DARK MATTER SEARCH AT ATLAS EXPERIMENT

◆Mono-Higgs (H->bb)

◆Missing Transverse Energy Significance Improvement

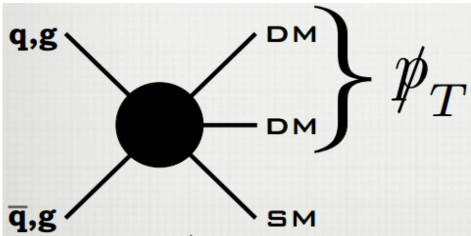
DILIA MARÍA PORTILLO

Biennale
07/10/2016

DARK MATTER SEARCH: MONO-HIGGS SIGNATURES

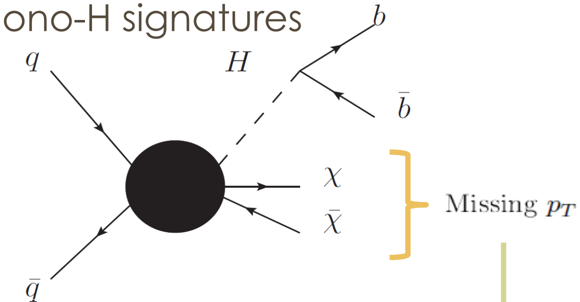
Dark matter as a WIMP → Mono-X:

- Standar Model (SM) particles ('X') recoils against missing transverse momentum



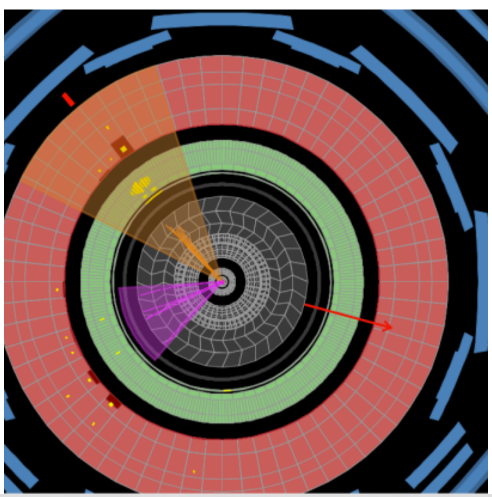
Mono-H Plan:

- Analysis of DM pair production in association with a H-> bb in the mono-H signatures

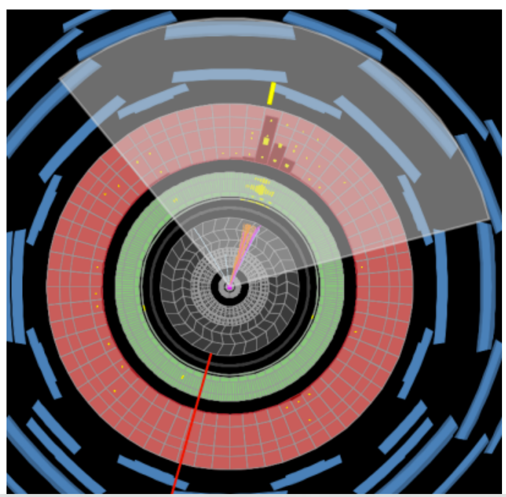


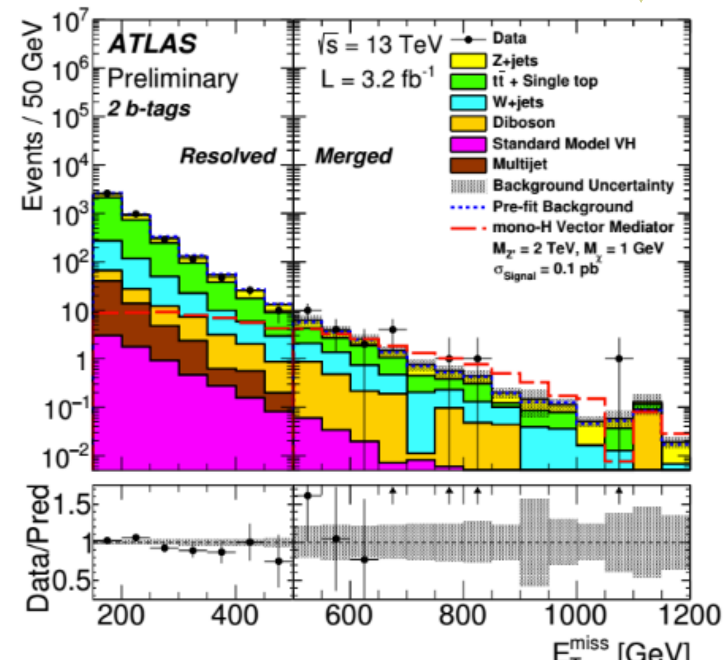
The decay of the Higgs is reconstructed as a high momentum bb system with either a pair of smalll radius jets or a single large radius jet.

Resolved Regime



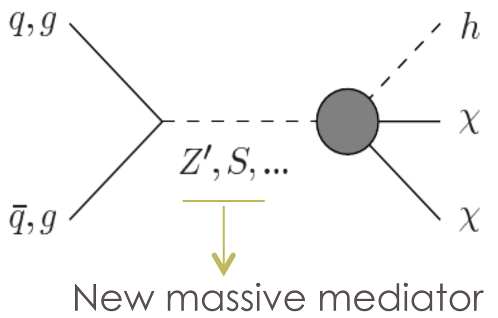
Merged Regime



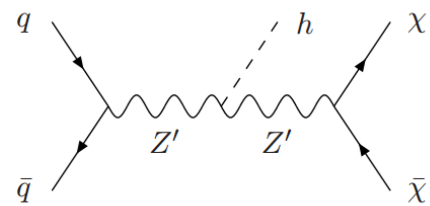


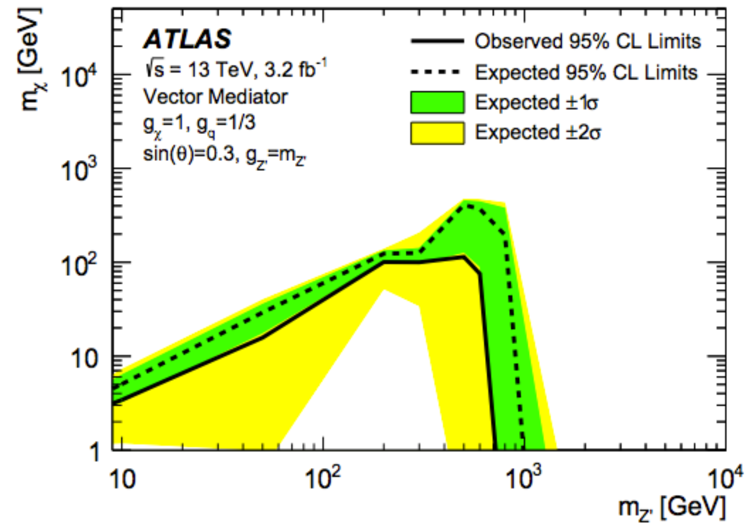
- https://indico.cern.ch/event/443590/contributions/1098915/attachments/1287774/1916321/Event_Displays.pdf
- I am part of the analysis team for the first first monoHbb paper ATLAS-CONF-2016-019

SIMPLIFIED BENCHMARK MODELS

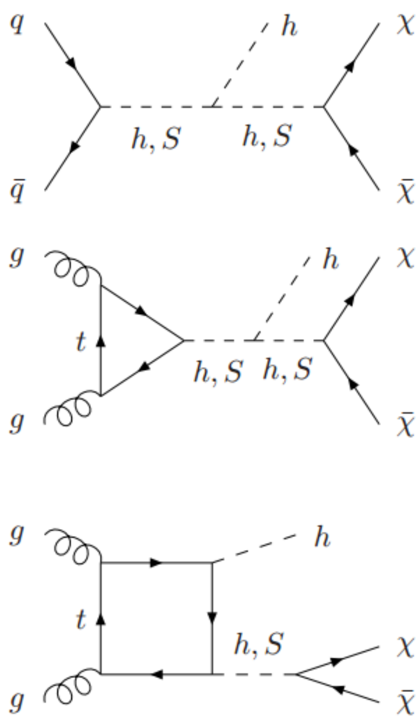


- Baryonic Z'

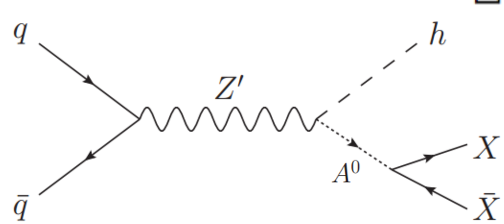


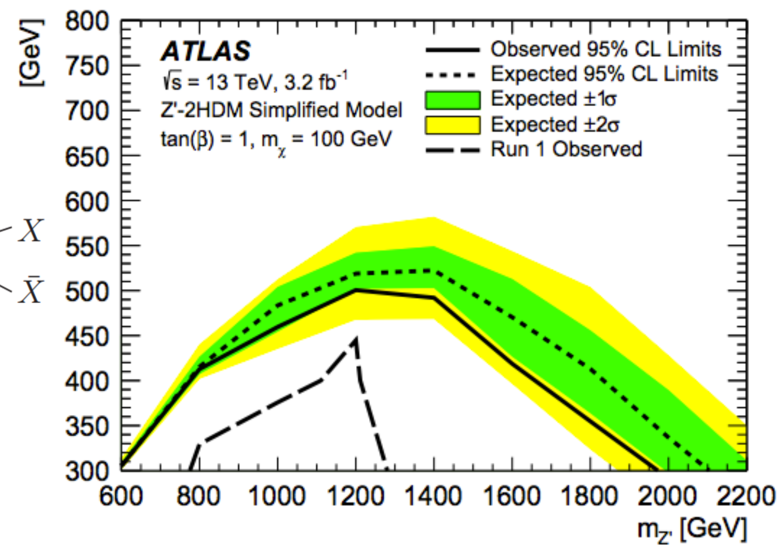


- Scalar Mediator



- Z'-2HDM





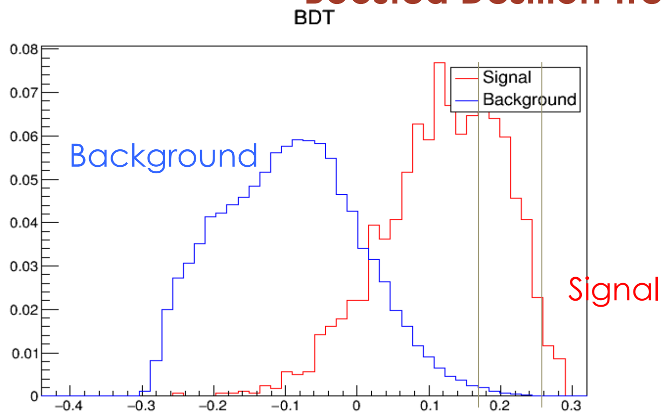
https://indico.cern.ch/event/443590/contributions/1098919/attachments/1193046/1732211/mono-Hmeeting_23-11-2015.pdf

POTENCIAL IMPROVEMENT IN THE MONO-H (H->bb) SEARCH

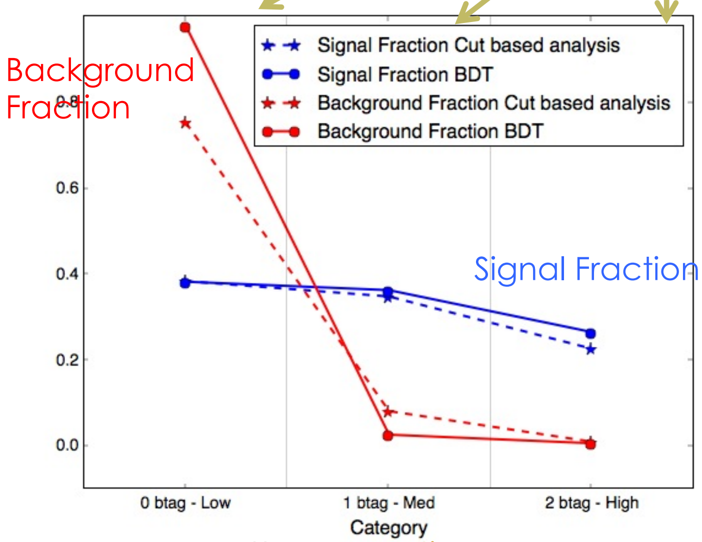
Multi Variate Analisis on Scalar Mediator Model

Plan: Study a multivariate approach for mono-Higgs search and compare the potential improvement to the current cut-based analysis.

Boosted Desition Tree (BDT) on Merged Regime (MET > 500 GeV)



- In the training of the BDT we considered kinematic variables, variables related to the Anti-QCD cuts and b-tag information.
- BDT distribution is divided in categories in order to compare w.r.t. cut based analysis categories.



Better separation power.

	Cuts Based Analysis			BDT		
	0 btag	1btag	2 btag	Low	Medium	High
Signal Significance	0.039	0.105	0.178	0.034	0.177	0.23
Combined Sign.	0.2104368			0.2929455		

	Low	Medium	High	Global Improvement
Improvement	-12.40%	68.80%	29.50%	39.20%

This potential was presented at the mono-W/Z/H hadronic meeting https://indico.cern.ch/event/443590/contributions/2175810/attachments/1277340/1895734/MVA_Mono-H_14-05-2016.pdf

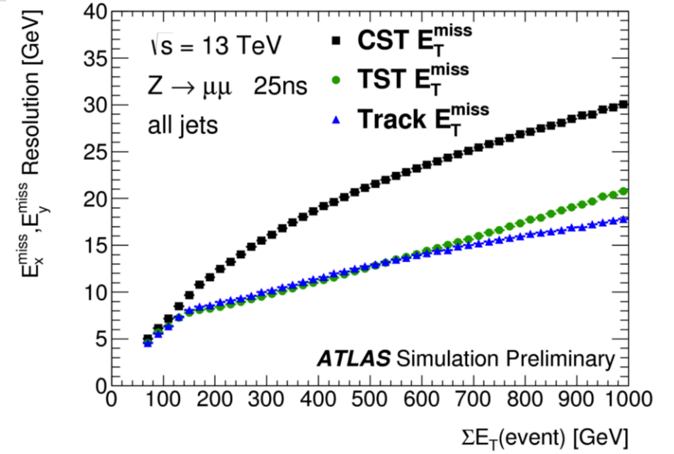
MISSING TRANSVERSE ENERGY SIGNIFICANCE IMPROVEMENT
ATLAS QUALIFICATION TASK

Events in which the reconstructed Met is either consistent with contributions solely from particle-measurement resolutions and efficiencies or consistent with genuine Met can be identified by evaluating the Met Significance S.

Met Significance is currently defined as:

$$S = \frac{\text{Met}}{\sqrt{\text{Ht}}}$$

Ht: Proxie for Met error. Event based quantities and correlations are not entering in the calculation.

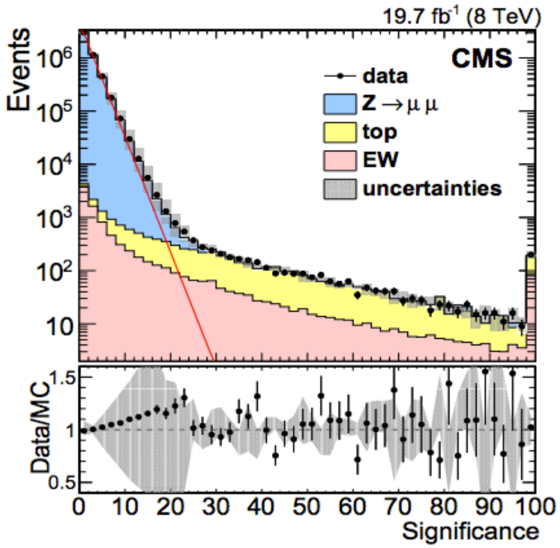


Object Based Met Significance

How likely is it that this METmeas is TRUE MET, and not simply a result of measurement error or other effects?

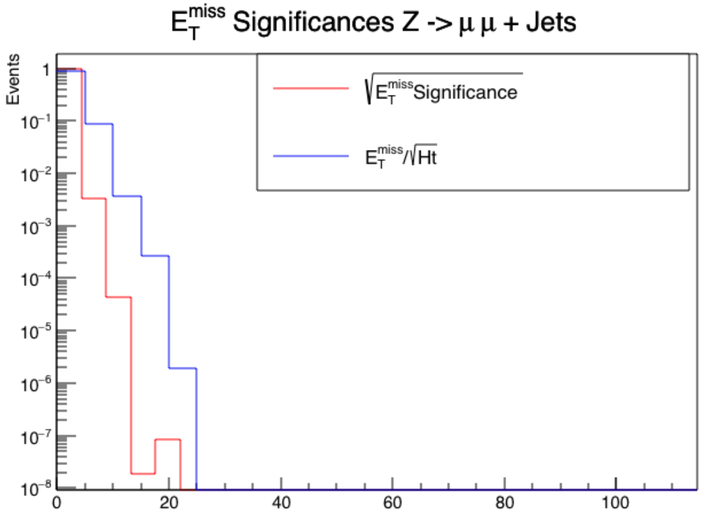
$$S(\vec{E}) \equiv 2ln\left(\frac{\mathcal{L}(\vec{E}=\sum_i \vec{E_{T_i}})}{\mathcal{L}(\vec{E}=0)}\right) \rightarrow S \sim \left(\sum_i \vec{E_{T_i}}\right)^{\dagger}\left(\sum_i V_i\right)^{-1}\left(\sum_i \vec{E_{T_i}}\right)$$

The log-likelihood ratio definition of the significance becomes a χ^2 variable



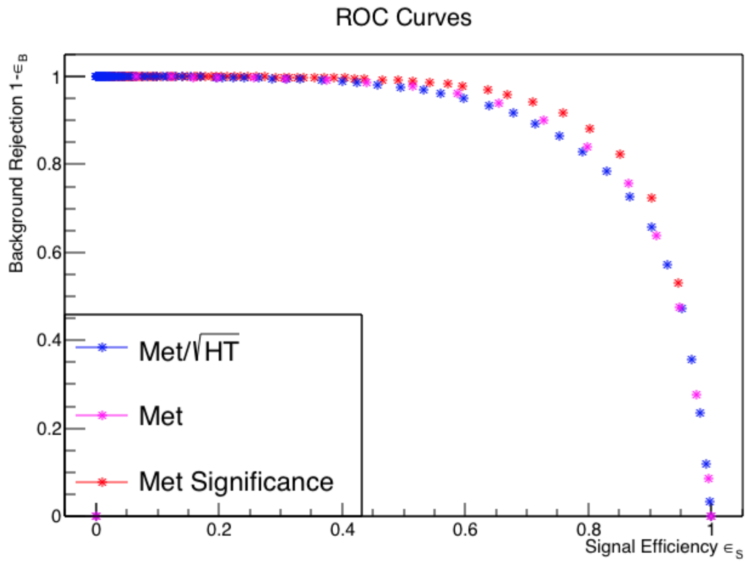
First implementation of a Object based Met Significance

- Z-> $\mu\mu$ + jets (no genuine Met)
- Considering
  - Jet resolutions
  - A constant soft term resolution as an approximation.



First test of Met significances Separation power

- Z($\mu\mu$)+jet (Bakground) and ZZ($\mu\mu\nu\nu$)+jet (Signal)



Met Significance with a constant resolution of the soft term has a better performance compare w.r.t. Met/Sqrt(Ht)

This is our performance starting point. We are missing…

- Properly implement the Soft Term resolutions
- Other objects resolutions

THANKS

BACKUP SLIDES

MVA STRATEGY ON SCALAR MEDIATOR MODEL

Study a multivariate approach for mono-Higgs search and compare the potential improvement to the current cut-based analysis. In this presentation the first case study for the simplified scalar model. MVA approach will be later optimized for other simplified models and for mediator/wimp masses.

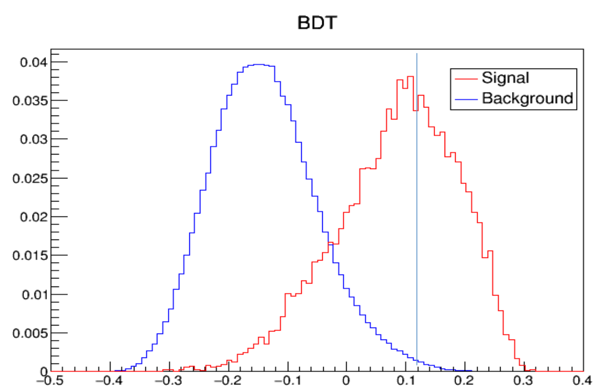
BDT ON THE SCALAR MEDIATOR S.M

In the training of the BDT we considered the Mbb, the Pt, eta of the bb system, the variables related to the AntiQCD cuts and the number of btag jets, fat jets, addicional jets, Antikt4 jets and the MET and MPT.

ALL MET RANGE

The BDT method was trained for all the MET range in order to take advantage of the whole statistics.

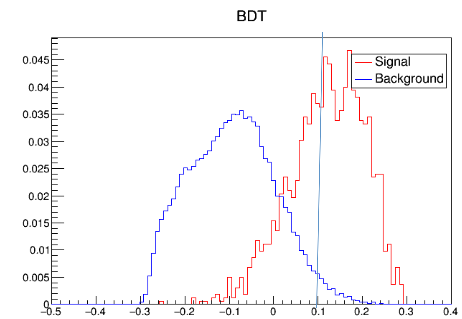
- **Classifier Output Distribution**



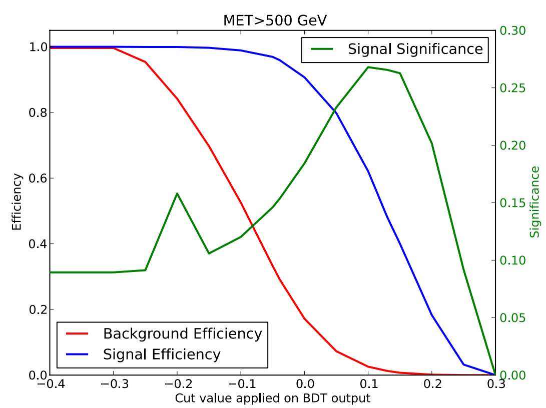
MET > 500 GeV

We are interested in the high MET category .

- **Classifier Output Distribution**

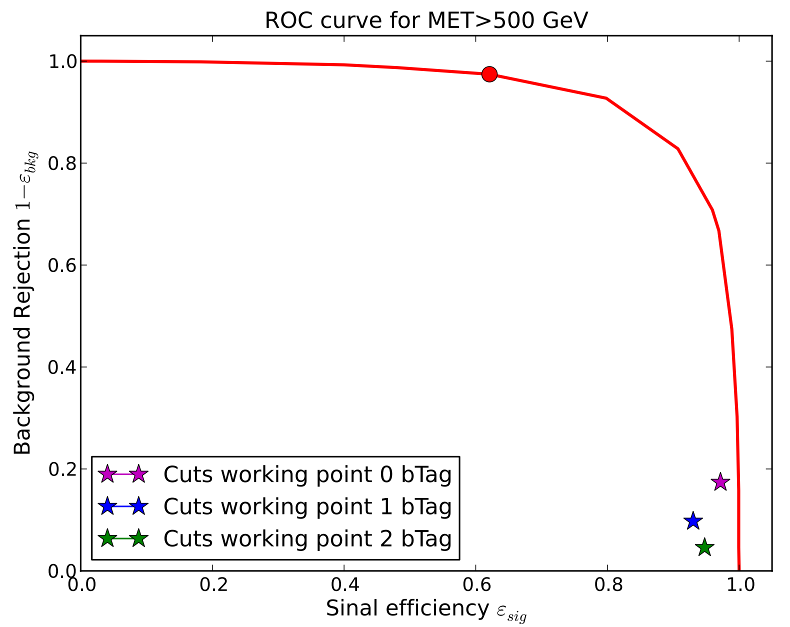


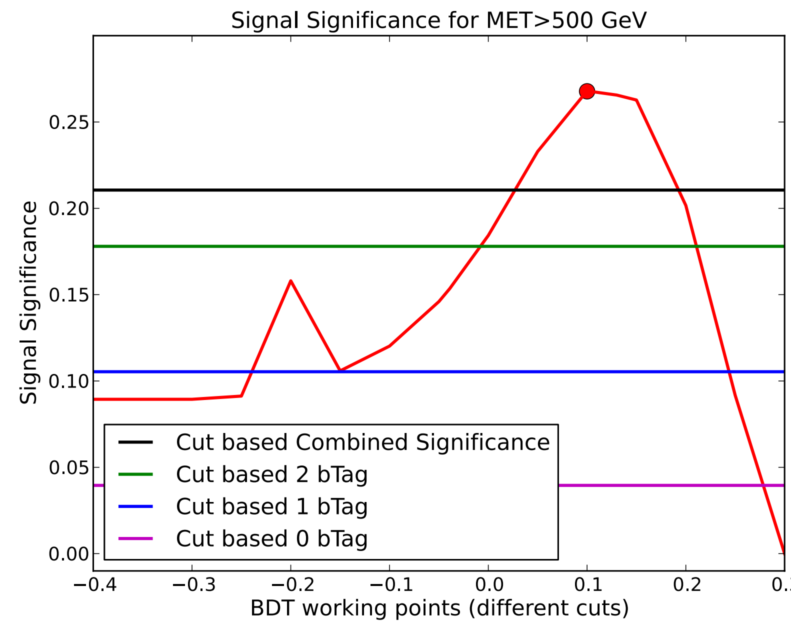
- **Classifier Cut efficiencies**



BDT USED AS SINGLE CUT VS CUT BASED BJET CATEGORIES FOR MET > 500 GEV

Comparison in btag categories for the merged channel





	500 < MET			
Signal Events *	0.27755			
Background Events *	9.37206			
	Cut Based Analysis			BDT
	0 b tag	1btag	2 btag	BDT > 0.1
Signal Events after cutting *	0.105569	0.0956614	0.0618382	0.172337
Background Events after cutting *	7.04099	0.730347	0.058999	0.241865
Signal efficiency	0.971698	0.930233	0.947644	0.620922
Background rejection	0.173781	0.0974908	0.0457236	0.974193
Signal significance	0.03953	0.10536	0.177984	0.267777
Combined significance	**0.210575**			**0.267777**

- Combined Significance

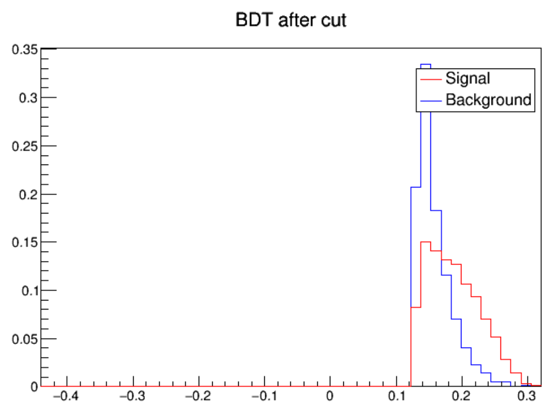
$$S = \sqrt{S^2_{0-btag} + S^2_{1-btag} + S^2_{2-btag}}$$

Significance improvement is +28%

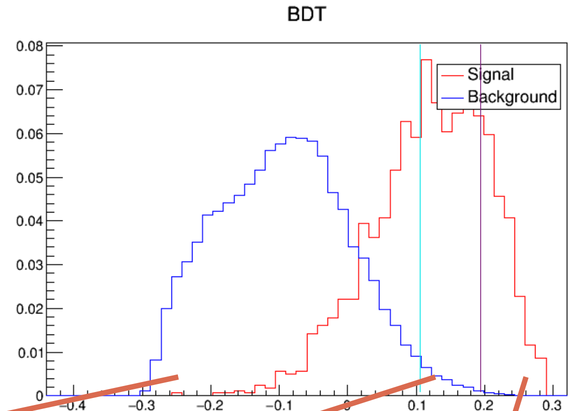
B-TAG JETS VS BDT CATEGORIES

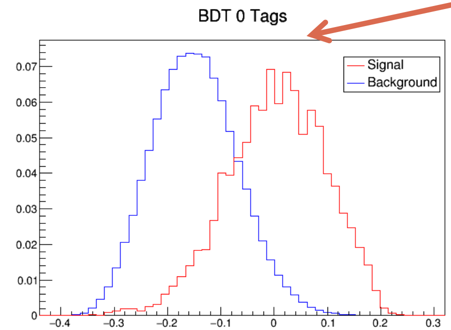
The number of b-tag jets is a variable considered in the BDT training.

