DARK MATTER SEARCH AT ATLAS EXPERIMENT

◆Mono-Higgs

◆Missing Transverse Energy Significance Improvement

DILIA MARÍA PORTILLO

JRJC 2016
09/12/2016

CONTENT

Dark Matter at the LHC

- Dark Matter Evidence
- Dark Matter at Colliders
- Mono-X signatures

Mono-Higgs Signatures

- Mono-Higgs Sigatures Interpretations
- Simplified Models for Mono-H (H->bb)
  - Scalar Mediator
  - Z' Vector Mediator
  - Z'-2HDM

Potential improvement with Multivariance Analysis for Mono-H

- Multivariate Analysis on Scalar Model for:
  - Merge Regime
  - BDT-/b-tag categories

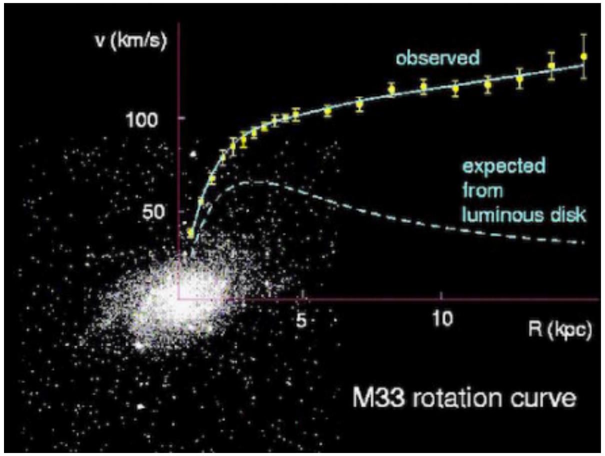
Missing Transverse Energy Signifiance Improvement

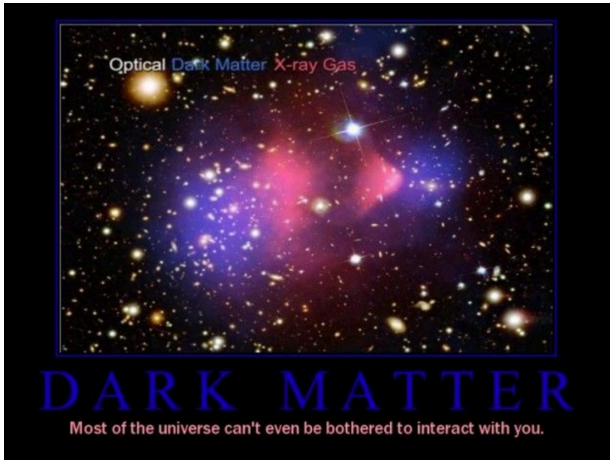
- Object based Missing Transverse Energy Significance Approach
- First performance Study

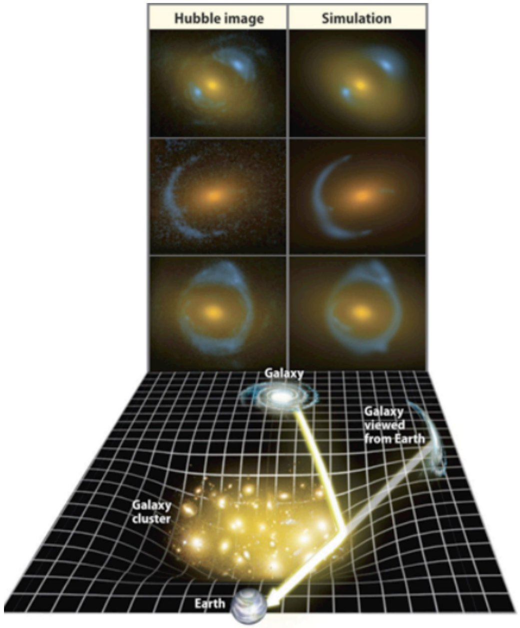
DARK MATTER EVIDENCE

Evidence of Dark Matter (DM) → Cosmological Observations

- Galaxies rotation curves
- X-ray observation of galaxy collitions
- Gravitational lensing

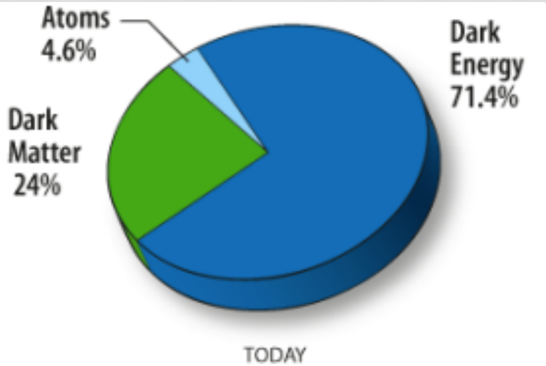






There is no evidence yet for non-gravitational interactions between DM and Stardar Model particles

Particle nature of DM is completely unknown and it represent ~25 % of the content of the Universe!

