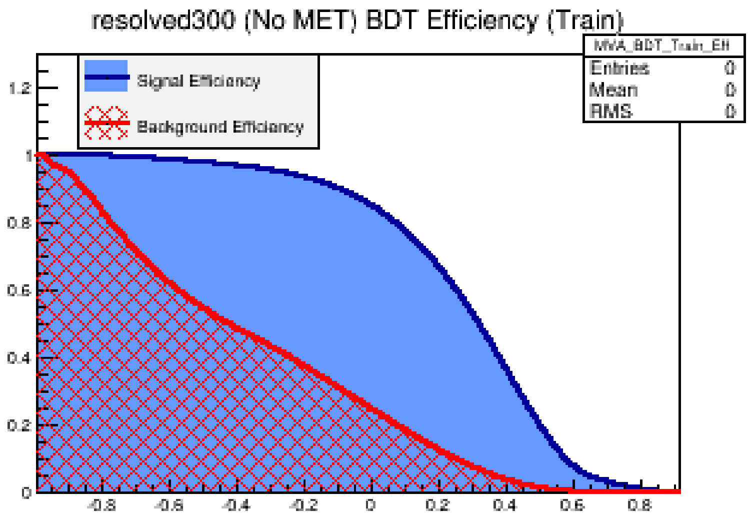


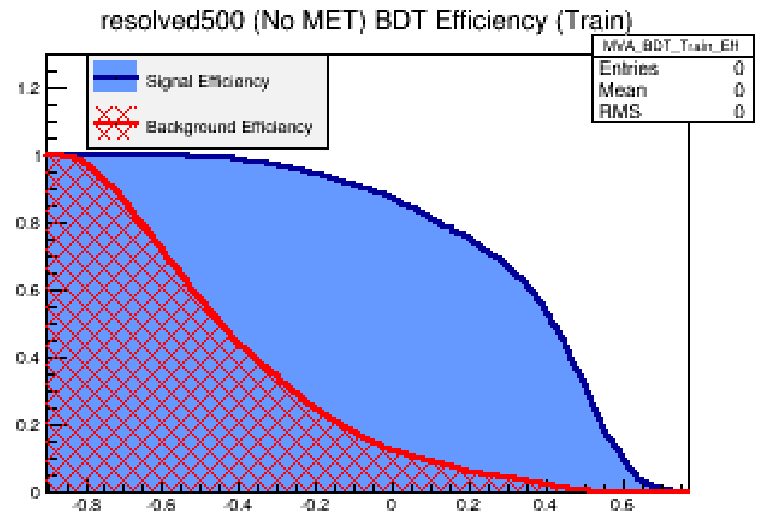
BDT test (points) and train (filled histogram) sample outputs for signal (blue) and background (red).