Then, instead of having a single sharp cut on BDT

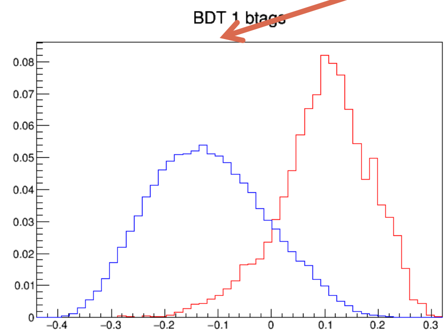


We can define BDT categories that can be related to 0, 1 and 2 btags.

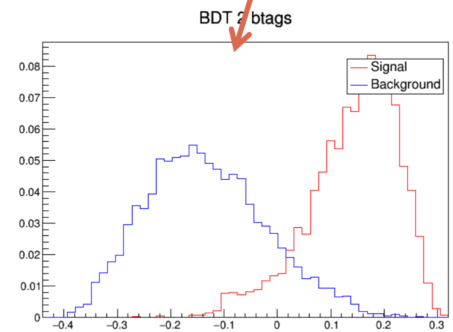




- ✧ Low BDT category



- ✧ Medium BDT category



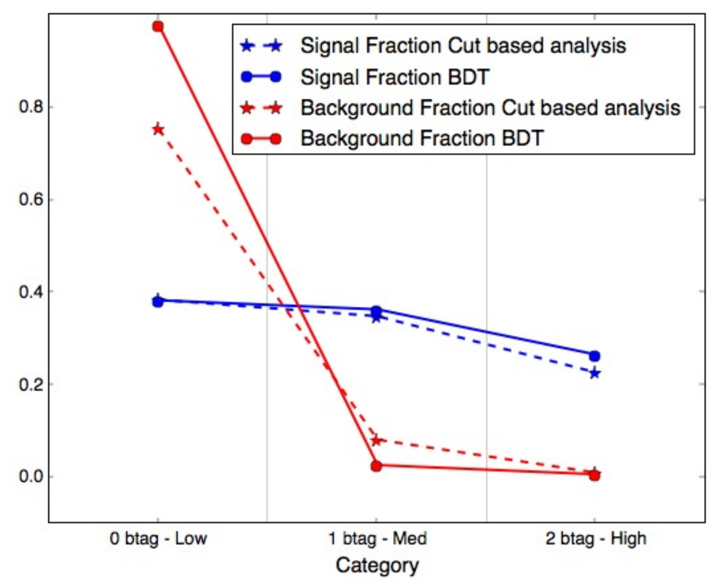
- ✧ High BDT category

Low – medium – high BDT categories are defined by maximizing the combined significance at first order.

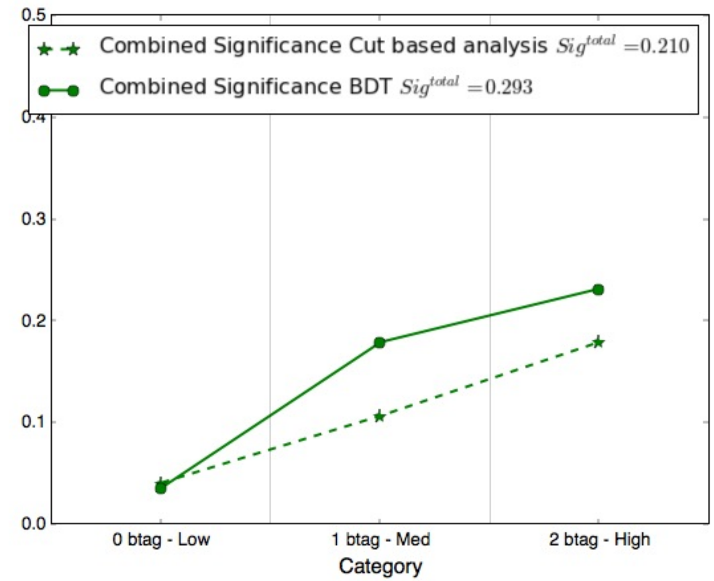
COMPARISON BETWEEN BDT CLASSIFIER AND CUTS BASED ANALYSIS FOR MET >500 GEV

Now we can do a better comparison between the cut based analysis in btag categories and the BDT classifier in low-medium-high categories.

BDT achieved to have more signal-like events in the best category (High BDT) and more background-like events in the worst category (Low BDT) meaning a better separation power.



For Medium (related to 1 btag) and High (2 btag) categories we obtain an approx. improvement of +69% and +30% respectively. And a Global improvement of +39% in Signal Significance.

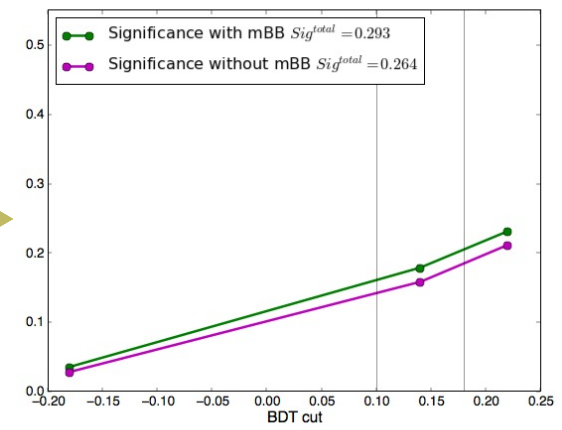


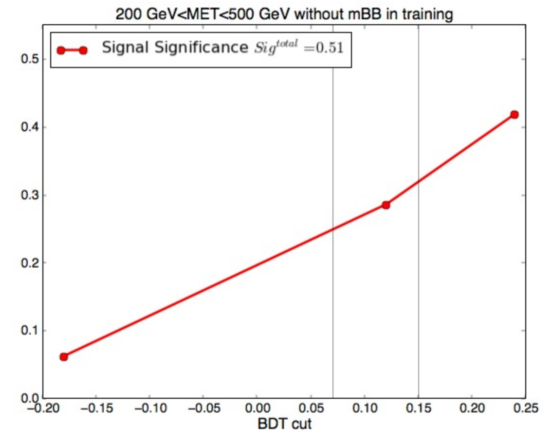
	Cuts Based Analysis			BDT			
	0 btag	1btag	2 btag	Low	Medium	High	
Signal Events Expected*	0.105569	0.0956614	0.0618382	0.105227	0.0997611	0.0726001	
Background Events Expected *	7.04099	0.730347	0.058999	9.1536	0.215417	0.0267651	
Purity	0.01499348813	0.130980753	1.04812285	0.01149569568	0.4631069043	2.712491267	
Signal Significance (improvement)	0.0394901	0.105255	0.177892	0.0345821(-12.4%)	0.177698 (+68.8%)	0.230314(+29.5%)	Global Improvement
Combined Significance	0.2104368			0.2929455			39.20%

MONO-HIGGS MULTIVARIATE SEARCH PLANS

- Implement the BDT training for the resolved regime on the scalar mediator model with a very light DM particle.
- Study the possibility of not considering the mBB variable in the training so we can have the possibility of fitting it with the BDT.
- Revisit the merged-resolved separation with MVA approach.
- Extrapolate this results for other masses of the DM and mediator particles in this simplyfied model.

- Perform a MVA optimization for
  - Z' mediator
  - Z'+2HDM





This potential improvement in signal significance for the scalar simplified model was presented at the mono-W/Z/H hadronic meeting (23 May 2016) https://indico.cern.ch/event/443590/contributions/2175810/attachments/1277340/1895734/MVA_Mono-H_14-05-2016.pdf

QUALIFICATION TASK: MET SIGNIFICANCE DEFINITION

- Events in which the reconstructed Met is either consistent with contributions solely from particle-measurement resolutions and efficiencies or consistent with genuine Met can be identified by evaluating the Met Significance S.

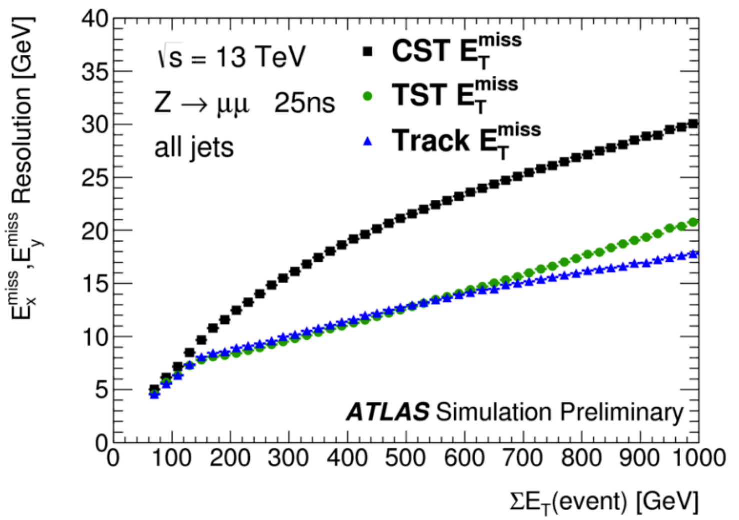
- A high value of S is an indication that the observed Met is not well explained by resolution smearing alone, suggestions that the event may contain unseen objects such as neutrinos or more exotic weakly interacting particles.

Met Significance is currently defined as:

$$S = \frac{\mathrm{Met}}{\sqrt{\mathrm{Ht}}} \qquad S = \frac{\mathrm{Met}}{\sqrt{\mathrm{Sumet}}}$$

Where *Sumet* and *Ht* are used as proxies for Met error. These are event based quantities and correlations are not entering in the calculation.

We want a MET significance that is based on the uncertainties for all objects that enter the calculation of MET.



OBJECT BASED MET SIGNIFICANCE

The DØ-CMS approach

- How likely is it that this METmeas is TRUE MET, and not simply a result of measurement error or other effects?

This can be evaluated with the log-likelihood ratio of measuring the total observed transverse momentum to the likelihood of the null hypothesis.

$$S(\vec{E}) \equiv 2ln\left(\frac{\mathcal{L}(\vec{E} = \sum_i \overrightarrow{E_{T_i}})}{\mathcal{L}(\vec{E} = 0)}\right)$$

On a event-by-event basis, *S* evaluates the p-value that the observed MET is consistent with a null hypothesis, given the full event composition.

If we assume that …

- The sum of all the truth transverse momentum is equal to zero $\sum \overrightarrow{e_{T_i}} = 0$
- The difference $\vec{\epsilon_i} = \overrightarrow{E_{T_i}} - \overrightarrow{e_{T_i}}$ has a gaussian probability density function

$$p_i(\vec{\epsilon_i} \mid \overrightarrow{e_{T_i}}) \equiv P_i(\vec{\epsilon_i} + \overrightarrow{e_{T_i}} \mid \overrightarrow{e_{T_i}}) = P_i(\overrightarrow{E_{T_i}} \mid \overrightarrow{e_{T_i}}) \sim exp\left[-\frac{1}{2}\left(\overrightarrow{\epsilon_{T_i}}\right)^{\dagger} V_i^{-1}\left(\overrightarrow{\epsilon_{T_i}}\right)\right]$$

…the likelihood for two objects is given by:

$$\mathcal{L}(\vec{E}) = \int P_1(\overrightarrow{E_{T_1}} \mid \overrightarrow{e_{T_1}}) P_2(\overrightarrow{E_{T_2}} \mid \overrightarrow{e_{T_2}})\, \delta\left(\vec{E} - (\overrightarrow{E_{T_1}} + \overrightarrow{E_{T_2}})\right) d\overrightarrow{E_{T_1}} d\overrightarrow{E_{T_2}}$$

$$\mathcal{L}(\vec{E}) \sim exp\left[-\frac{1}{2}\left(\vec{E}\right)^{\dagger}\left(\sum_i V_i\right)^{-1}\left(\vec{E}\right)\right]$$

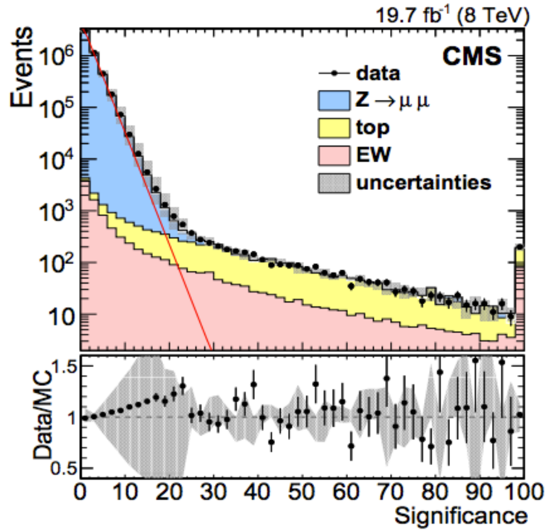
OBJECT BASED MET SIGNIFICANCE

Significance Definition

$$S = 2ln\left(\frac{\mathcal{L}(\overrightarrow{E} = \sum_i \overrightarrow{\epsilon_{T_i}})}{\mathcal{L}(\overrightarrow{E} = 0)}\right)$$

The log-likelihood ratio definition of the significance becomes a χ^2 variable for gaussian uncertainties distributions:

$$S \sim \left(\sum_i \overrightarrow{E_{T_i}}\right)\left(\sum_i V_i\right)^{-1}\left(\sum_i \overrightarrow{E_{T_i}}\right)$$



The observed spectrum conforms to a χ^2 distribution in the core region (red line).

Plan

- Implement this definition in ATLAS and evaluate if it will actually improve analysis needing met significance

Also, this definition can be also improved considering…

- That the sum of the truth missing energy is equal to a scale of energy
- Parameterize the uncertainties distributions in the likelihood without the assumption of gaussian distributions

DEFINITION OF OBJECT BASED MET SIGNIFICANCE

Covariance Matrix

- The significance is defined as the log-likelihood ratio of measuring the total observed transverse momentum to the likelihood of the null hypothesis.

$$S(\overrightarrow{E}) \equiv 2ln\left(\frac{\mathcal{L}(\overrightarrow{E} = \sum_i \overrightarrow{E_{T_i}})}{\mathcal{L}(\overrightarrow{E} = 0)}\right)$$

- Assuming gaussian uncertainties distributions:

$$S \sim \left(\sum_i \overrightarrow{E_{T_i}}\right)^{\dagger} \left(\sum_i V_i\right)^{-1} \left(\sum_i \overrightarrow{E_{T_i}}\right)$$

- For each object contributing to the MET, the covariance matrix is calculated as:

$$V^{x,y} = \begin{bmatrix} \sigma_x^2 & \sigma_x\sigma_y \\ \sigma_x\sigma_y & \sigma_y^2 \end{bmatrix}$$

Where the measurements in the x and y components are 100% correlated.

- This matrix is rotated into a coordinate system with the x axis parallel to the total Met axis. Then, the total covariance matrix is calculated as the sum of all the covariance matrices from each object contributing to the Met:

$$\mathbf{V} = \sum_i R(\Phi(\mathrm{Met}))\, V_i^{x,y}\, R(\Phi(\mathrm{Met}))^{-1}$$

DEFINITION OF AN OBJECT BASED MET SIGNIFICANCE

Covariance Matrix

$$\mathbf{V} = \begin{bmatrix} \sigma_{\parallel}^2 & \sigma_{\parallel\perp}^2 \\ \sigma_{\parallel\perp}^2 & \sigma_{\perp}^2 \end{bmatrix}$$

Where $\sigma_{\parallel}^2$ is the variance in the direction of the Met, $\sigma_{\perp}^2$ is the variance perpendicular to the Met and $\sigma_{\parallel\perp}^2 = \rho\,\sigma_{\parallel}\sigma_{\perp}$ is the associated covariance.

Met Significance

- In this coordinate system, parallel and perpendicular to the total measured Met, the Met Significance can be simplified:

$$S \sim \left(\sum_i \overrightarrow{E_{T_i}}\right)^{\dagger} \mathbf{V}^{-1} \left(\sum_{\mathbf{i}} \overrightarrow{E_{T_i}}\right)$$

$$S \sim \frac{\left(\sum_i \overrightarrow{E_{T_i}}\right)^2}{\sigma_{\parallel}^2\,(1-\rho^2)}$$

- Where the correlation coefficient is:

$$\rho = \frac{\sigma_{\parallel\perp}^2}{\sqrt{\sigma_{\parallel}^2\,\sigma_{\perp}^2}}$$

NOTE: If

$$\rho^2 = 1$$

In this case the definition becomes:

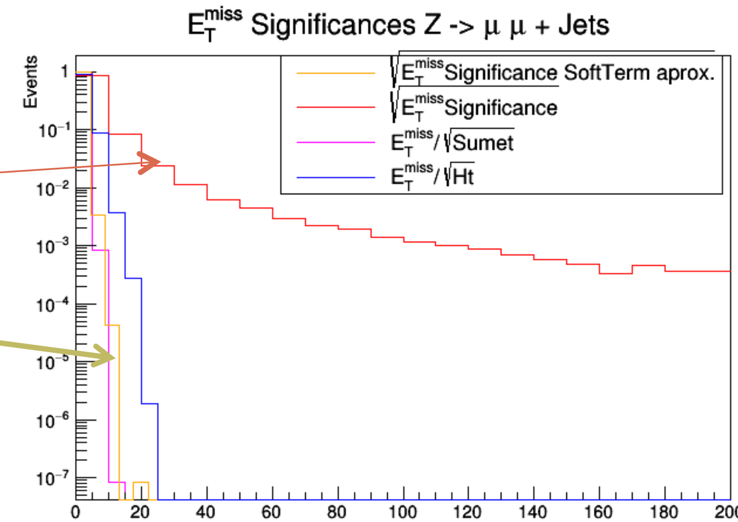
$$S \sim \frac{\left(\sum_i \overrightarrow{E_{T_i}}\right)^2}{\sigma_{\parallel}^2}$$

For $\rho^2 \geq 0.9$

Object based Met Significance

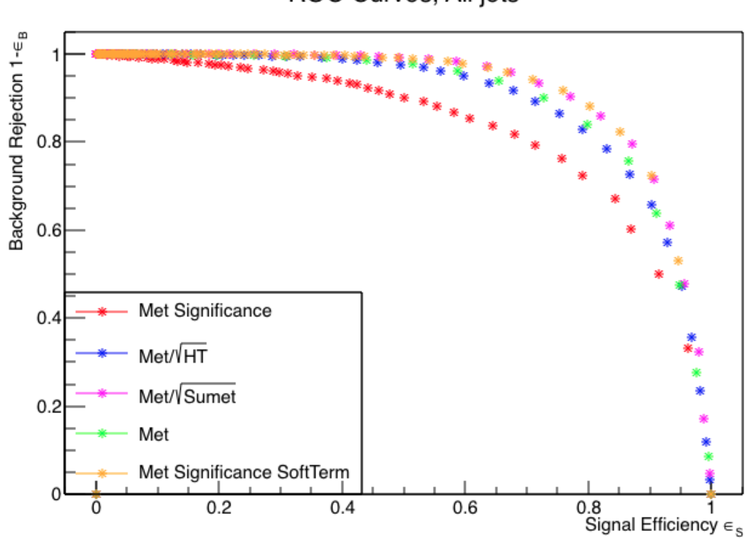
- **Considering jet resolutions**
- **Z-> μ μ + jets (no genuine Met)**
  - The Met error is under estimated (due to the missing if the other objects resolutions)so we obtain a over estimated distribution.
  - First approximation for the resolution of the Soft terms

$$S \sim \frac{\left(\sum_i \overrightarrow{E_{T_i}}\right)^2}{\sigma_{\|}^2\,(1-\rho^2)}$$



First test of Met significances Separation power

- Z(μ μ)+jet (BKG) and ZZ(μ μ ν ν)+jet (SGN)



Met Significance with a constant resolution of the soft term has a comparable performance w.r.t Met/sqrt(Sumet)

This is our performance starting point. We are missing…

- Properly implement the Soft Term resolutions
- Other objects resolutions

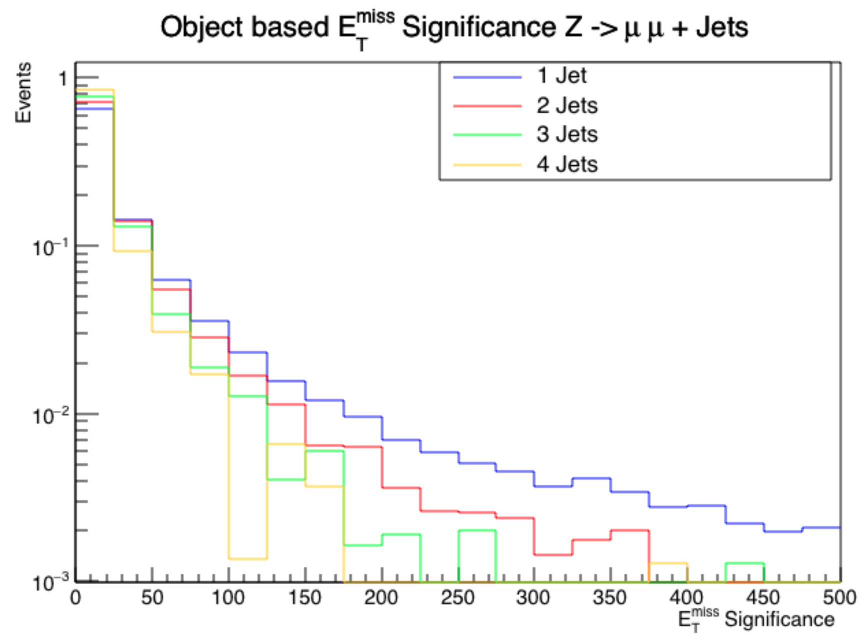
FIRST MET SIGNIFICANCE IMPLEMENTATION

Object based Met Significance for **Z-> μ μ + jets**

Considering only the jet resolution
(i.e. without soft term and lepton resolutions).