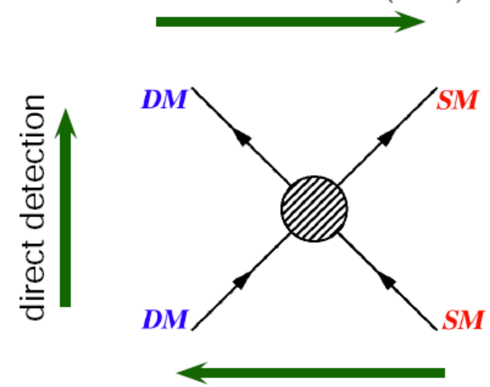


DARK MATTER (DM) AT COLLIDERS

What do we know about DM?

- How much: $\Omega \sim 0.26$
- Cold (Non-relativistic during structure formation)
- Non-baryonic
- Massive
- Electrically neutral (Dark)
- Stable
- Weakly interacting

How to detect DM?



Searches at LHC

Missing Transverse Energy (MET)

Pros and cons of DM searches @ LHC

- Pros: Independent of astrophysical uncertainties, wide range of DM/SM interactions, LHC can reach thermal cross sections.
- Cons: «Invisible» things are hard to see, DM, may not be distinct enough from SM bkg, requires interaction with some component of a proton

Which particle Theory?

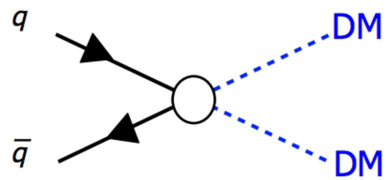
- EFT's
- Simplified Models
- Supersymmetry
- UV complete theories
- Higgs Portal
- More Exotic (Long-lived decays, hidden valleys)

Which signatures?

- SUSY searches
- Dijets
- Exotic Decays
- Mono-X

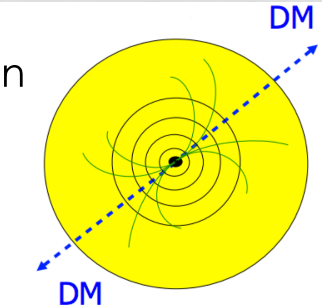
MONO-X SIGNATURES

The basic diagram of DM production and detection



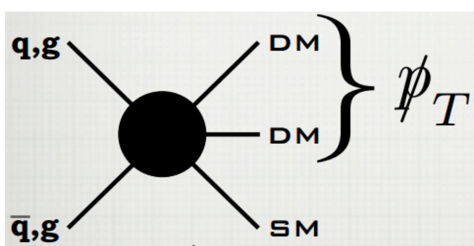
DM doesn't interact in the detector

2 back-to-back invisible particles

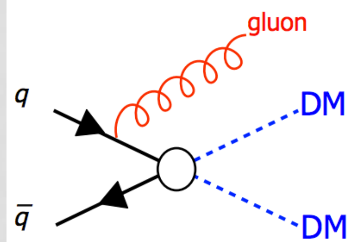


The "Mono-X" Topology:
Standar Model (SM) particles ('X') recoils against missing transverse momentum

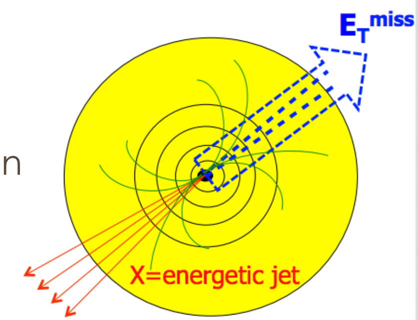
$$pp \to \not{E}_T + X$$



E.g. If **q** radiates a **'X'**



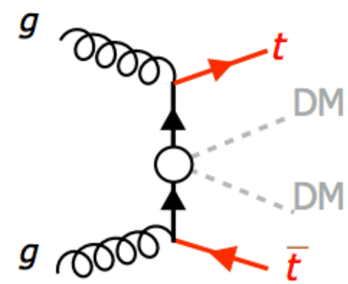
E and P conservation

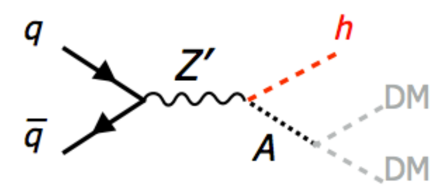


One name, many processes…

- ✧Jet + E_T^{miss}
- ✧W/Z (→lep/jet) + E_T^{miss}
- ✧t/b + E_T^{miss}
- ✧γ + E_t^{miss}
- ✧H + E_T^{miss}