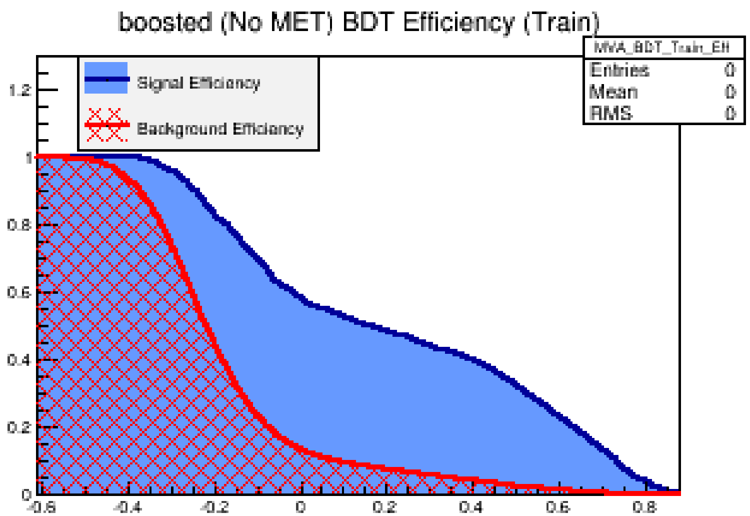


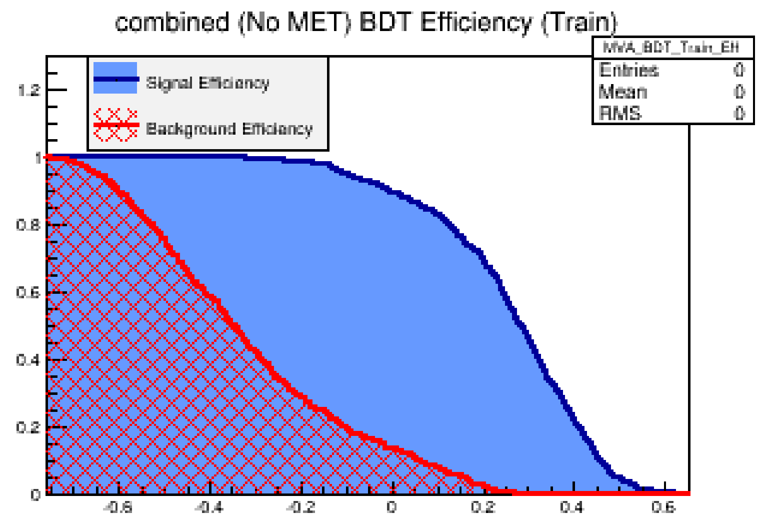


BDT Results



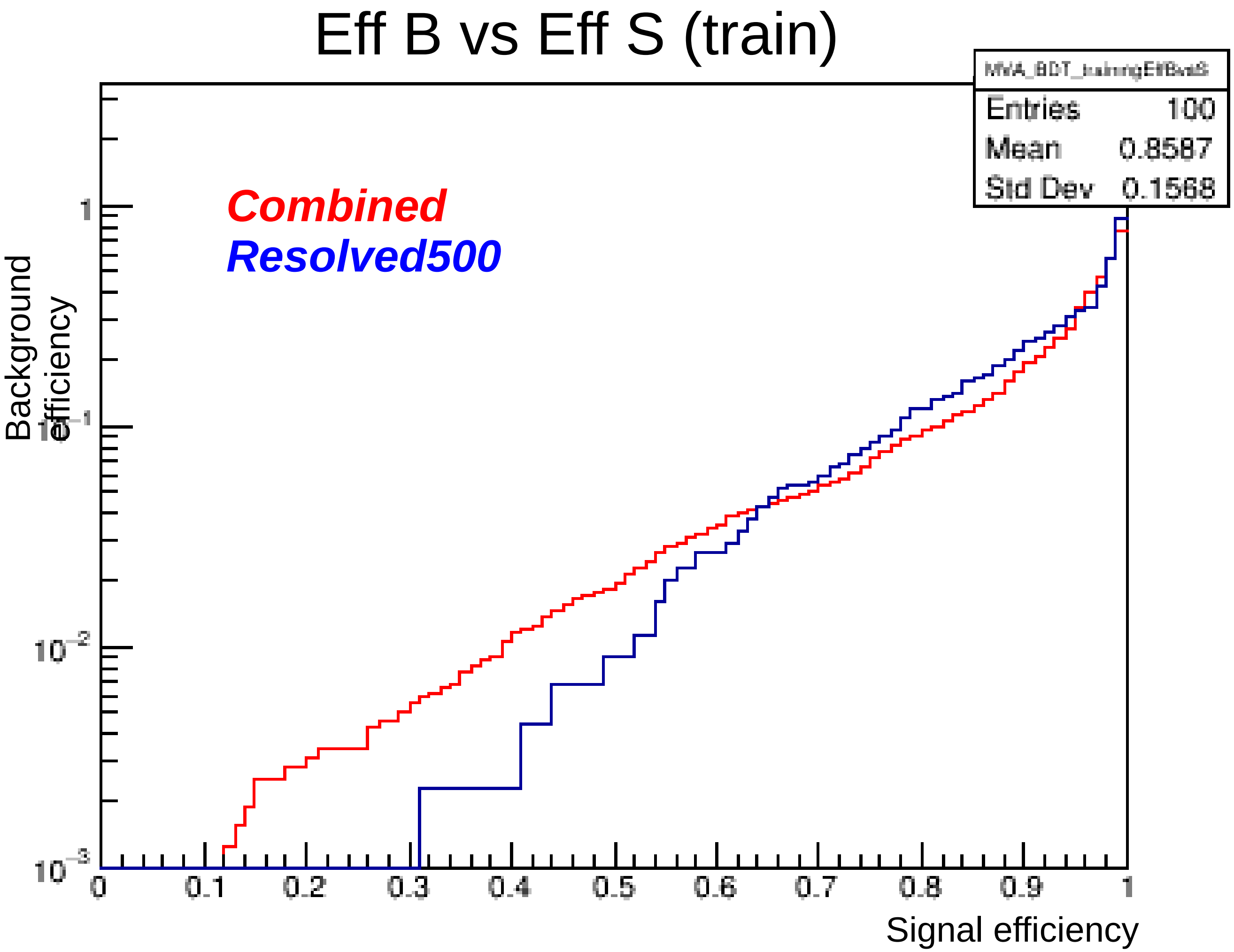






Efficiency of event selection (training) of signal (blue) and backgrounds (red)

BDT results



- Plot of the significance background efficiency vs the signal efficiency for combined (red) and resolved500 (blue).
- It appears that the combined method provides a better significance for the signal events, however uncertainties are large.

Conclusion

We looked to compare resolved and boosted analysis methods so as to refine our search for the Higgs boson and its properties. From our results we conclude that:

- For higher transverse momentum, the **boosted** analysis is **more efficient** in event reconstruction;
- Better energy calibration for fat jet reconstruction may improve the **mass scale** and **resolution** for the **boosted** analysis;
- **Combined** analysis may **improve the significance** within the pTV range [300,500] (GeV);
- Need more **high energy data** to provide greater clarity on the significance and resolution of the boosted analysis.