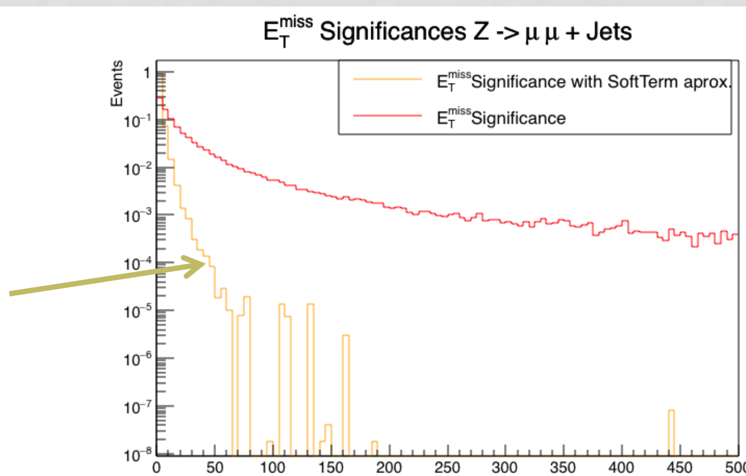
The Met error in the denominator of the Significance is under estimated (due to the missing if the other objects resolutions)so we obtain a over estimated significance distribution.

$$S \sim \frac{\left(\sum_i \overrightarrow{E_{T_i}}\right)^2}{\sigma_{\parallel}^2\,(1-\rho^2)}$$



We can have a first approximation (very rought from 0 Jet) for the resolution of the Soft term, assuming:

- Constant resolution in the x and y direction of 13.5 GeV (see backup slides)

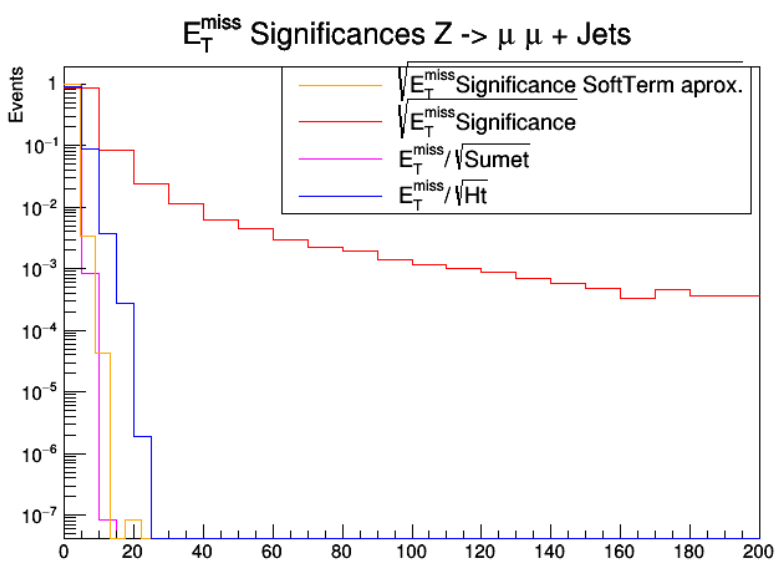


MET SIGNIFICANCES SEPARATION POWER

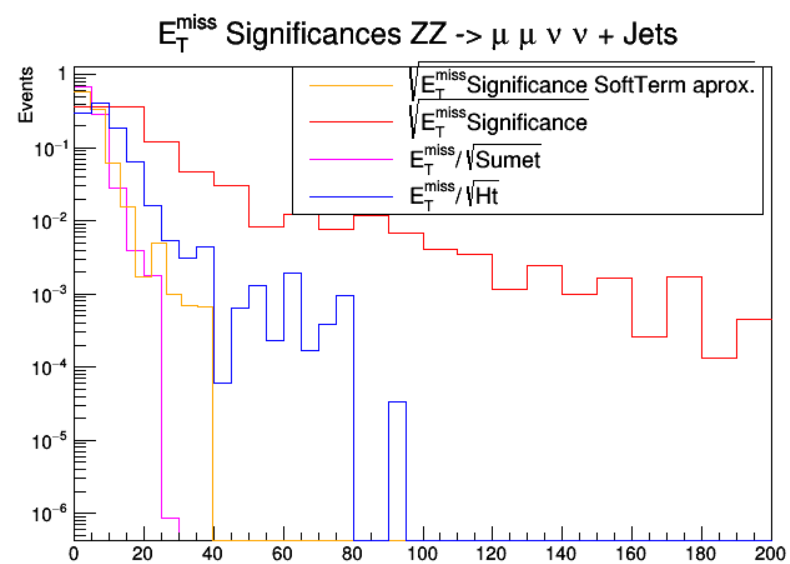
- Comparison of the separation power between
  - Z($\mu\ \mu$)+jet (BKG) and ZZ($\mu\ \mu\ \nu\ \nu$)+jet (SGN)
  - Cutting on MET - cutting on MET significance - cutting on MET/sqrt(HT) or cutting on MET/sqrt(Sumet)

Sample	MC Generator	MC channel number	Run number	xSection	KFactor	Filter Efficiency
Zmumu	Powheg	361107	222525	1951	1.026	1
Dibosonllvv	Sherpa	361068	222525	12.76	0.91	1

Signal and Background Met Significance Distributions



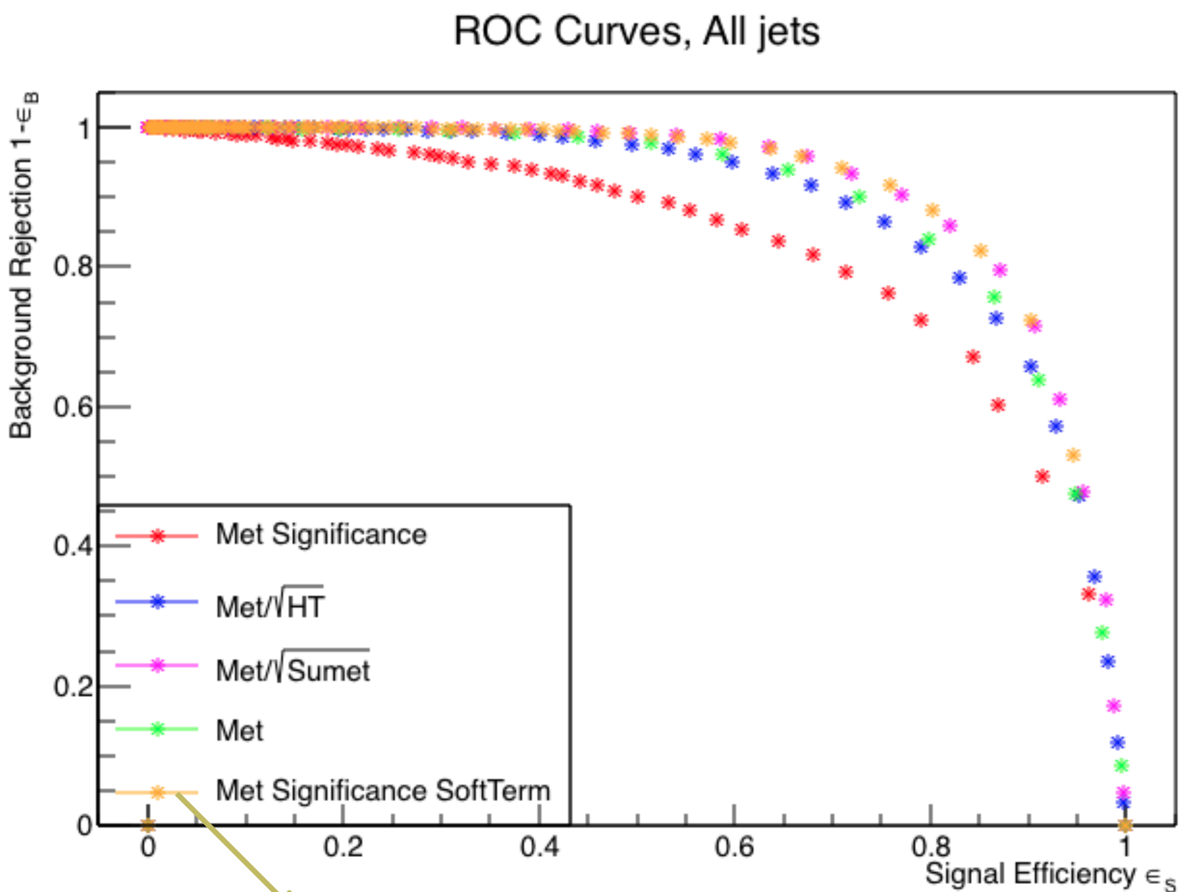
Background



Signal

MET SIGNIFICANCES SEPARATION POWER

ROC Curve



Important: MET significance calculated considering just jet resolutions

Met Significance with a constant resolution of the soft term (as an approximation) has a comparable performance w.r.t Met/sqrt(Sumet)

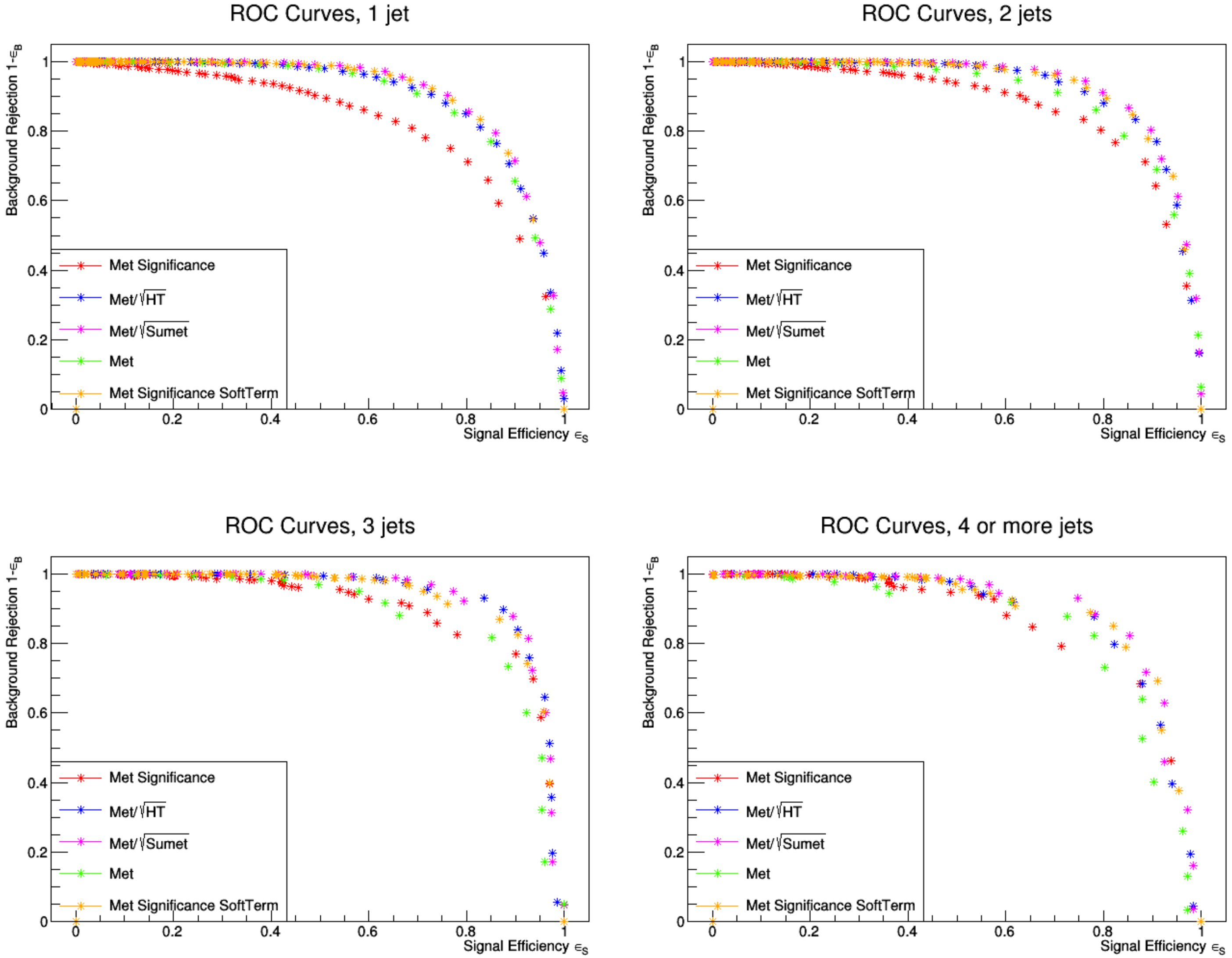
We need to properly implement the Soft Term resolutions!

…And we are also missing the other objects resolutions

Starting point. We can do a better implementation of the soft Term

MET SIGNIFICANCES SEPARATION POWER

ROC Curve for different Jet Multiplicities



At higher jet multiplicities

- Contributions from jet resolution are bigger than the one from soft term and muons.
- Performance of Met Sig. with jet resolutions approaches the other ROC curves

MET SIGNIFICANCES SEPARATION POWER

Plan

- We need to parameterize the Soft Term resolutions which is a important contribution in the study of the Zmumu Met significance.
- Study the performance of the object based Met Significance considering the soft term resolutions.
- After this first part on the study of the algorithm and the study of its performance, we will start to prepare tools to make this available in ATLAS.

FIRST STEPS TOWARD OBJECT BASED MET SIGNIFICANCE

$$S \sim \left(\sum_i \overrightarrow{E_{T_i}}\right)\left(\sum_i V_i\right)^{-1}\left(\sum_i \overrightarrow{E_{T_i}}\right)$$

The covariance matrices V are estimated by propagating the expected resolutions for all the objects

- First test: Build the V in an easy environment, start selecting Z->mumu+1 jet. Why?
- Assuming contribution from resolution from jets bigger than the one from muon, we can neglect muons resolution for the moment.
- Selecting events with only 1 jet should be easier to calculate V
- The only terms participating the MET resolutions are jets and soft terms.
- Select events with low soft term activity to be able to first focus on jet contribution and to validate the method.
- Selecting events in which MET and Jet are aligned. Statistically more likely the unbalance comes from jet resolution than Soft Term activity.

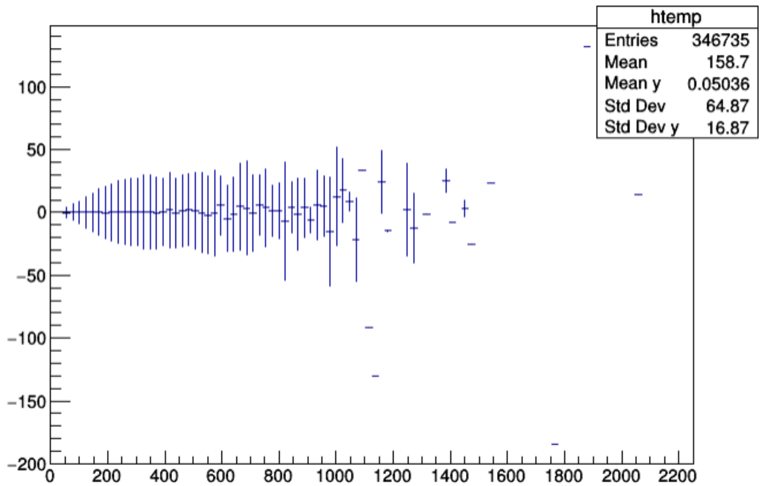
- Once the first MET significance is calculated, planning to compare the separation power between Z($\mu\mu$)+jet (BKG) and ZZ($\mu\mu\nu\nu$)+jet (SGN)
  - cutting on MET - cutting on MET significance - cutting on MET/sqrt(HT) or cutting on MET/sqrt(Sumet)
- Depending on these first tests, we will start adding complexity in the final state, adding jets, adding other backgrounds (ttbar), and other objects
  - Most of the CP groups already contacted – Paolo
- After this first part on the study of the algorithm and the study of its performance, we will start to prepare tools to make this available in ATLAS.

FIRST STEPS TOWARD OBJECT BASED MET SIGNIFICANCE

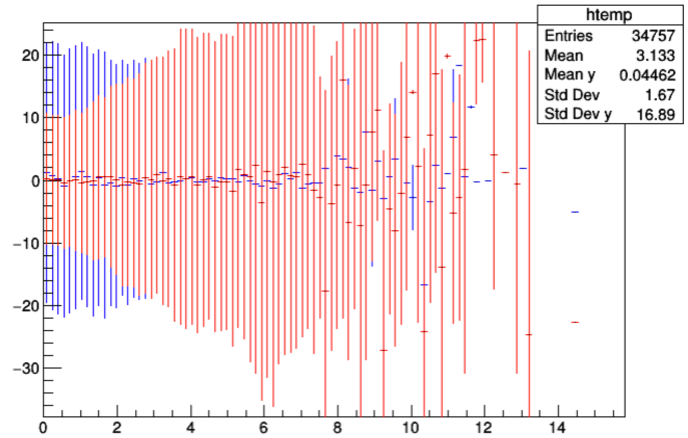
First case study : 1Jet recoiling against a Z boson in Zmumu

We are going to start with the Met Significance definition implementation for one Jet and suppressing the MET soft terms. And comparing with the current definition $S = \frac{\text{Met}}{\sqrt{\text{Sumet}}}$

Met x component resolution with respect to sumet

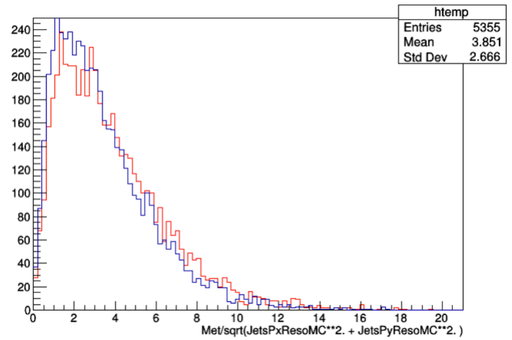


Met x component resolution with respect to Px resolution of the Jet



QUALIFICATION TASK PLANS

- First study with one Jet suppressing the Met soft term in order to understant the jet contribution to the significance and implement the significance considering the Met-error propagation with one Jet.
- Compare this jet Met Significance with respect to the current definition $S = \frac{\text{Met}}{\sqrt{\text{Sumet}}}$

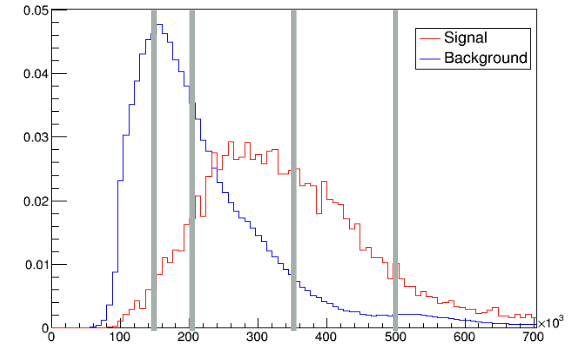


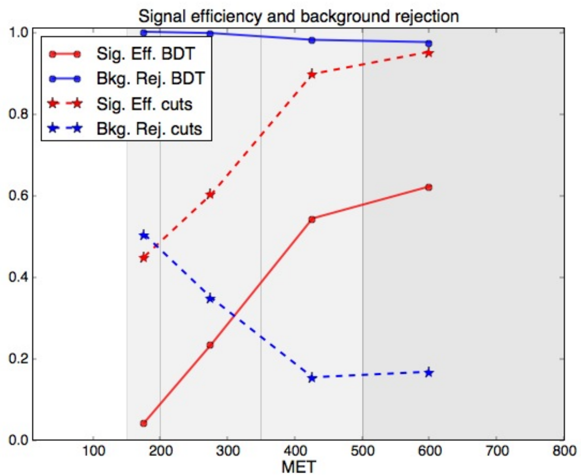
- Evaluate the potential improvement with this new definition with a ROC curve.
- Implement the covariance matriz and the likelihood ratio for more objects.

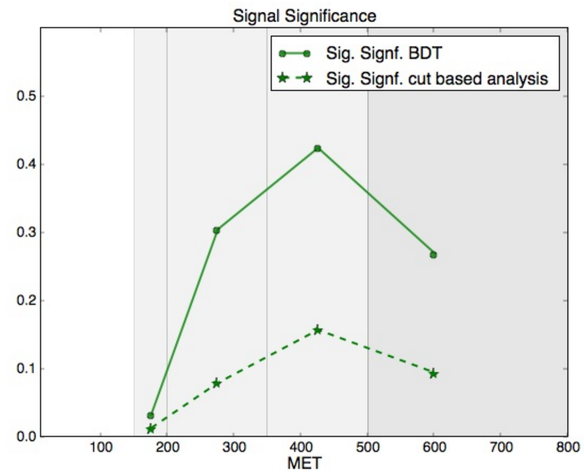
NAÏVE COMPARISON BETWEEN BDT CLASSIFIER AND CUTS BASED ANALYSIS

As a first aproximation we can compare the performance of the BDT classifier used as a single cut variable w.r.t. the cut based analysis (without b-tag caterories in this page).

- **MET categories**







MET CATEGORY [GeV]	ALL MET RANGE		150 < MET < 200		200 < MET < 350		350 < MET < 500		500 < MET	
Signal Events *	1.98418		0.138348		0.910875		0.610098		0.27755	
Background Events *	210.02		56.5411		74.8985		13.9164		9.37206	
	Cut Based	BDT > 0.1318	Cut Based	BDT > 0.09	Cut Based	BDT > 0.1318	Cut Based	BDT > 0.14	Cut Based	BDT > 0.1
Signal Events after cutting *	1.42071	0.696183	0.0620004	0.0056364	0.548098	0.213159	0.546731	0.330669	0.263886	0.172337
Background Events after cutting *	96.4075	0.746975	28.0949	0.0268668	48.9232	0.280533	11.7927	0.281255	7.81377	0.241865
Signal efficiency	0.716021	0.350868	0.448148	0.0407407	0.601727	0.234015	0.896137	0.541994	0.950769	0.620922
Background rejection	0.540959	0.996443	0.503107	0.999525	0.346807	0.996254	0.152606	0.97979	0.16627	0.974193
Signal significance	**0.143639**	**0.579517**	**0.0116843**	**0.0312636**	**0.0779258**	**0.303372**	**0.155642**	**0.422712**	**0.0928483**	**0.267777**

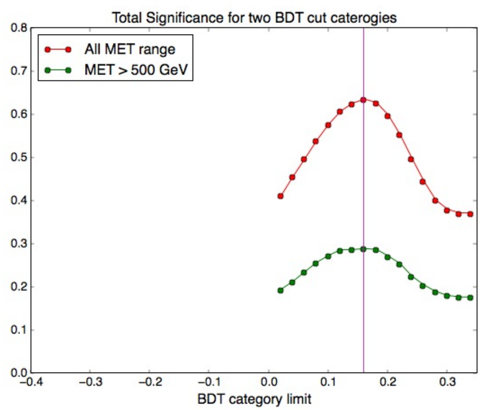
* From now on, the number of signal and background events are associated to the TestTree which correspond to half of the statistics

DEFINING BDT CATEGORIES

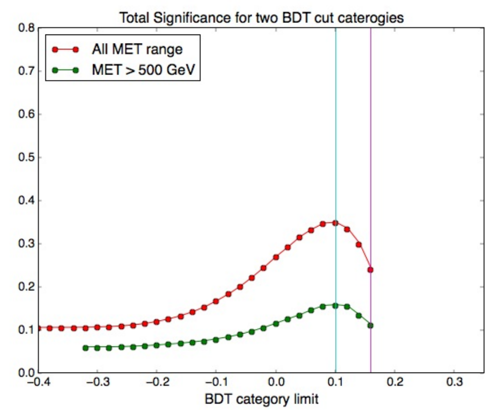
Low – medium – high BDT categories are defined by maximizing the combined significance at first order.

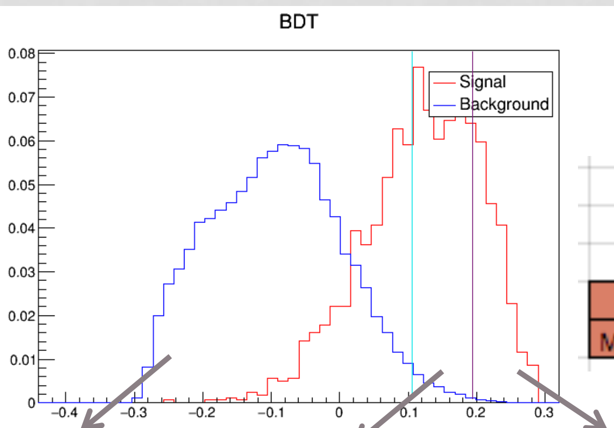
$$S = \sqrt{S^2_{LOW} + S^2_{MEDIUM} + S^2_{HIGH}}$$

Medium and high categories: The combined significance is calculated variating the the BDT limit between the categories starting from BDT=0 to the right.



Low and medium categories: Starting from the limit that maximises the medium and high categories, a second maximum is found for the combined significance to the left.