Not only ISR, **'X'** may be more closely connected to DM production, e.g.:



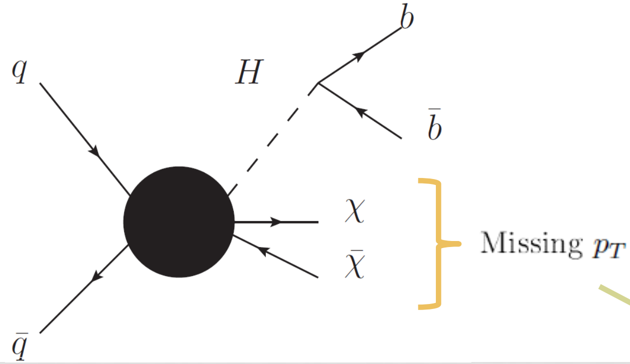


MONO-HIGGS SIGNATURES (H->bb)

DM pair production in association with a Higgs boson decaying in to b quarks H-> bb

- No Initial State Radiation Higgs

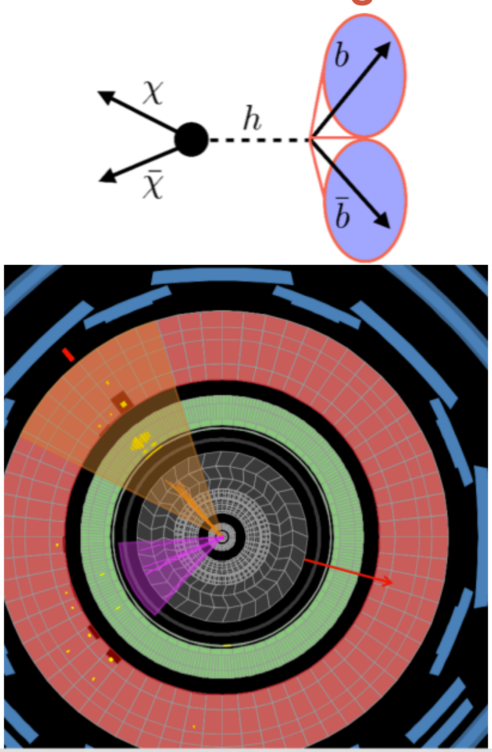
Provides direct probe of DM-SM coupling



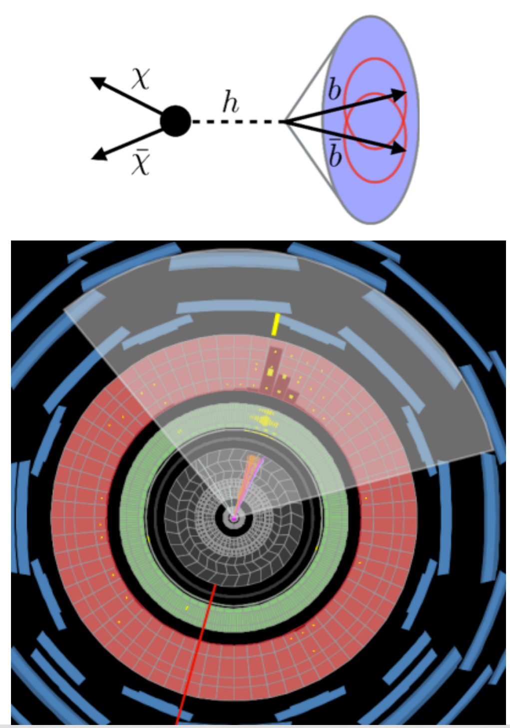
- H -> bb dominant decay mode

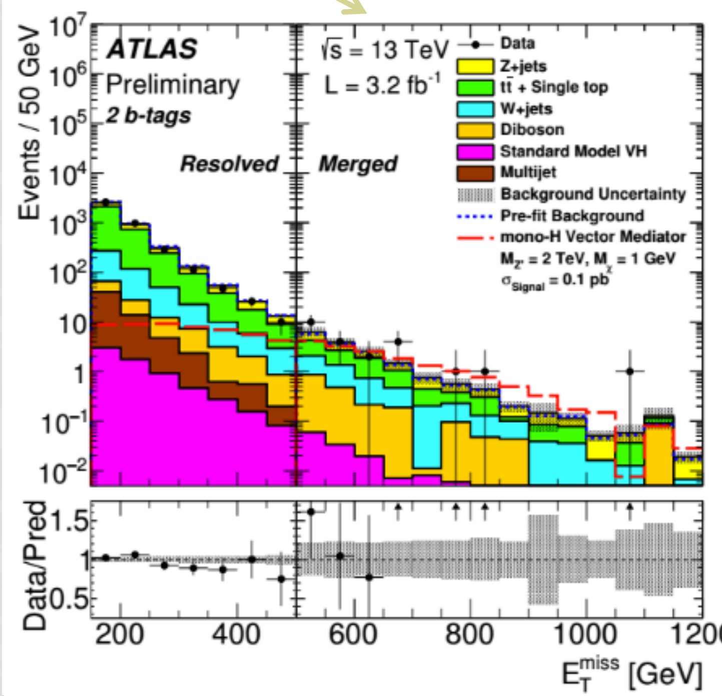
(largest cross section).

Resolved Regime



Merged Regime

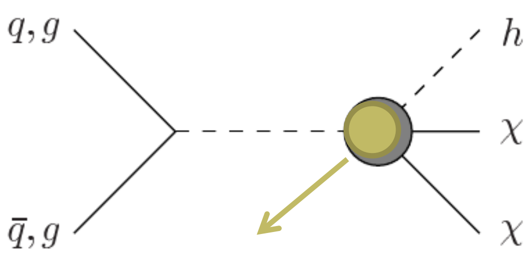




MONO-HIGGS SIGNATURES INTERPRETATIONS

Efective Field Theory framework

Contact operators



$$\lambda|\chi|^2|H|^2$$

$$\frac{1}{\Lambda}\bar{\chi}i\gamma_5\chi|H|^2$$

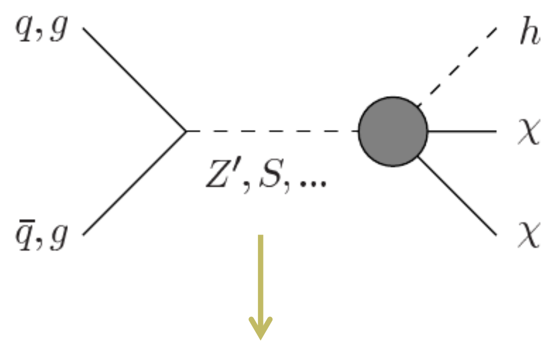
$$\frac{1}{\Lambda^2}\chi^\dagger\partial^\mu\chi H^\dagger D_\mu H$$