Low BDT category [-0.45, 0.1]

Medium BDT category[0.1, 0.18]

High BDT category [0.18, 0.4]

	Purity (Sig./Bkgr.)					
	Cuts Based Analysis			BDT		
	0 b tag	1btag	2 btag	Low	Medium	High
ALL MET	0.006138141	0.02003048	0.2326927	0.004973259	0.3321946	2.411031
MET > 500 GeV	0.01499348813	0.130980753	1.04812285	0.01149569568	0.4631069043	2.712491267

SCALAR MEDIATOR

Simple **potential**

DM couple to the SM only though the Higgs

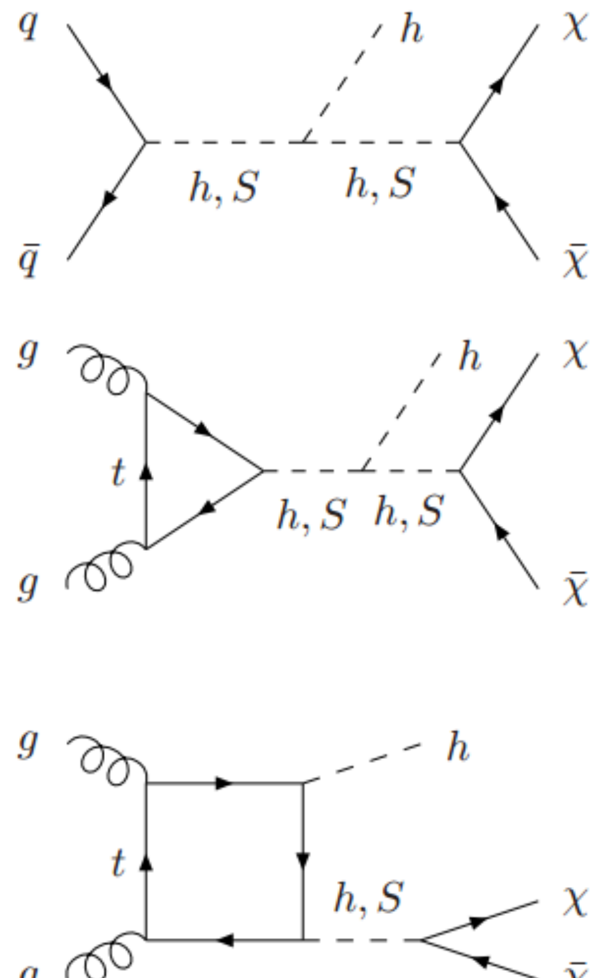
$$V \supset a|H|^2 S + b|H|^2 S^2 + \lambda_h |H|^4$$

Coupling of the scalar with DM ⟶ $-y_\chi \bar{\chi}\chi S$

After SSB ⟶ Mixing between h and S ⟶ θ

Quark and DM couplings terms:

$$-y_\chi \bar{\chi}\chi(c_\theta S - s_\theta h) - \frac{m_q}{v}\bar{q}q(c_\theta h + s_\theta S)$$



SCALAR MEDIATOR

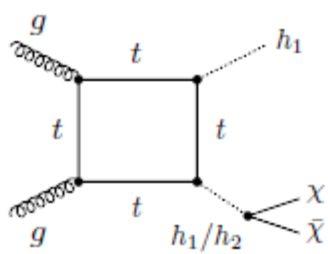
$$\mathcal{L}_{\text{fermion},H} \supset -\mu_s s^3 - \lambda_s s^4 - y_\chi \bar{\chi}\chi s - \mu_p s|H|^2 - \lambda_p s^2|H|^2$$

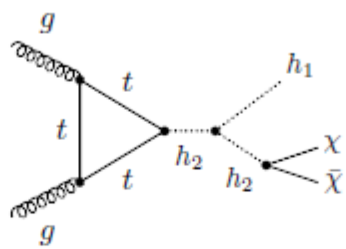
$$\langle s \rangle = 0$$

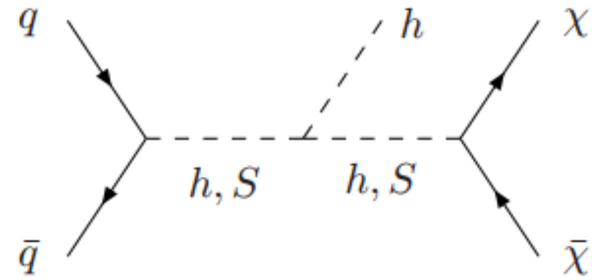
$$\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} h \\ s \end{pmatrix} \qquad \tan(2\theta) = \frac{2v\mu_p}{m_s^2 + \lambda_p v^2 - m_h^2}$$

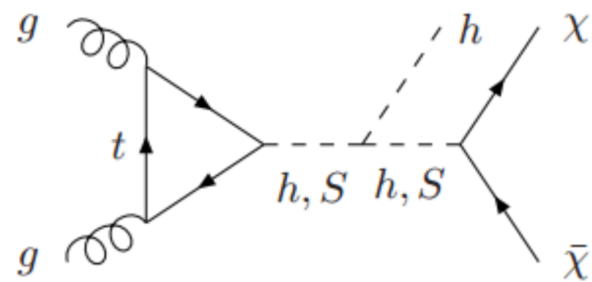
$\theta \to 0$ the dark sector is decoupled from the SM $\qquad \sin\theta \lesssim 0.4$

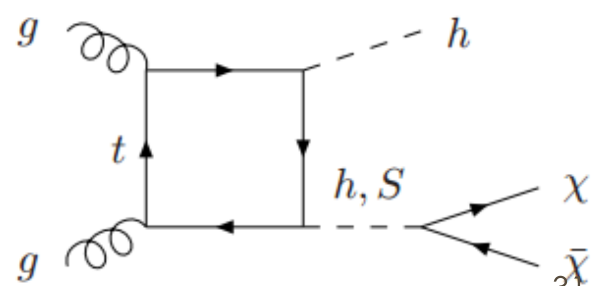
$m_{h_1} \simeq m_h$ and $m_{h_2} \simeq (m_s^2 + \lambda_p v^2)^{1/2}$











BOOSTED DESITION TREE

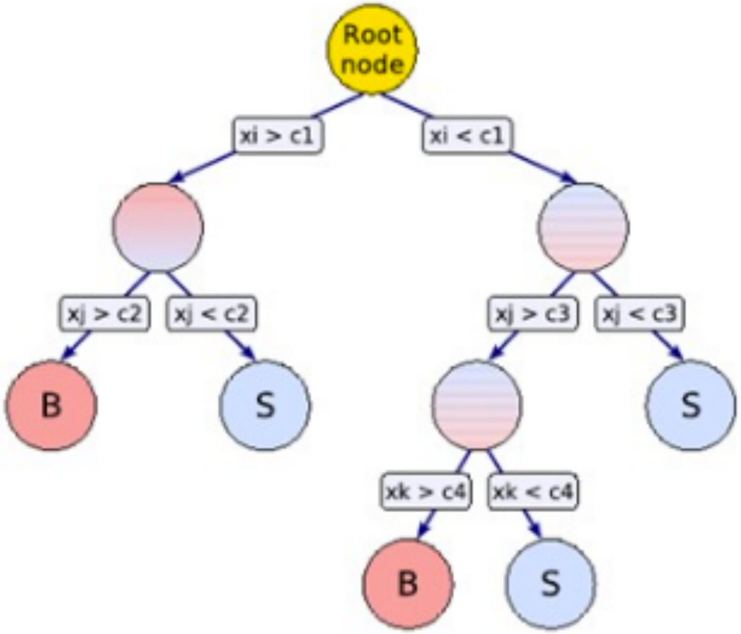
- Decision Tree

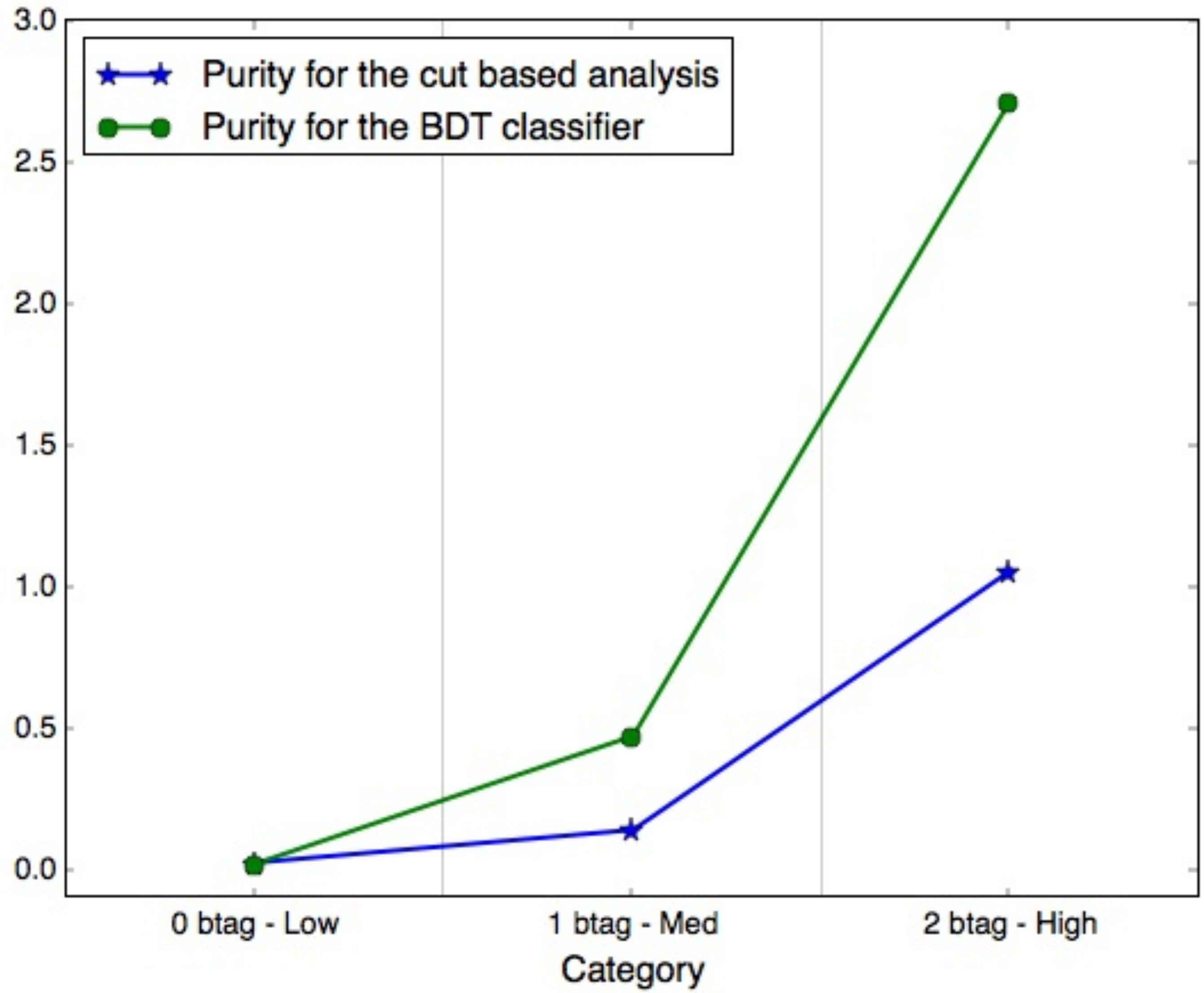
- Consecutive set of questions (nodes)
  - Only two possible answers per question
  - Each question depends on the formerly given answers
- Final veredict (leaf) is reached after a given maximum number of nodes

- Random forest

- Random Forests is an ensemble method that combines different trees
- Final output is determined by the majority vote of all trees
- Boosting
  - Misclassified events are weighted higher so that future learners concentrate on these

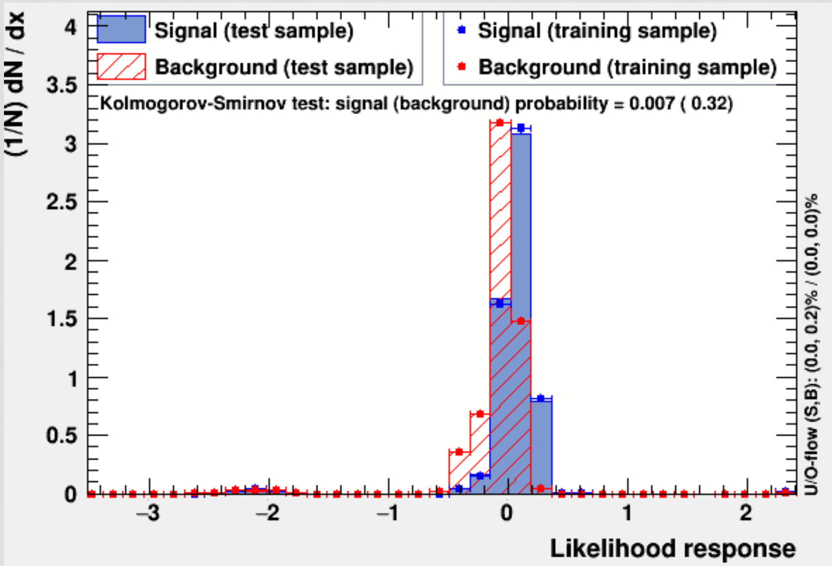
- Training a decision tree

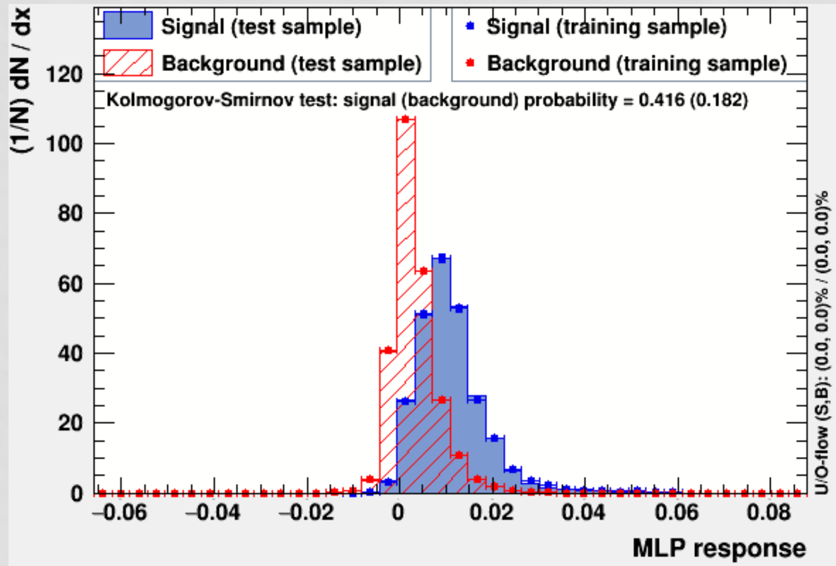




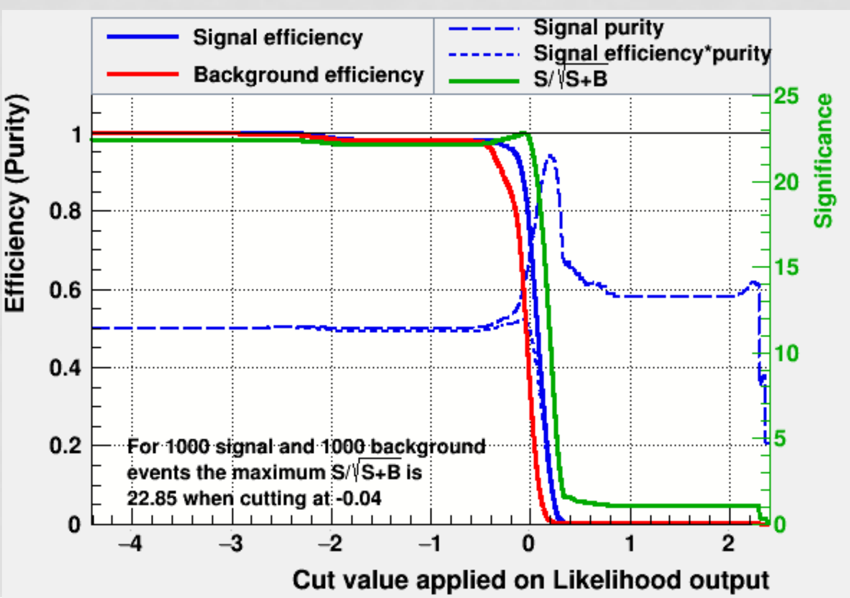
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Purity for the cut based analysis
Purity for the BDT classifier
0 btag - Low
1 btag - Med
2 btag - High
Category

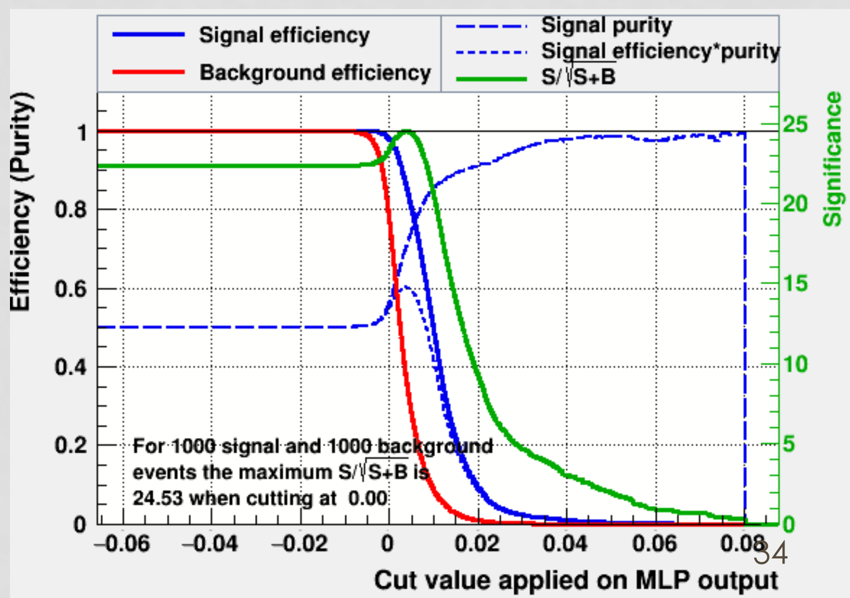
Classifier Output Distributions





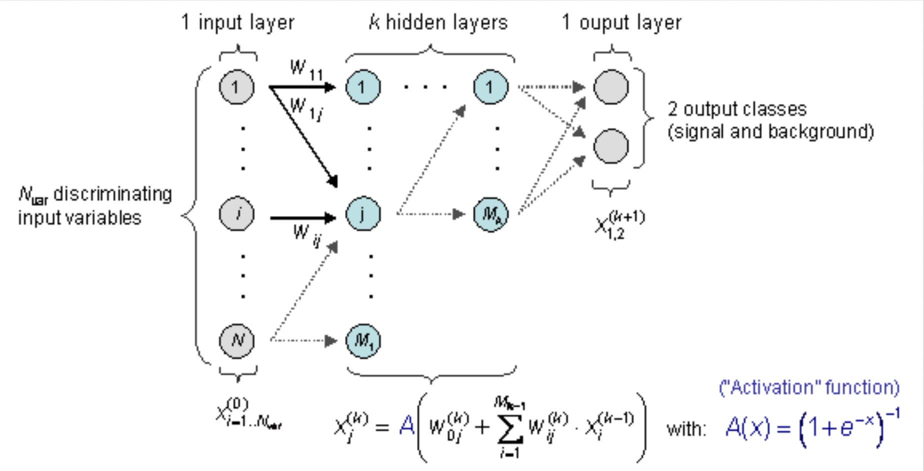
Classifier Cut Efficiencies





ARTIFICIAL NEURAL NETWORKS (NON-LINEAR DISCRIMINANT ANALYSIS)

- All ANNs belong to the class of Multilayer Perceptrons (MLP), which are feed-forward networks according to the following propagation schema:



- The input layer contains as many neurons as input variables used in the MVA. The output layer contains a single neuron for the signal weight. In between the input and output layers are a variable number of k hidden layers with arbitrary numbers of neurons. (While the structure of the input and output layers is determined by the problem, the hidden layers can be configured by the user through the option string of the method booking.)
- As indicated in the sketch, all neuron inputs to a layer are linear combinations of the neuron output of the previous layer. The transfer from input to output within a neuron is performed by means of an "activation function". In general, the activation function of a neuron can be zero (deactivated), one (linear), or non-linear. The above example uses a sigmoid activation function. The transfer function of the output layer is usually linear.

LIKELIHOOD (PDE APPROACH)

We define the likelihood ratio, R, for an event by the ratio of the signal to the signal plus background likelihoods. The individual likelihoods are products of the corresponding probability densities of the discriminating input variables used. In practice, TMVA uses polynomial splines fitted to histograms, or unbinned Gaussian kernel density estimators, to estimate the probability density functions (PDF) obtained from the distributions of the training variables.

MVA TRAINING VARIABLES AND SAMPLES

- Variables in training

Variable
EtaFatJet
mindphi
mBB
dPhiMETMPT
nJ
MET
btag
dPhiMETFatJet
MVH
nTrackJetsTagsOut
pTFatJet
MPT
nTrackJetsTagsIn
MEff
ntracktag
MEff3
PTVH
nFatJets

- Background Samples

Name	physically_descriptive_name	Xsec[pb]	#Events in Ntuples	#Events (1 fb-1)
ttbar	PowhegPythia_P2012_ttbar_allhad	696.04	316338	318.2851712
ZnunuB	Sherpa_CT10_ZnunuMassiveCBPt140_280_BFilter	60.02	1970516	7.1765914
ZnunuC	Sherpa_CT10_ZnunuMassiveCBPt140_280_CFilterBVeto	60.239	2393875	24.38173525
ZnunuL	Sherpa_CT10_ZnunuMassiveCBPt140_280_CVetoBVeto	60.23	2468799	28.3870013
ZZ	Sherpa_CT10_ZqqZvv	16.492	3627	4.63359232
ZvvH125	Pythia8_AU2CTEQ6L1_ZH125_nunubb	0.8696	43722	0.10035184
WZ	Sherpa_CT10_WlvZqq	12.543	41297	12.543
WW	Sherpa_CT10_WplvWmqq	25.995	27248	25.995

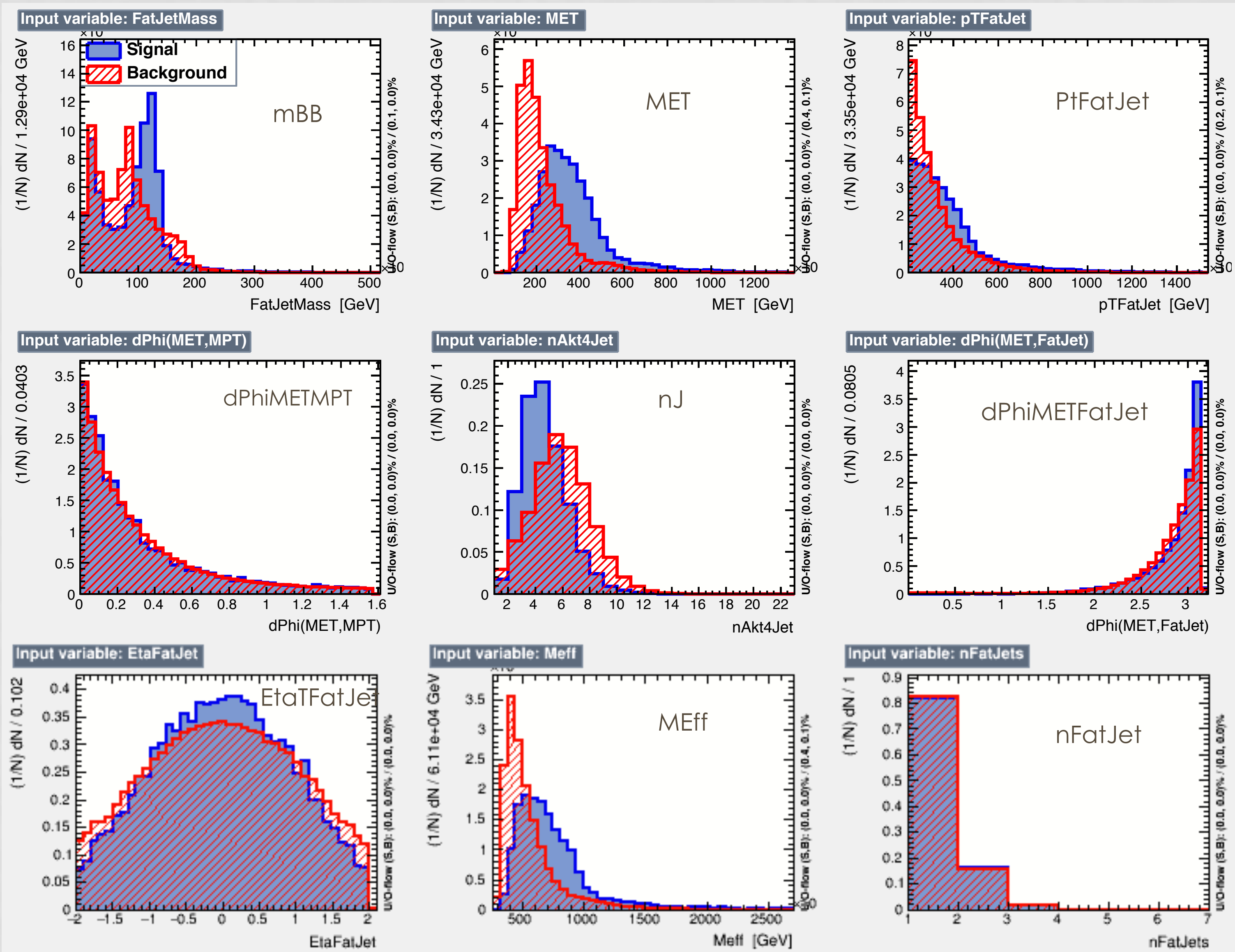
- Signal Sample

Name	physically_descriptive_name	Xsec[pb]	#Events in Ntuples	#Events (1 fb-1)
shxxbbms1000mx1	MadGraphPythia8EvtGen_A14NNPDF23LO_shxx_bb_ms1000_mx1	3.9708	23245	3.9708

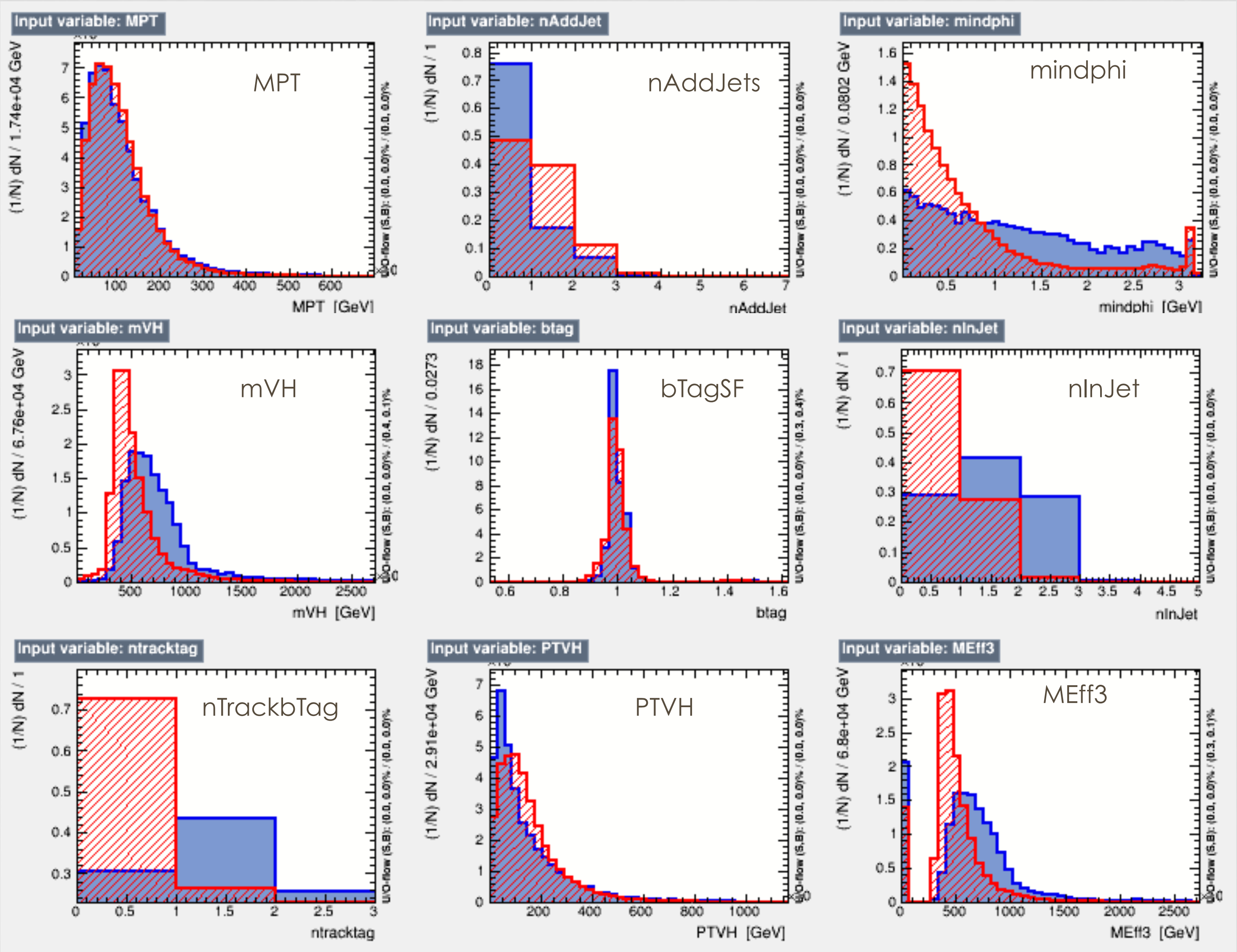
- Samples taken from CxAOD_00-14-05
- Normalizaded to 1 fb-1 of integrated luminosity

- For the time being only merged regime is considered.
- Preselection: At least one fat jet

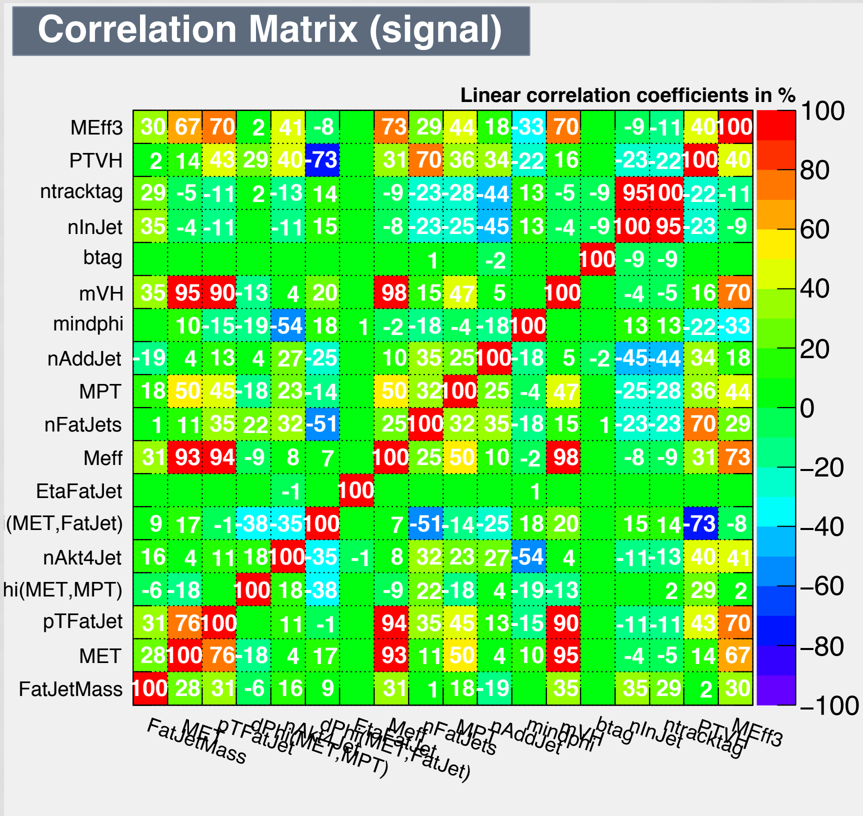
VARIABLES DISTRIBUTIONS

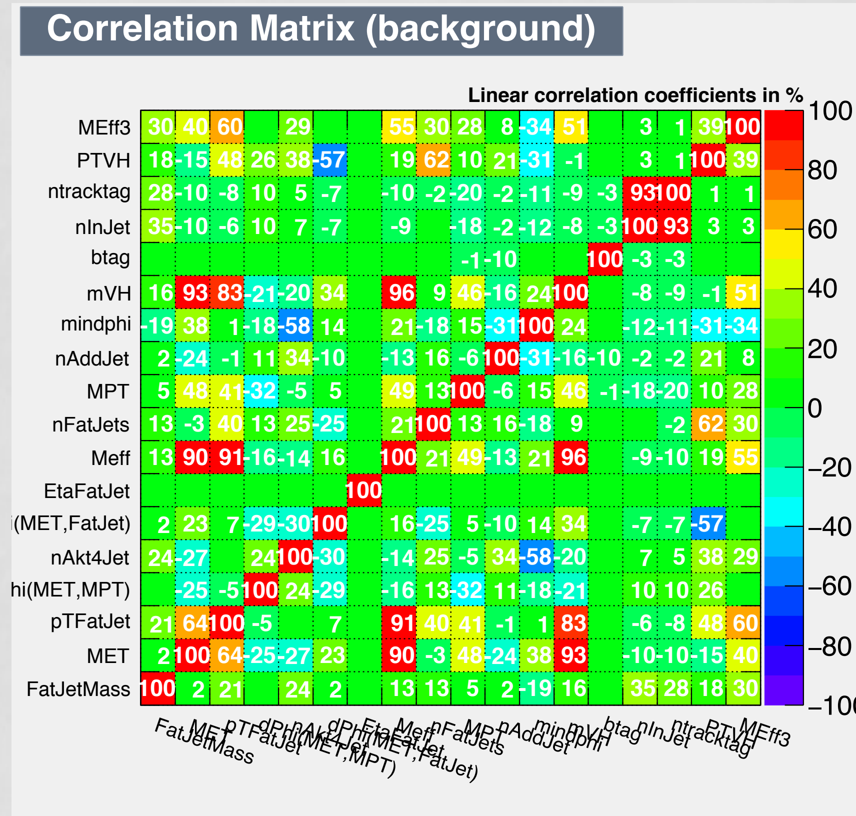


VARIABLES DISTRIBUTIONS



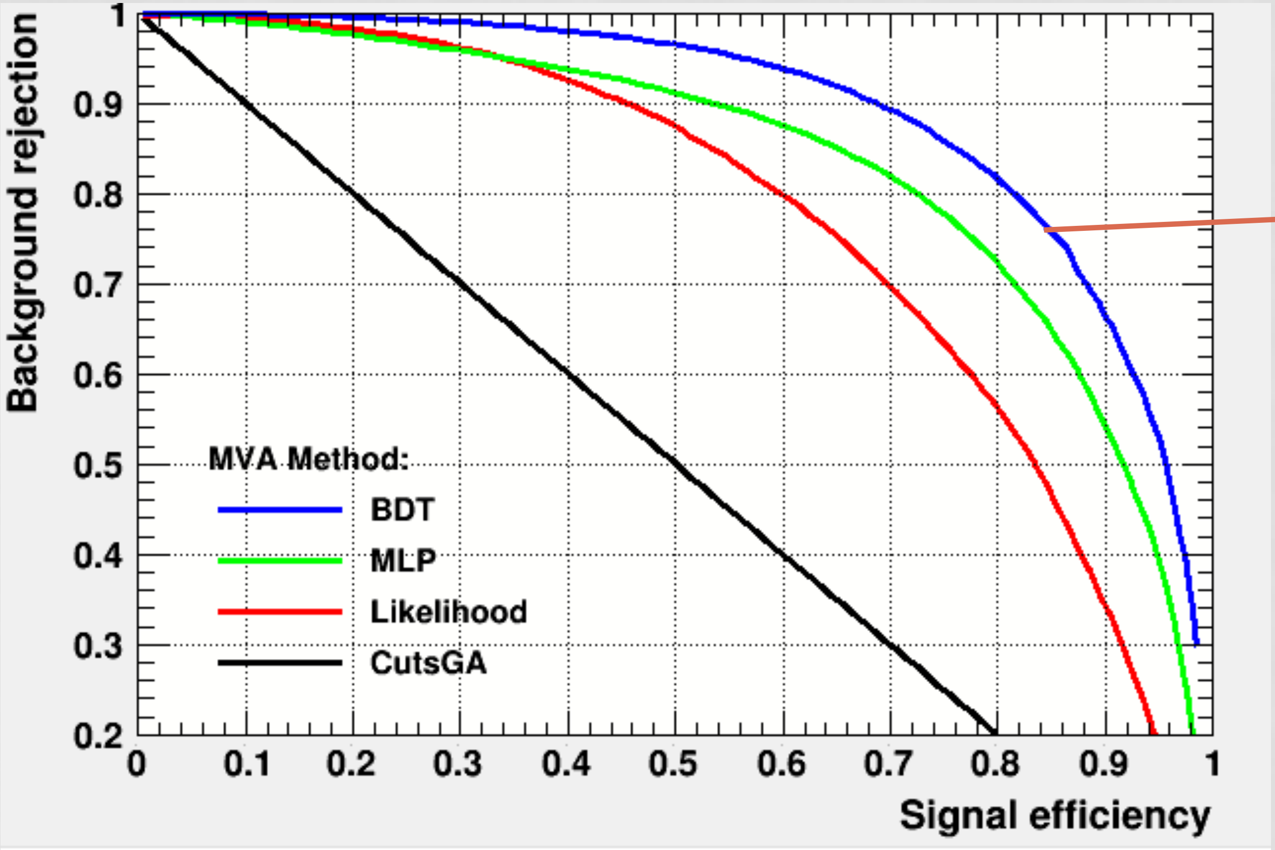
CORRELATION COEFFICIENTS





MVA ALGORITHMS COMPARISON

ROC (Receiver Operation Characteristics) curve shows the performance (characterize the quality of classification) for the MVA methods. The one that has that largest AUC (Area Under the Curve) has the best classification power between signal and background.



BDT classifier has the best performance

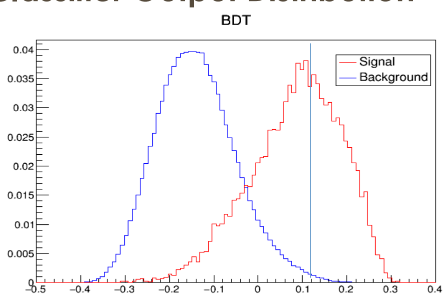
BDT

In the classifier distribution, the signal-like events are on the right (high BDT) and the background-like events are on the left (low BDT). A straightforward implementation of the BDT classifier is to perform a single cut on this variable in which the significance is maximized for a given number of signal and background generated events.

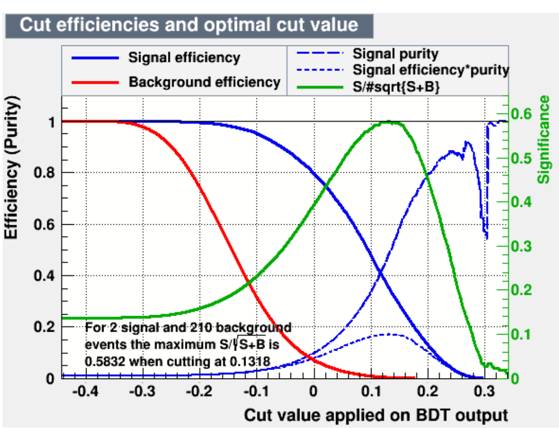
ALL MET RANGE

The BDT method was trained for all the MET range in order to take advantage of the whole statistics.

- **Classifier Output Distribution**



- **Classifier Cut Efficiencies**

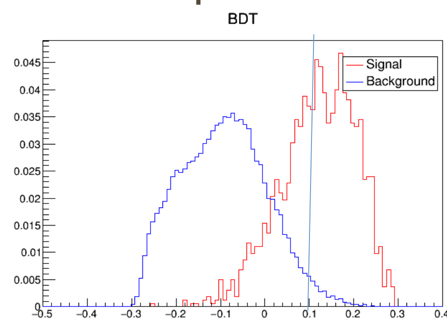


- Sign. Events: 1.98 (TestTree)
- Bckg. Events: 210.02 (TestTree)

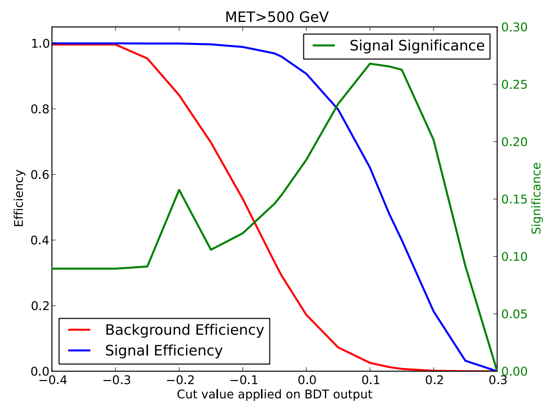
MET > 500 GeV

For now, we are interested in the high MET category .

- **Classifier Output Distribution**

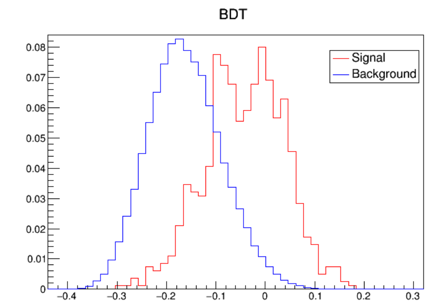


- **Classifier Cut efficiencies**

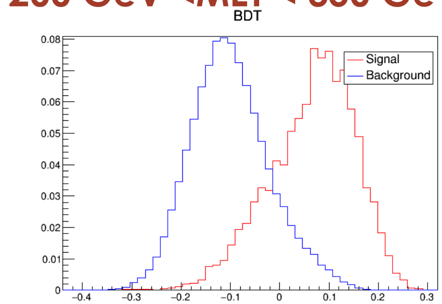


- Sign. Events: 0.28 (TestTree)
- Bckg. Events: 9.37 (TestTree)

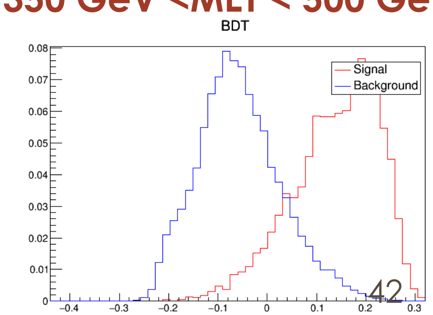
150 GeV <MET < 200 GeV



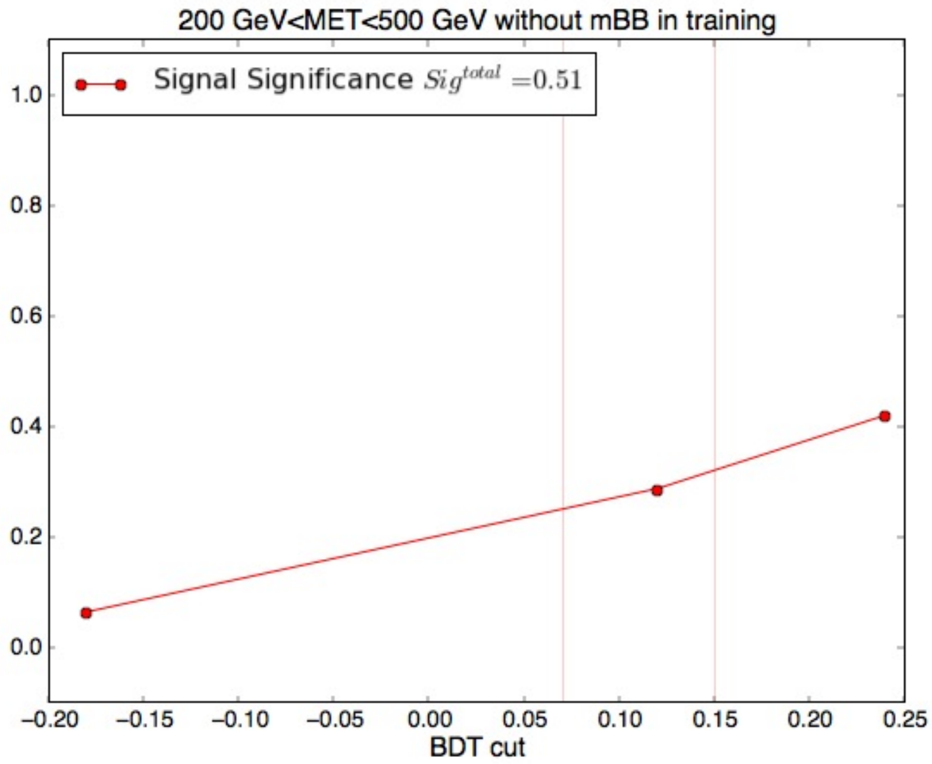
200 GeV <MET < 350 GeV

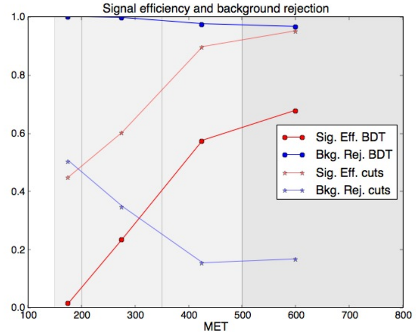


350 GeV <MET < 500 GeV



MERGED EVENTS WITH MET<500 GEV

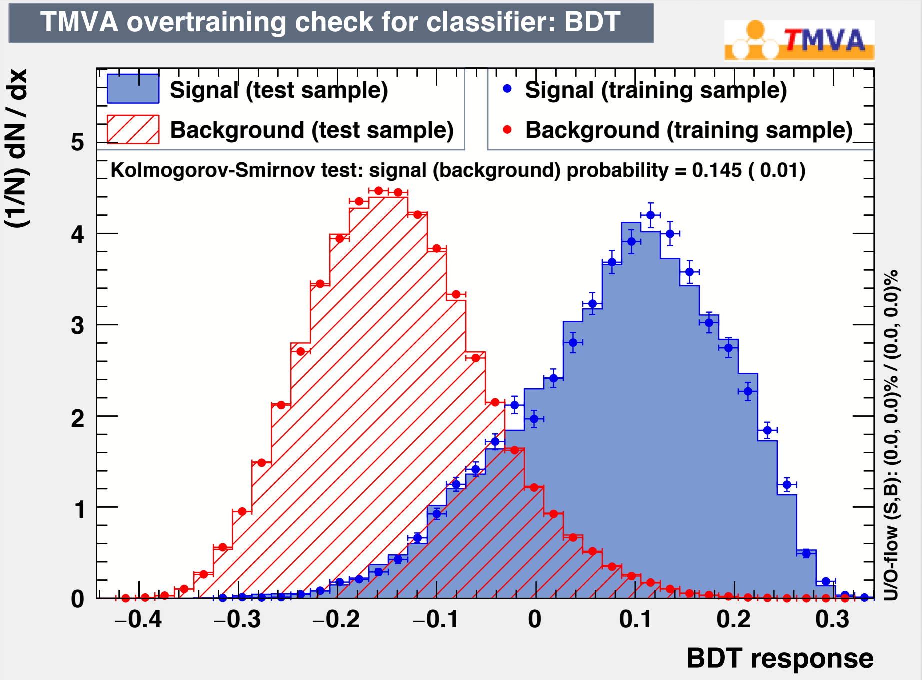




PRELIMINARY MVA ON SCALAR SM

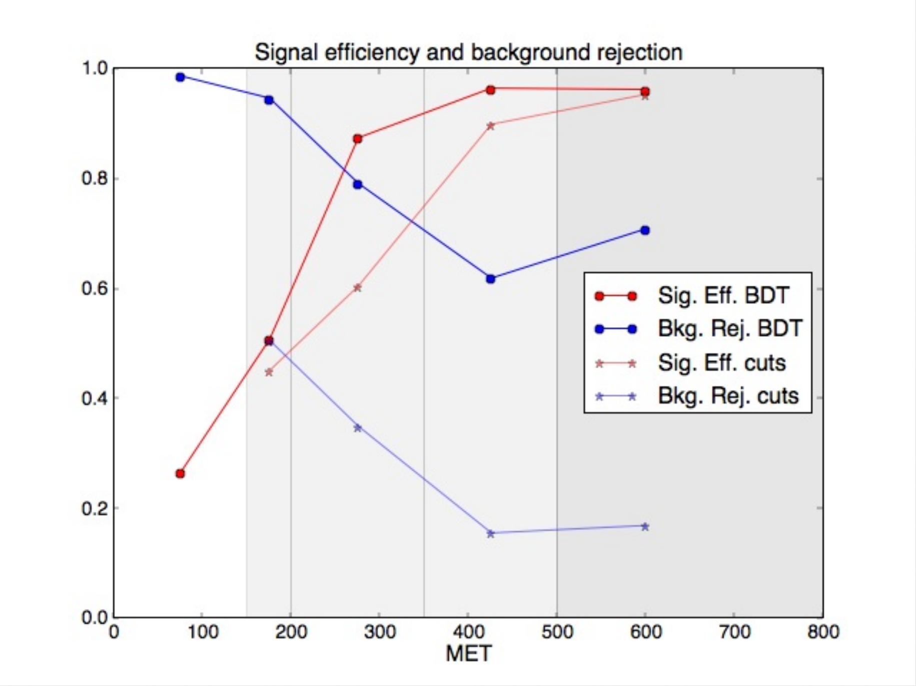
- Variables in training

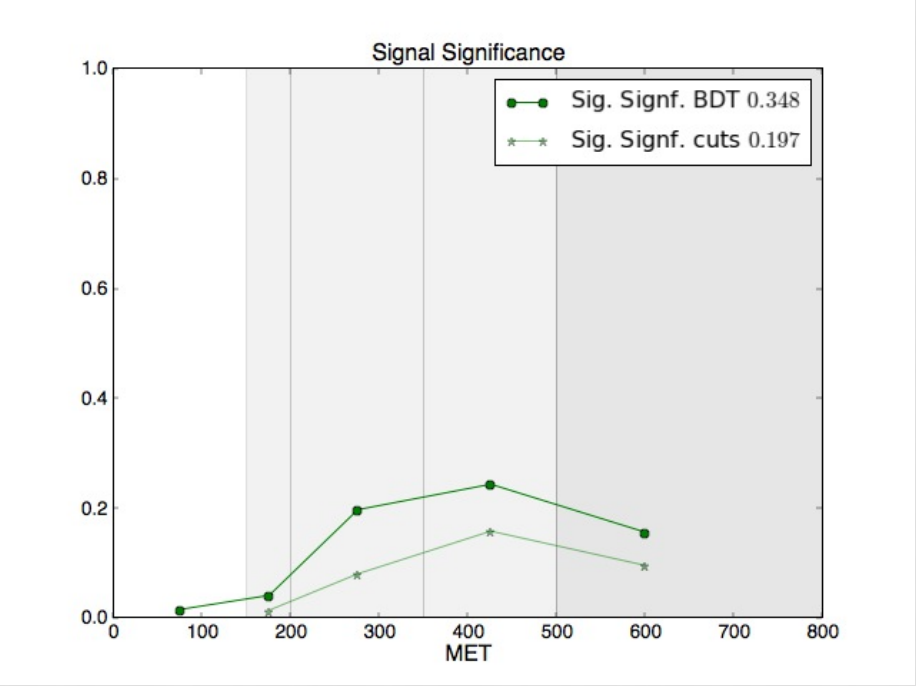
Rank	Variable	Variable Importance
1	EtaFatJet	9.03E-02
2	mindphi	8.67E-02
3	mBB	7.86E-02
4	dPhiMETMPT	7.41E-02
5	nJ	7.10E-02
6	MET	7.04E-02
7	btag	6.55E-02
8	dPhiMETFatJet	5.76E-02
9	MVH	5.12E-02
10	nTrackJetsTagsOut	4.94E-02
11	pTFatJet	4.58E-02
12	MPT	4.36E-02
13	nTrackJetsTagsIn	4.34E-02
14	MEff	4.13E-02
15	ntracktag	3.84E-02
16	MEff3	3.74E-02
17	PTVH	2.88E-02
18	nFatJets	2.65E-02

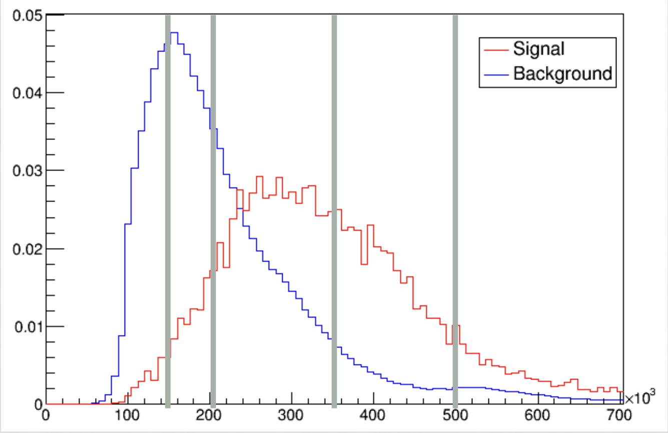


COMPARISON BETWEEN BDT CLASSIFIER AND CUTS BASED ANALYSIS

- MET categories





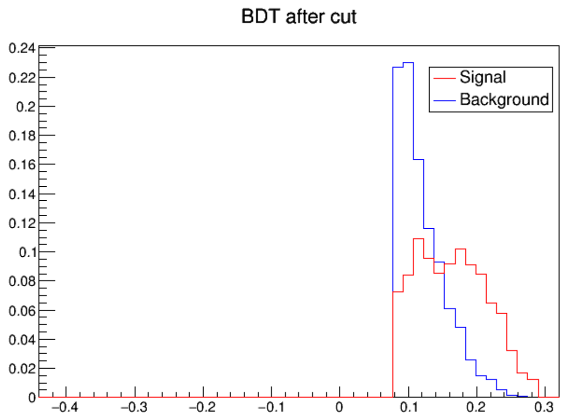


BDT cut = -0.0390 not optimized for the number of signal and background

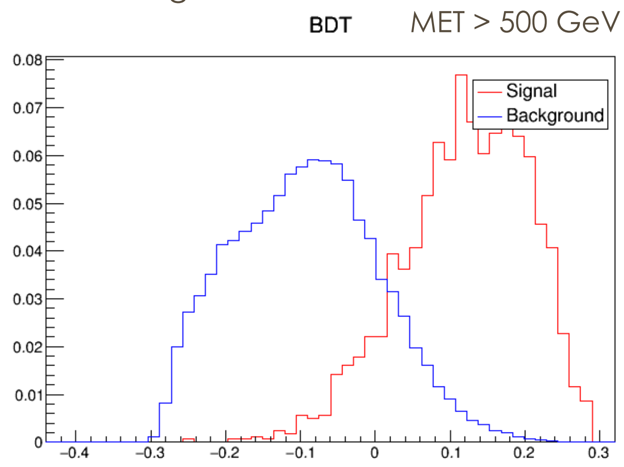
BDT CATEGORIES MET>500 GEV

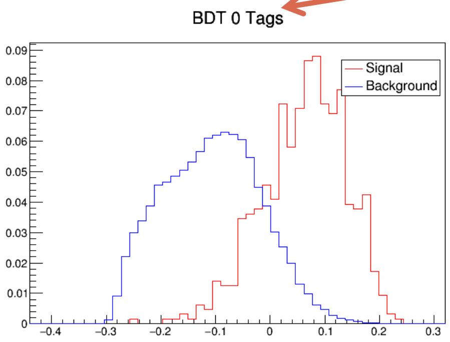
The number of b-tag jets is a variable considered in the BDT training.

Then, instead of having a single sharp cut on BDT

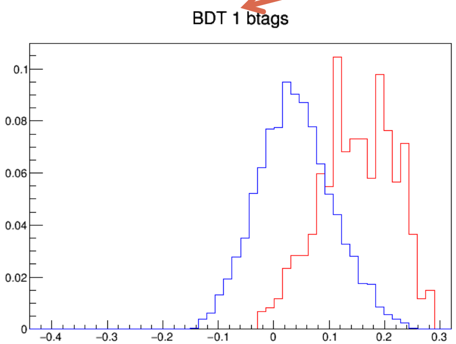


We can define BDT categories that can be related to 0, 1 and 2 btags.

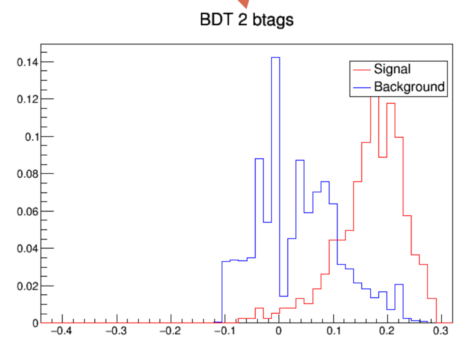




- Low BDT category



- Medium BDT category

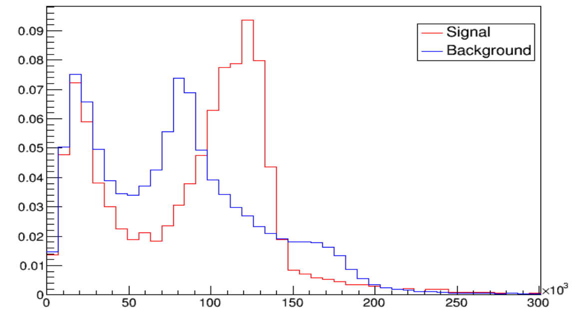


- High BDT category

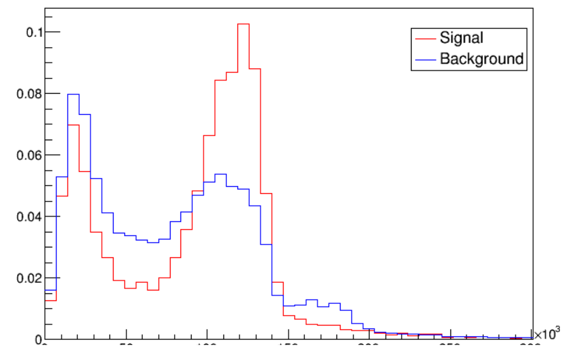
Distributions of the variables

ALL MET

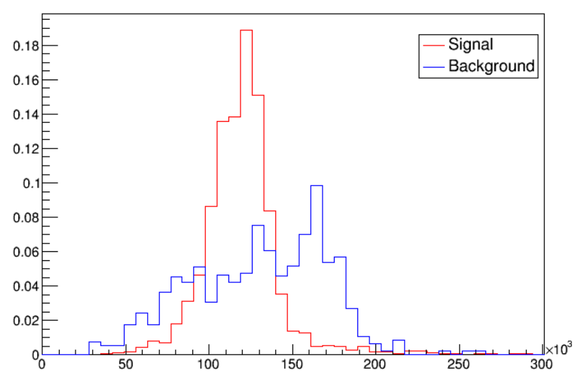
Signal mBB



mBB after BDT cut

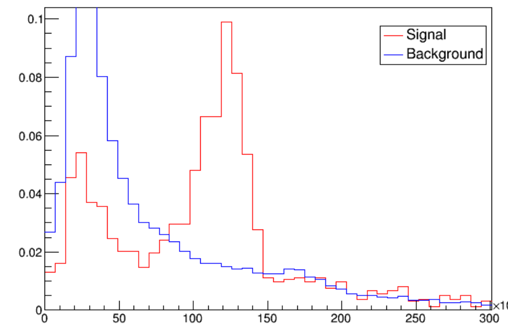


Signal mBB that pass mono-H selectin cuts

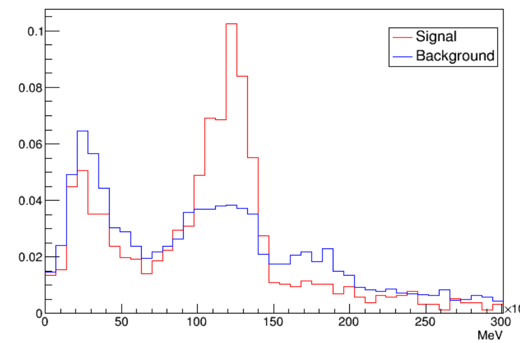


MET > 500 GeV

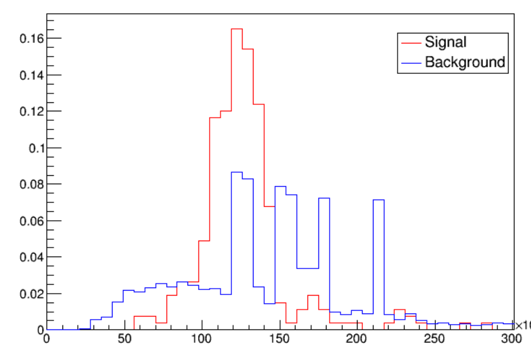
Signal mBB



mBB after BDT cut

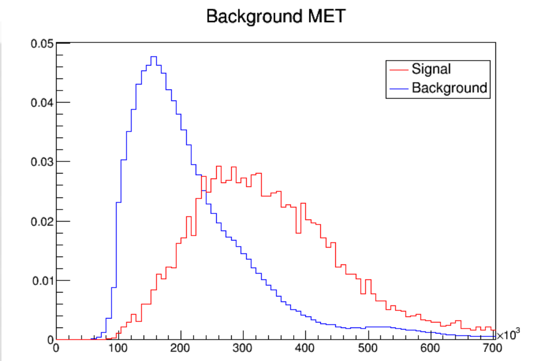


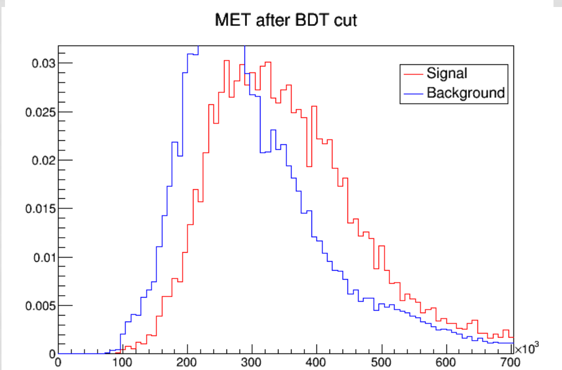
Signal mBB that pass mono-H selectin cuts

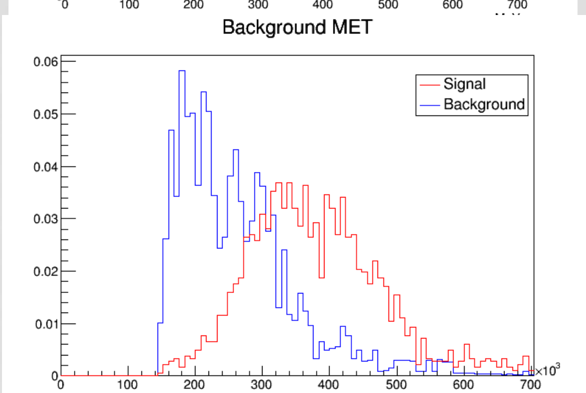


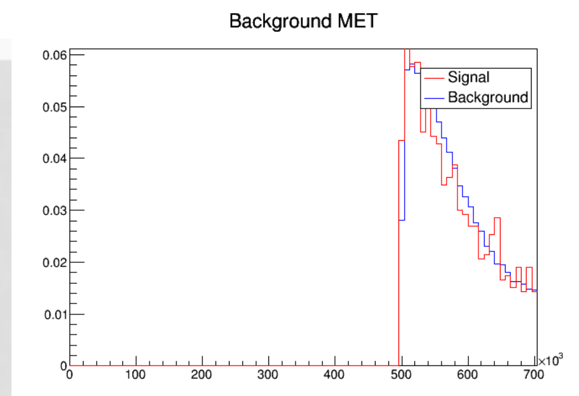
Distributions of the variables

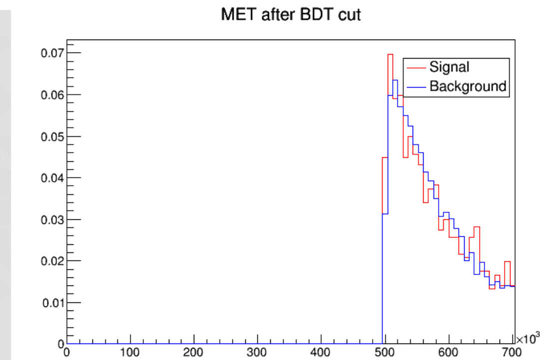
ALL MET | MET > 500 GeV

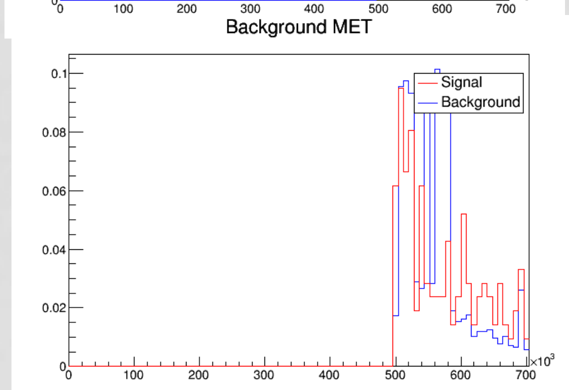








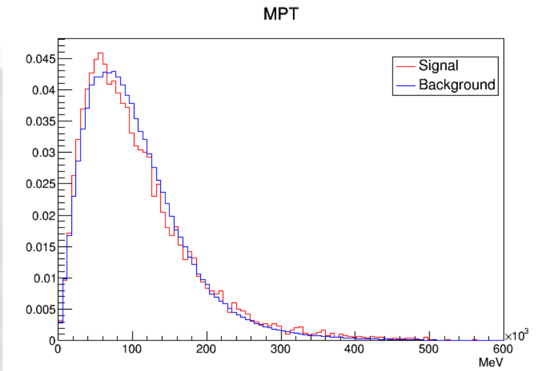


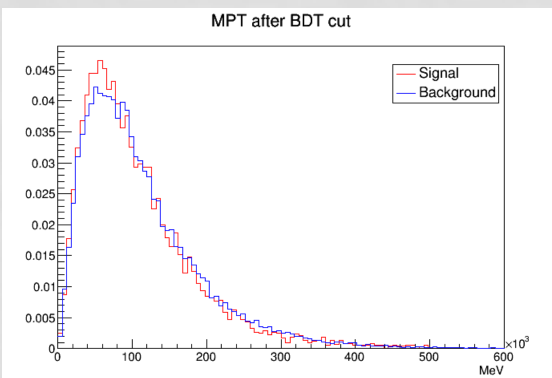


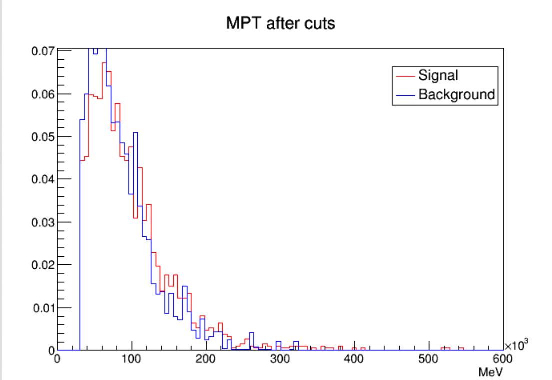
Distributions of the variables

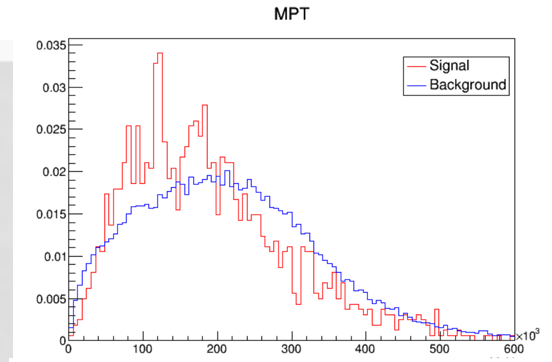
ALL MET

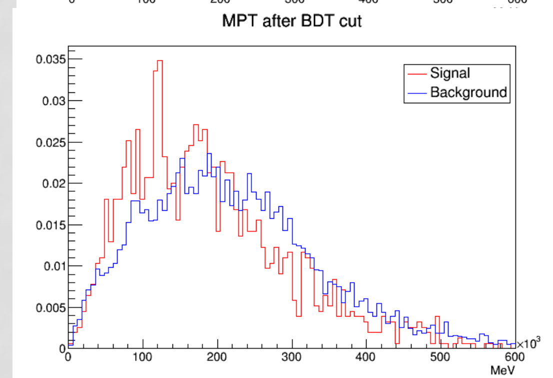
MET > 500 GeV

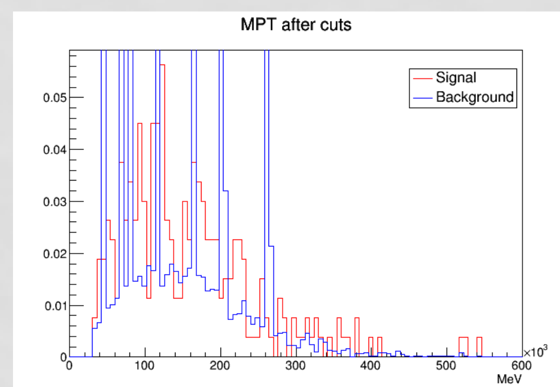






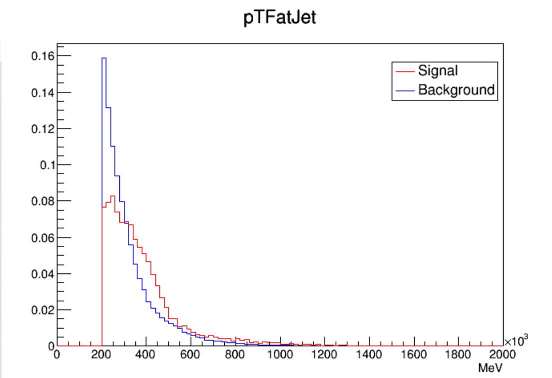


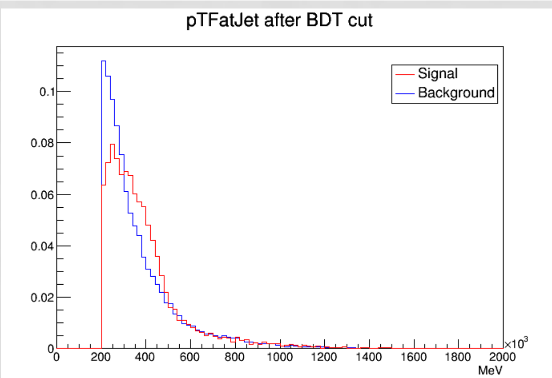


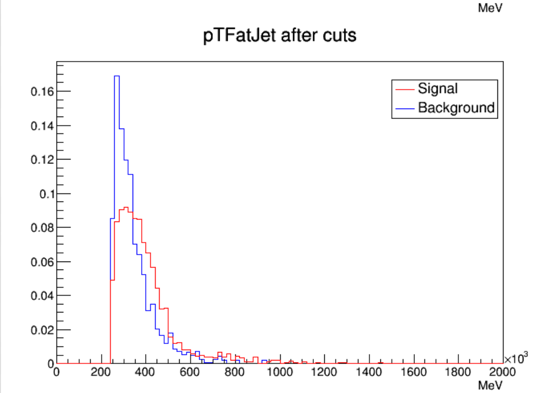


Distributions of the variables

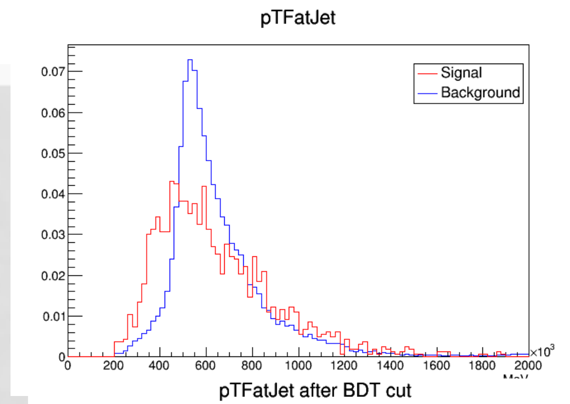
ALL MET

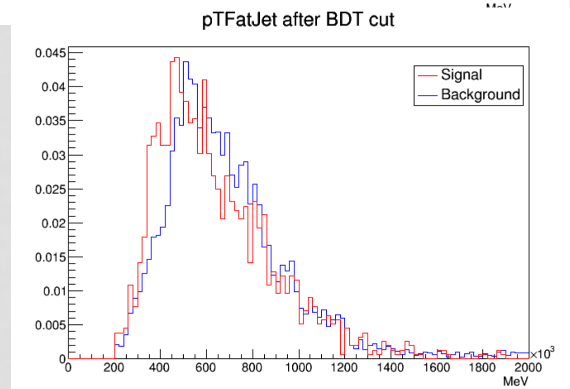


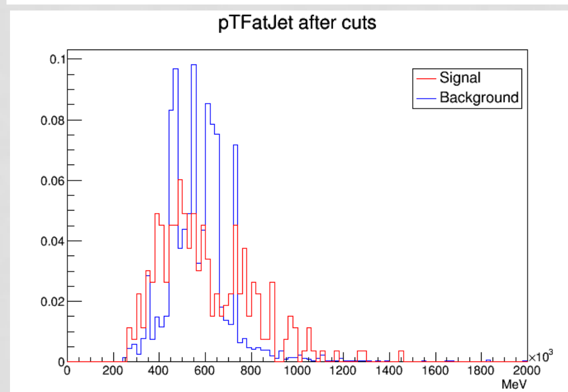




MET > 500 GeV



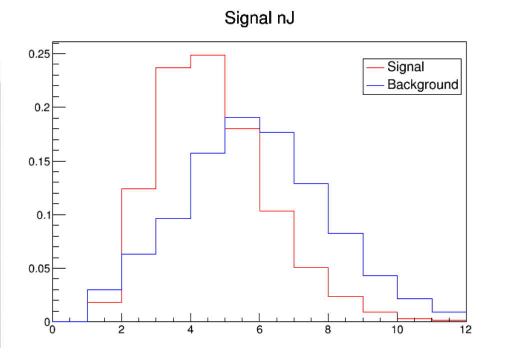


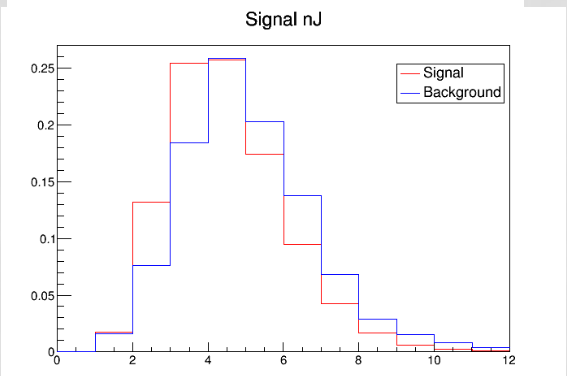


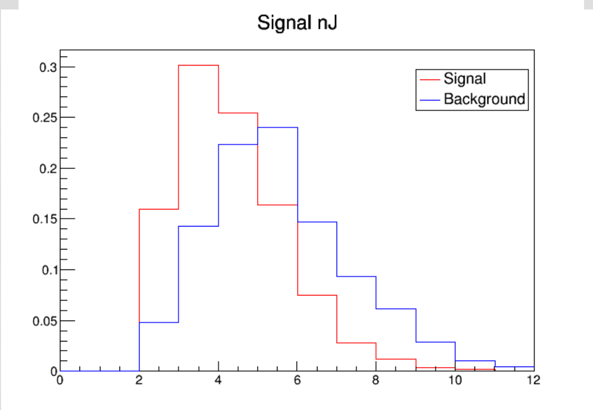
Distributions of the variables

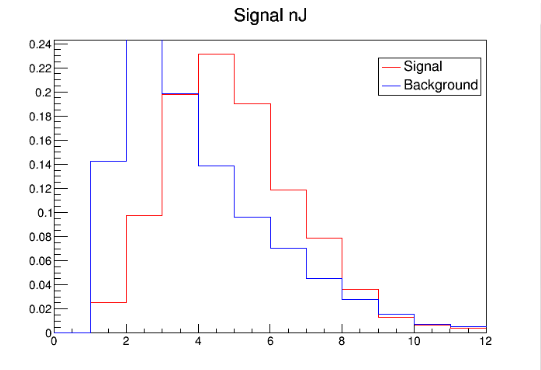
ALL MET

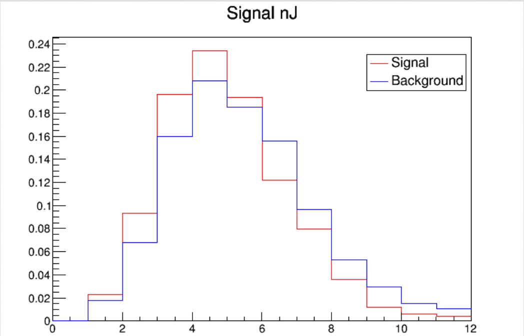
MET > 500 GeV

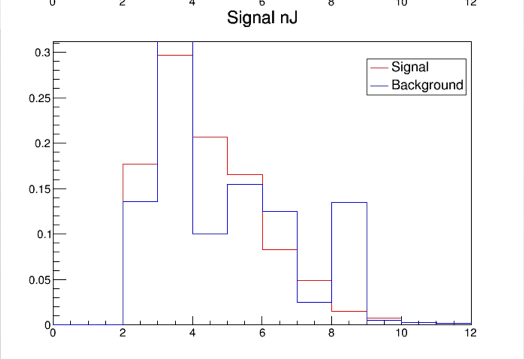






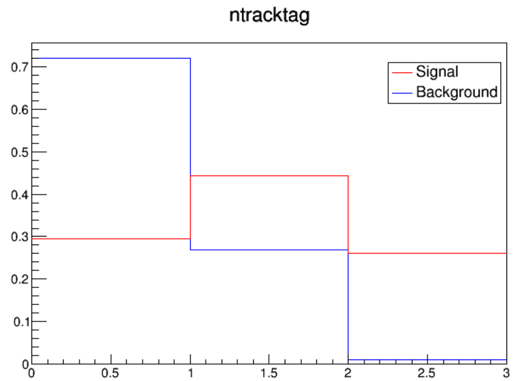


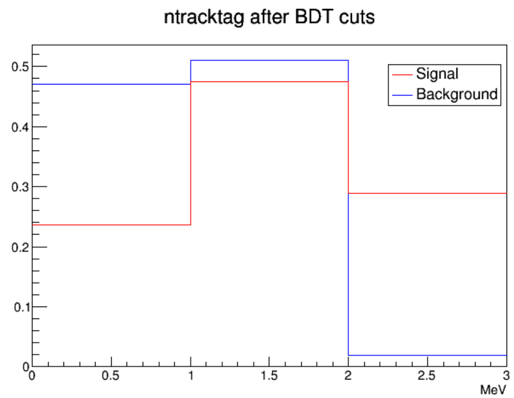


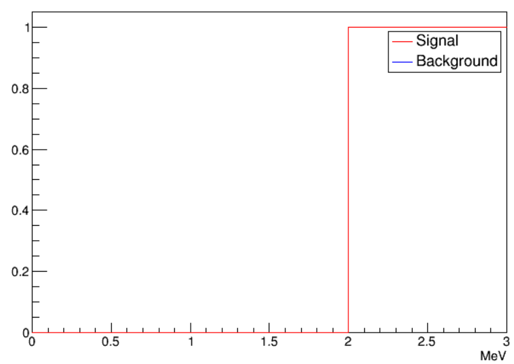


Distributions of the variables

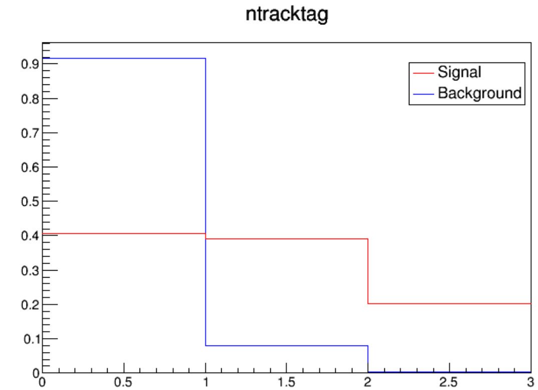
ALL MET

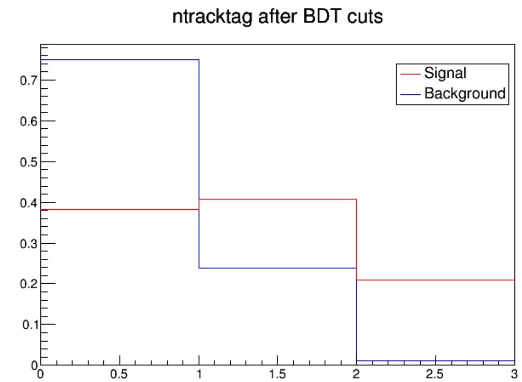


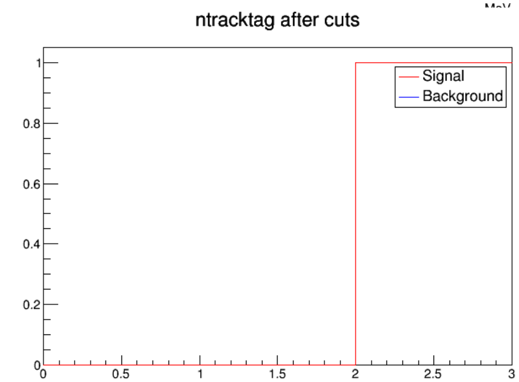




MET > 500 GeV

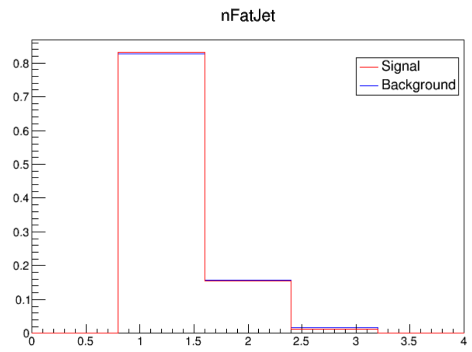


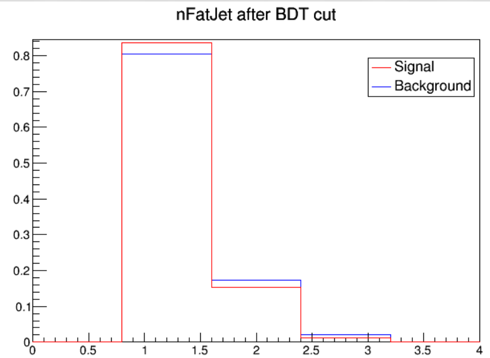


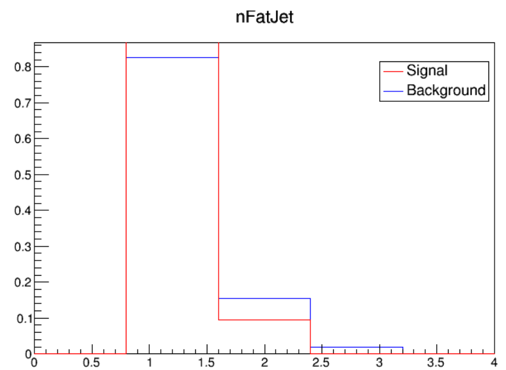


Distributions of the variables

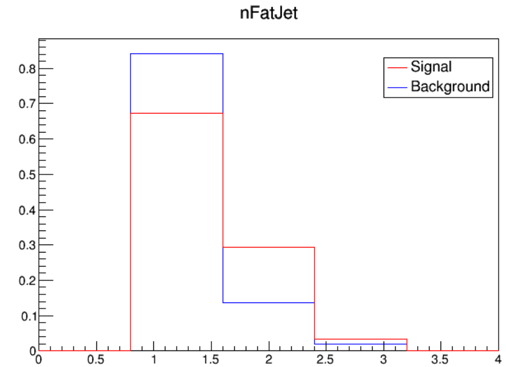
ALL MET

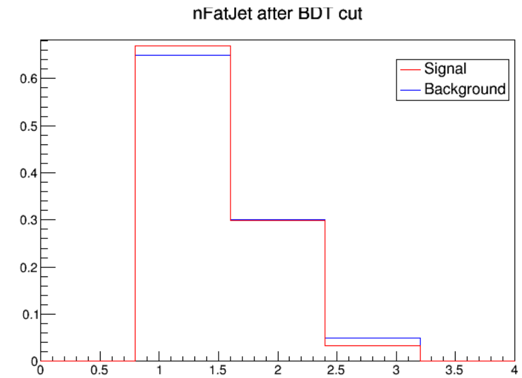


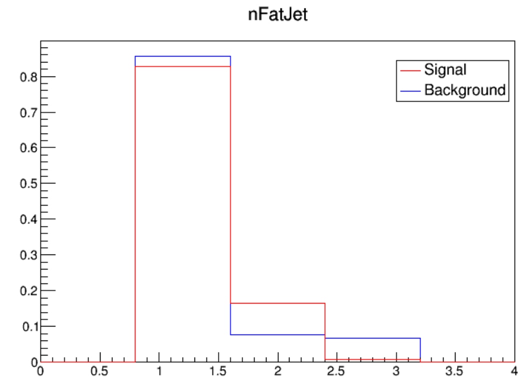




MET > 500 GeV

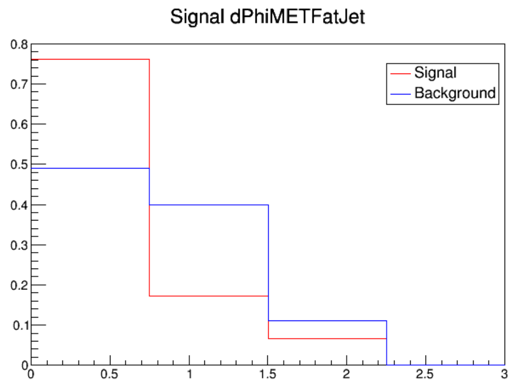


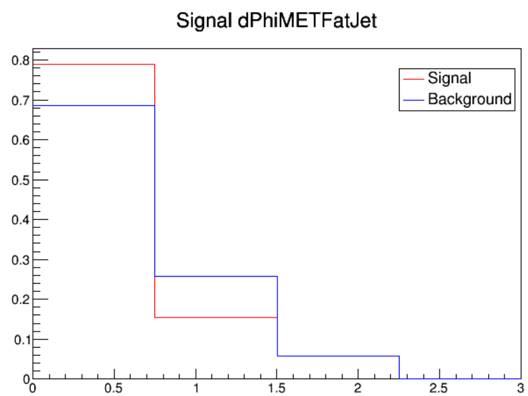


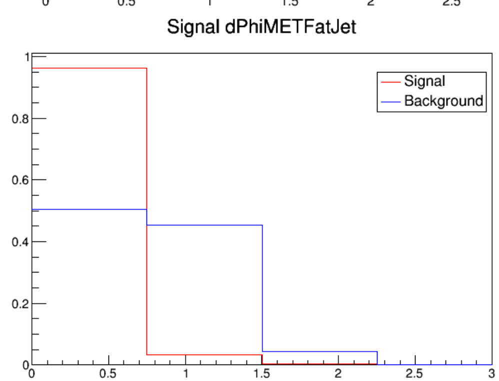


nTrackJetsTagsOut

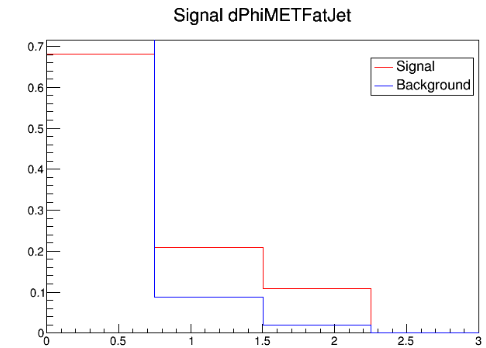
ALL MET

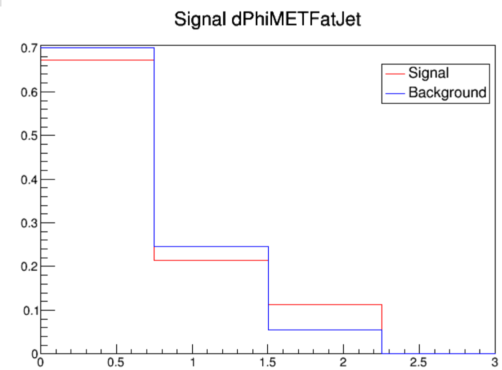


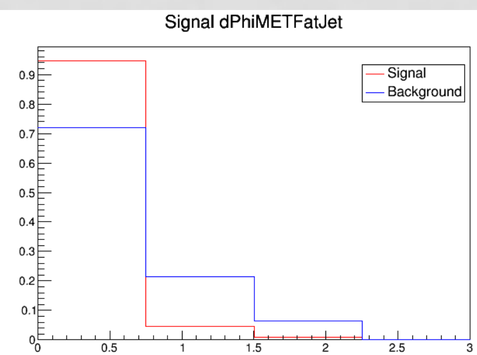




MET > 500 GeV

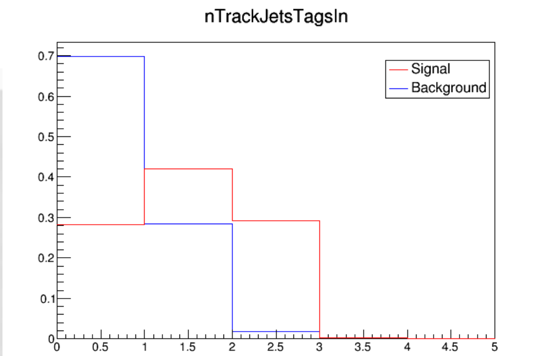


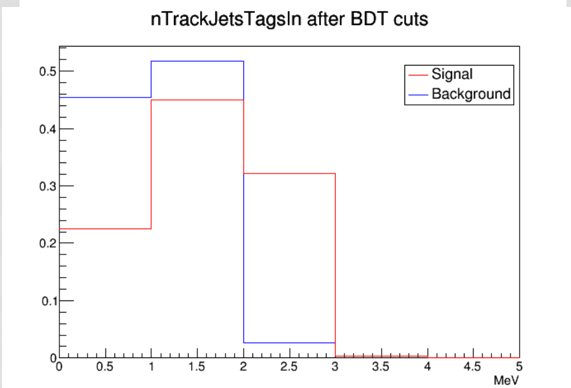


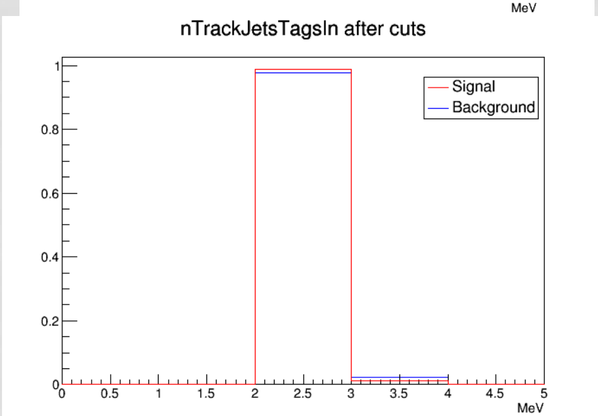


NtrackJetsTagsI n

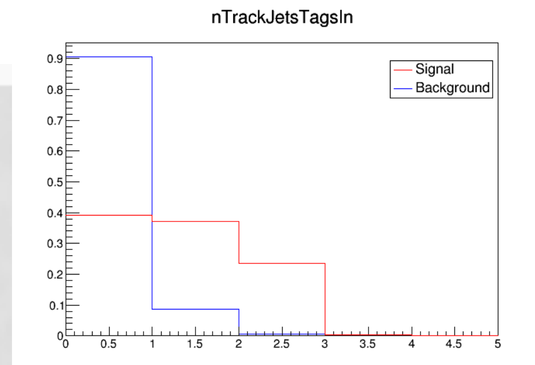
ALL MET

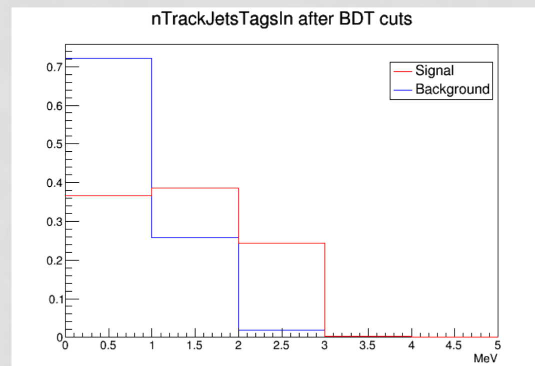


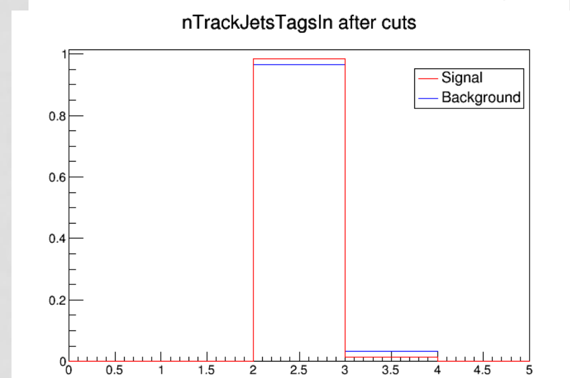




MET > 500 GeV



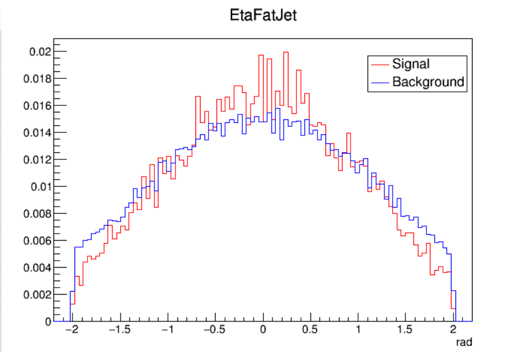


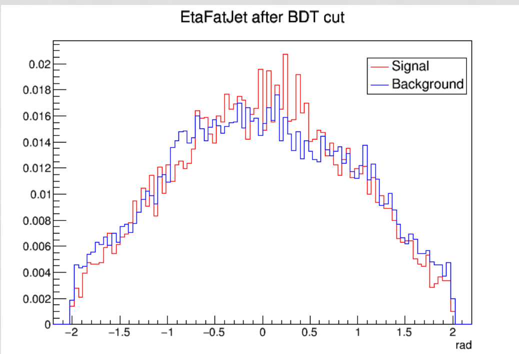


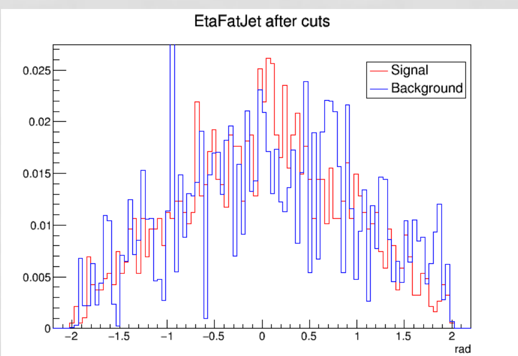
Distributions of the variables

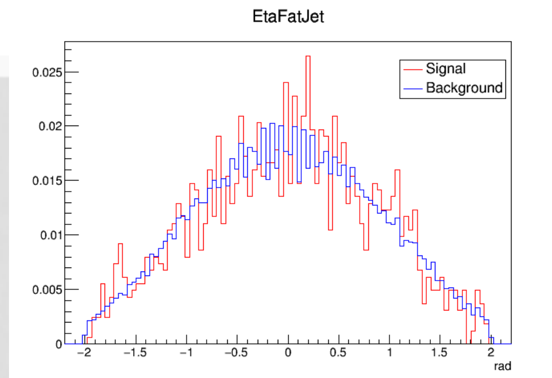
ALL MET

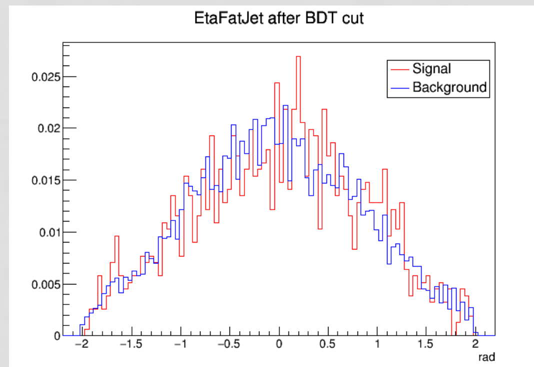
MET > 500 GeV

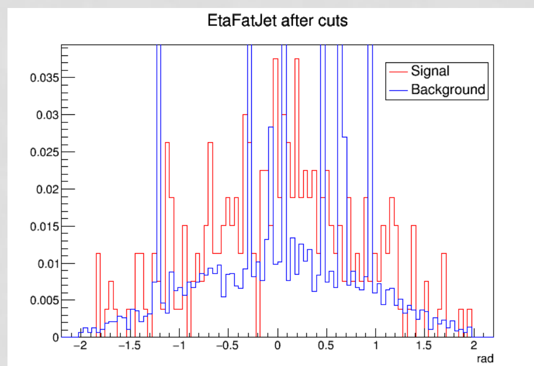






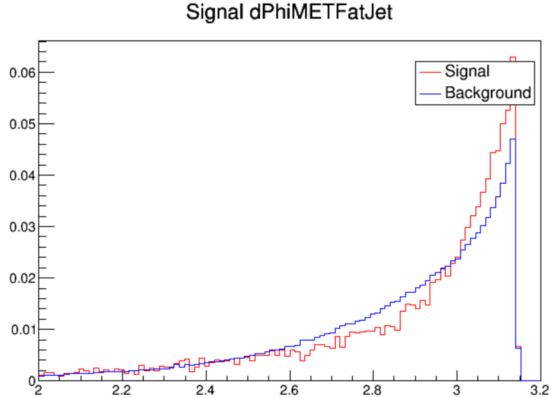


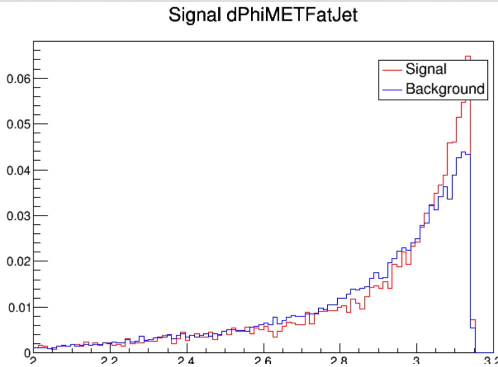


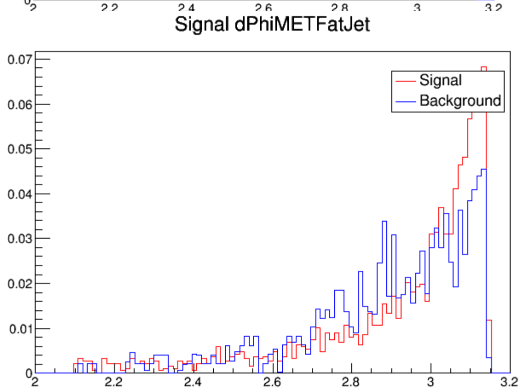


Distributions of the variables

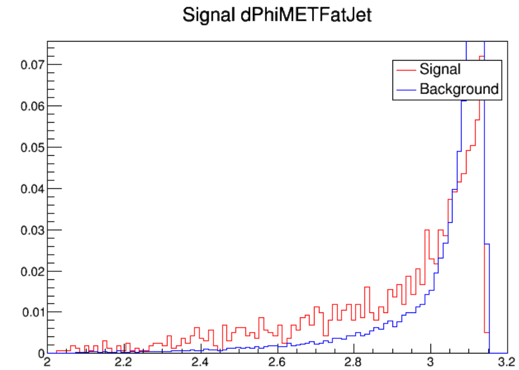
ALL MET

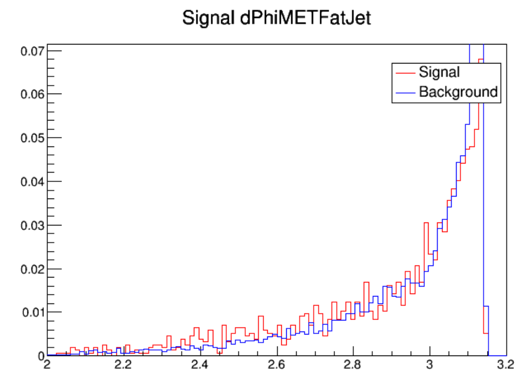


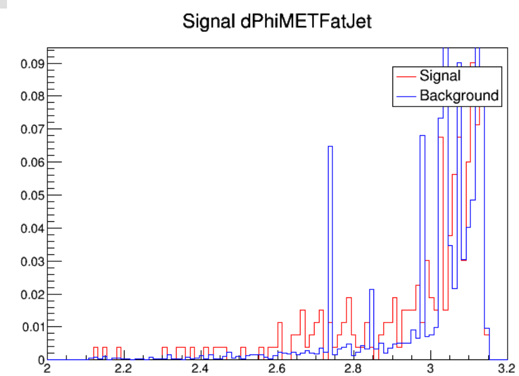




MET > 500 GeV



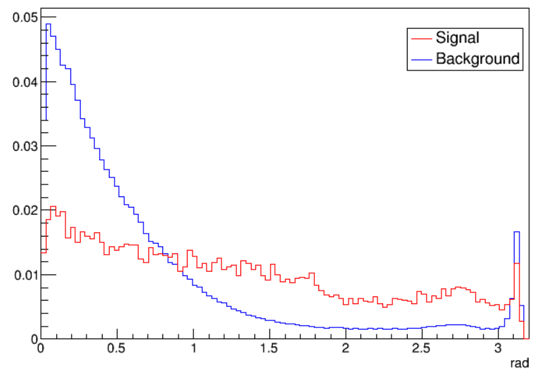




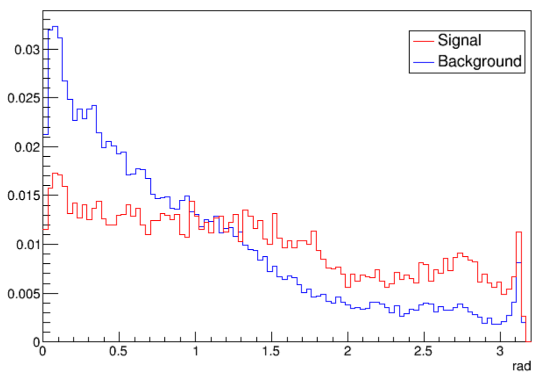
Distributions of the variables

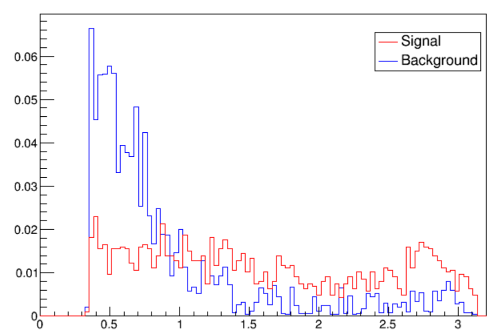
ALL MET

mindphi



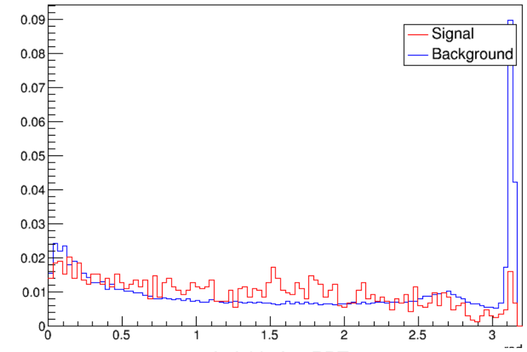
mindphi after BDT cut



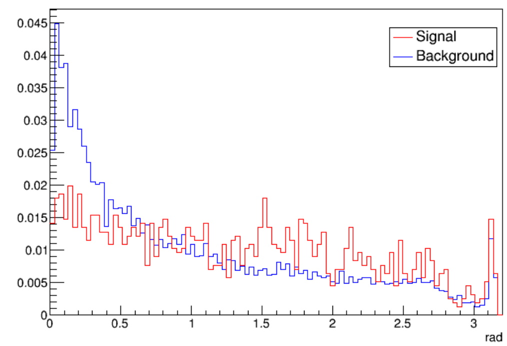


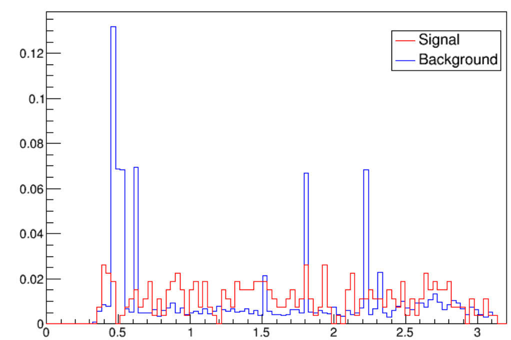
MET > 500 GeV

mindphi



mindphi after BDT cut

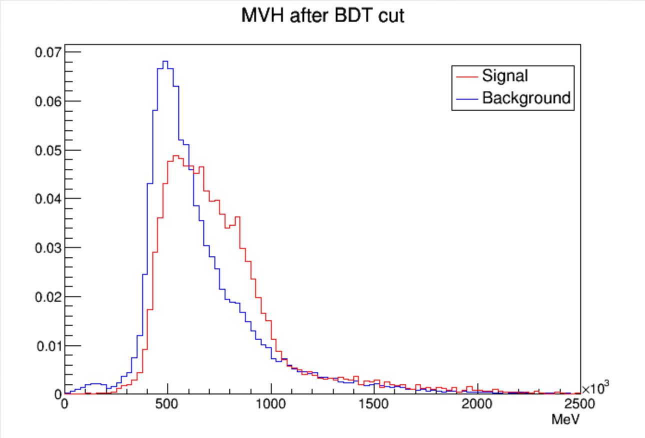


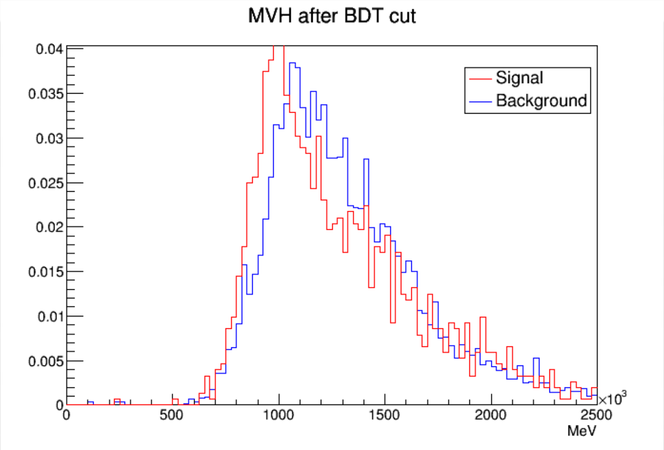


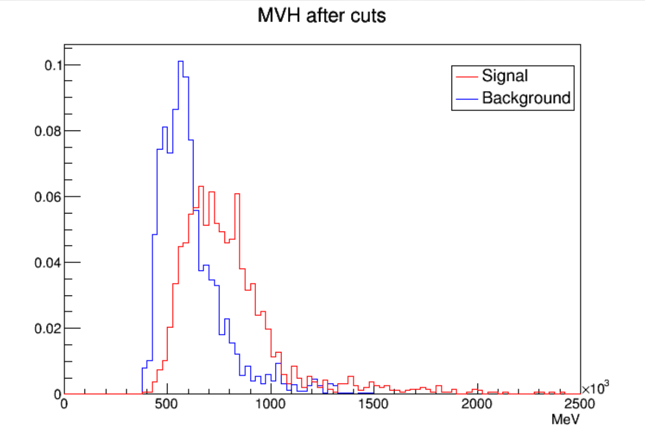
Distributions of the variables

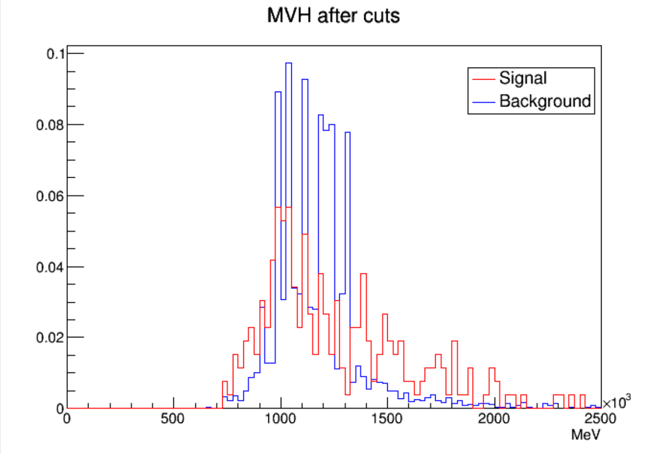
ALL MET

MET > 500 GeV





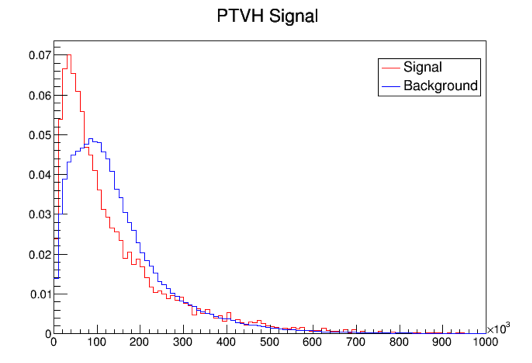


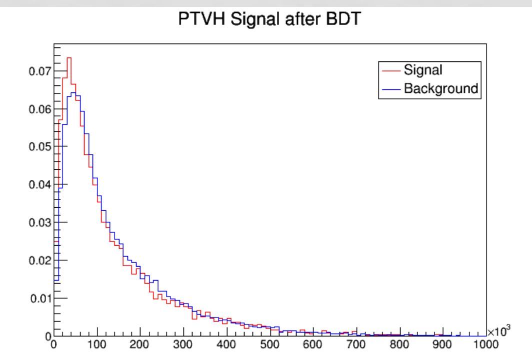


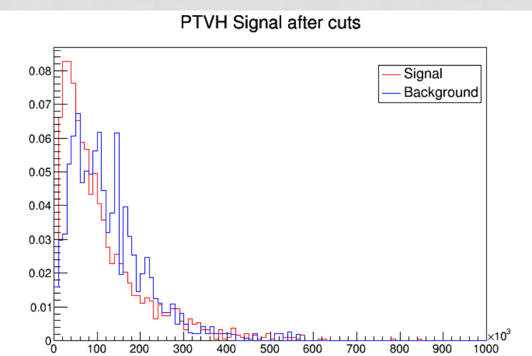
Distributions of the variables

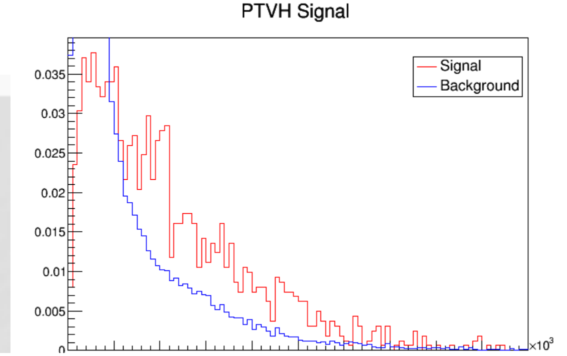
ALL MET

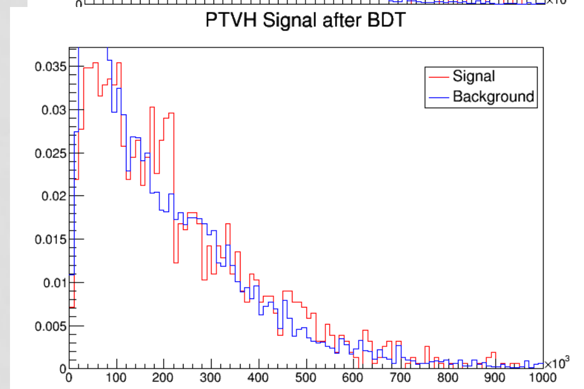
MET > 500 GeV

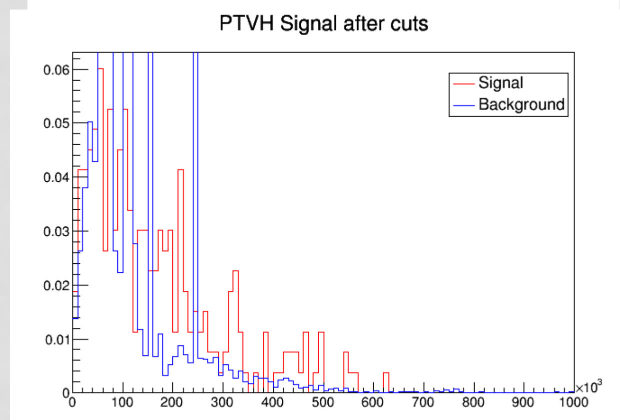








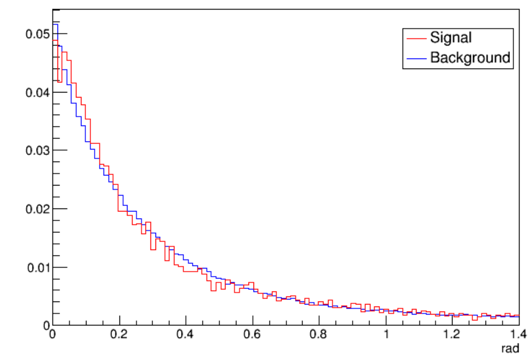




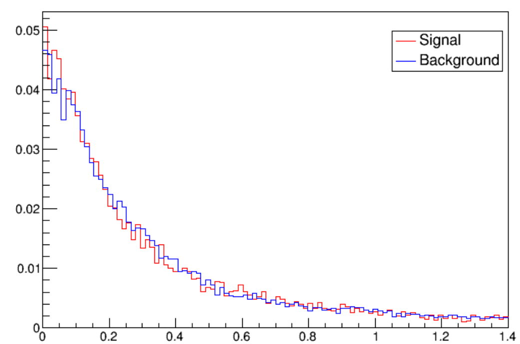
Distributions of the variables

ALL MET

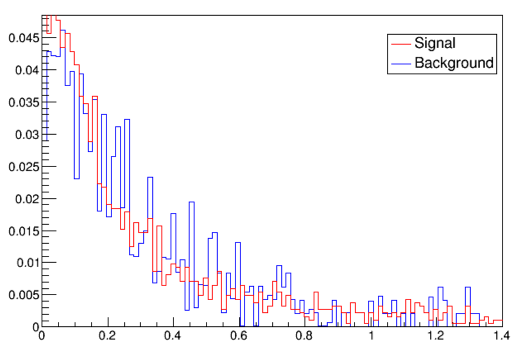
dPhiMETMPT



Signal dPhiMETMPT

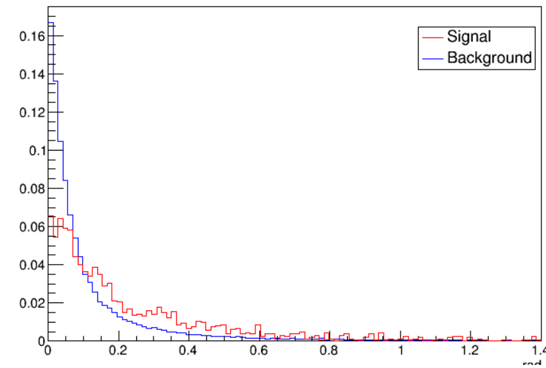


dPhiMETMPT after cuts

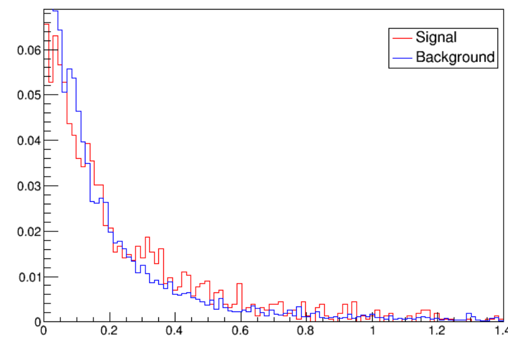


MET > 500 GeV

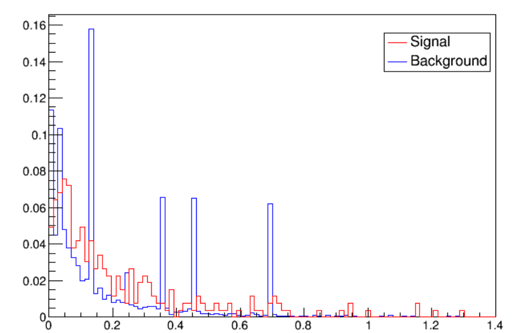
dPhiMETMPT



Signal dPhiMETMPT



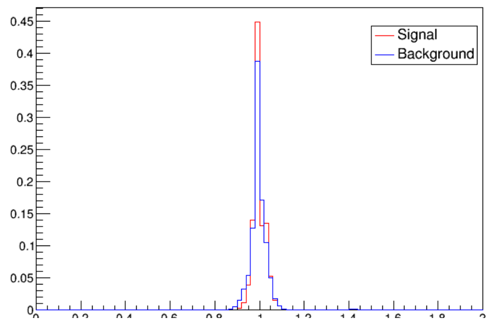
dPhiMETMPT after cuts



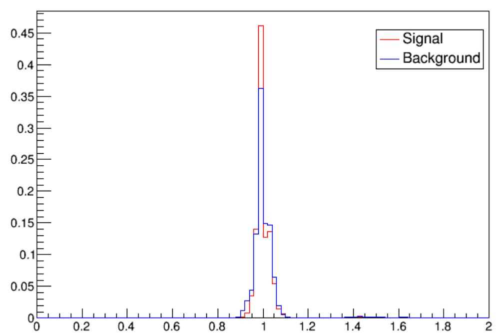
Distributions of the variables

ALL MET

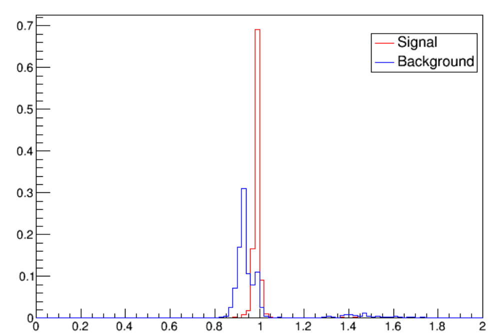
btag Signal



btag Signal after BDT

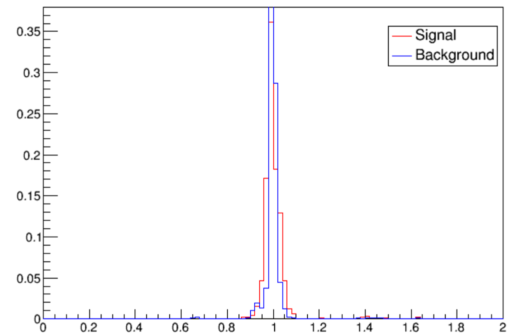


btag Signal after cuts

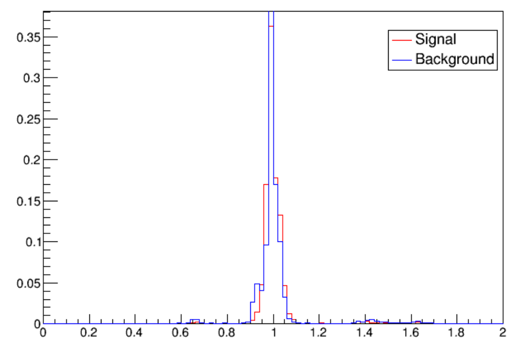


MET > 500 GeV

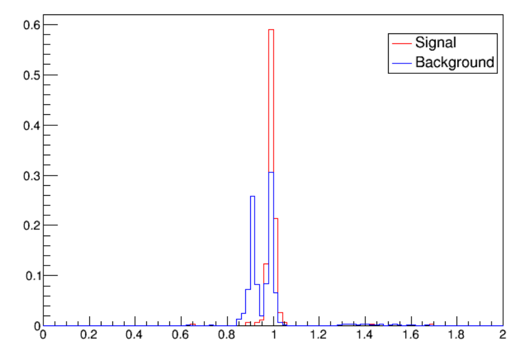
btag Signal



btag Signal after BDT

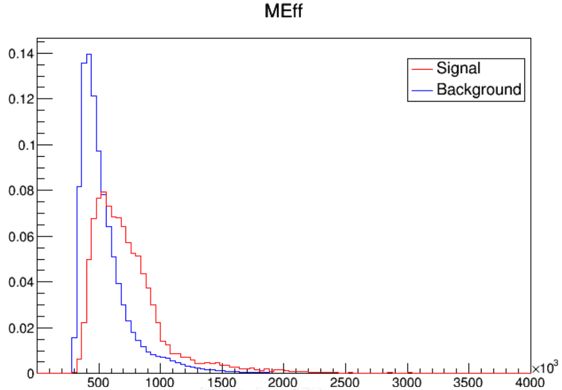


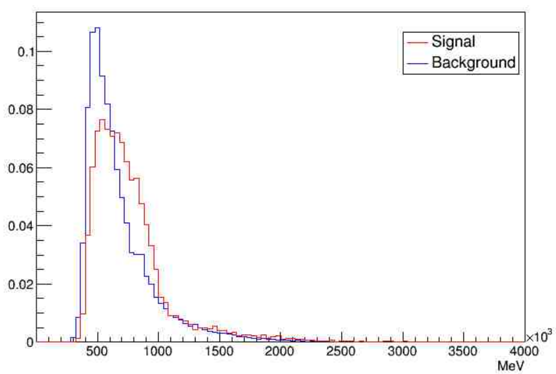
btag Signal after cuts

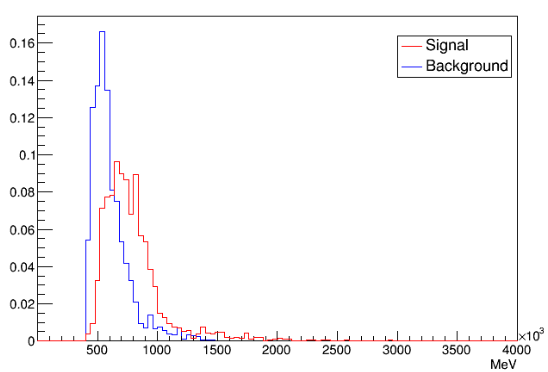


Distributions of the variables

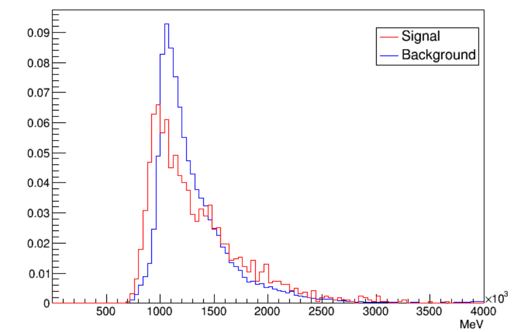
ALL MET

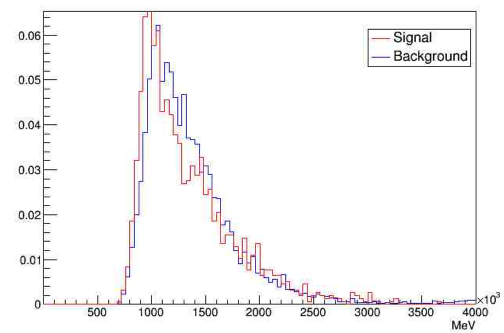


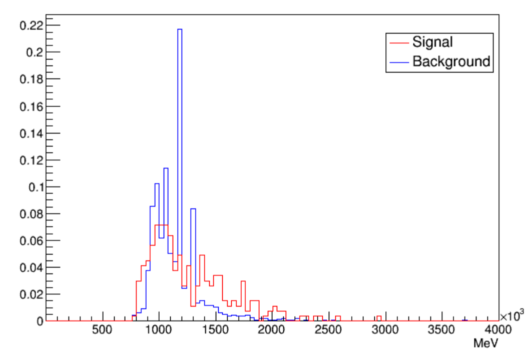




MET > 500 GeV



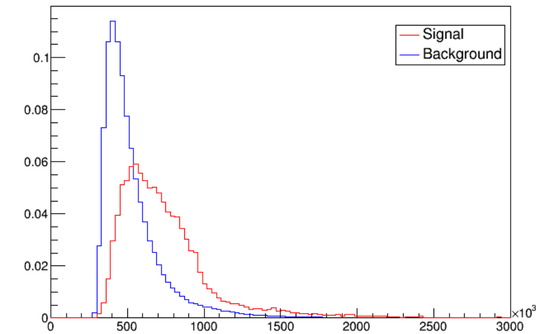




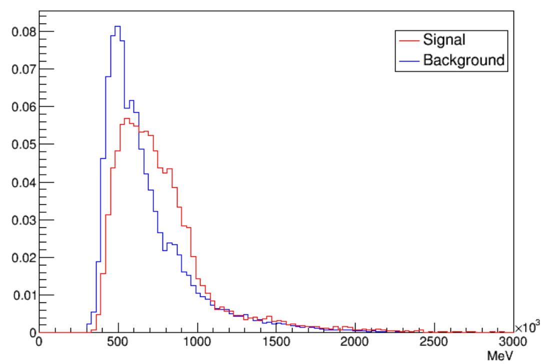
Distributions of the variables

ALL MET

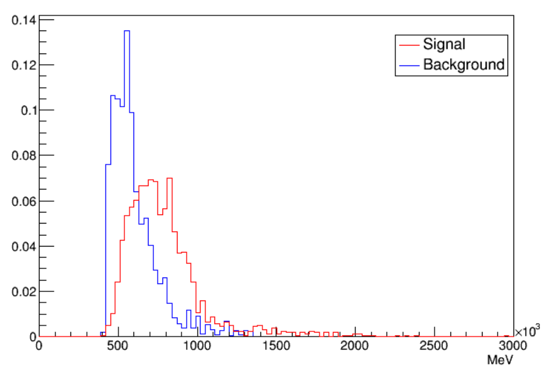
MEff3



MEff3 after BDT cuts

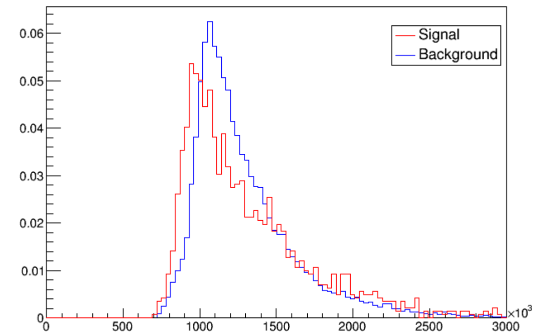


MEff3 after cuts

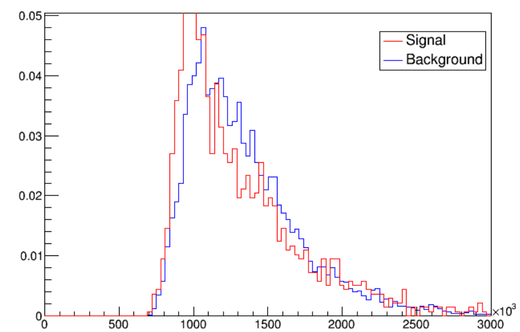


MET > 500 GeV

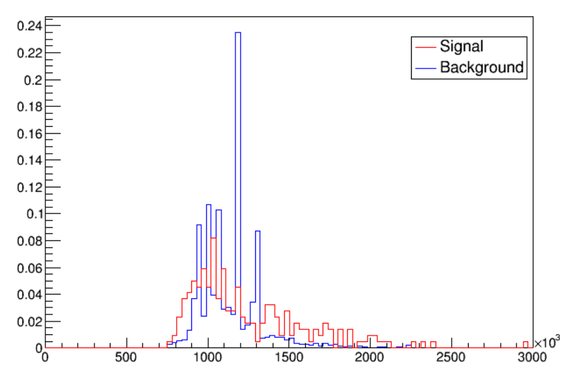
MEff3



MEff3 after BDT cuts



MEff3 after cuts



PLANS AND FIRST STUDIES

SAMPLES (Z+JETS, DI-BOSON … MC/DATA)