$$\frac{1}{\Lambda^4}\bar{\chi}\gamma^\mu\chi B_{\mu\nu}H^\dagger D^\nu H$$

- Pro
  - Generic interpretation
    - Model independent
- Con
  - Not valid at all momentum transfer

Simplified Models

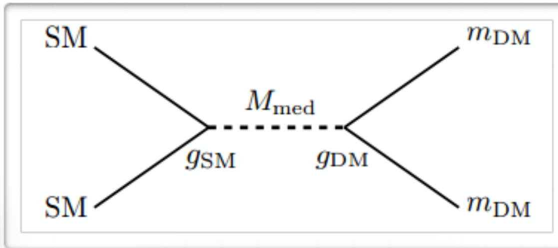
Minimal number of renormalizable operators



The DM and visible sectors are coupled through a new massive mediator

- Pro
  - UV complete
- Con
  - Less generic (specific number of parameters)
  - Too many exotic models that cannot be reduced to these models.

SIMPLIFIED BENCHMARK MODELS



Consider comprehensive set of diagrams for mediator

Vector	Axial-vector
Scalar	Pseudoscalar

DM

Dirac fermion	Scalar - real
Majorana fermion	Scalar - complex

Define simplified model with (minimum) 4 parameters

Mediator mass (M_{med})	DM mass (M_{DM})
g_q	g_{DM}

4-dimensional problem, projecting limits onto a 2-D plane

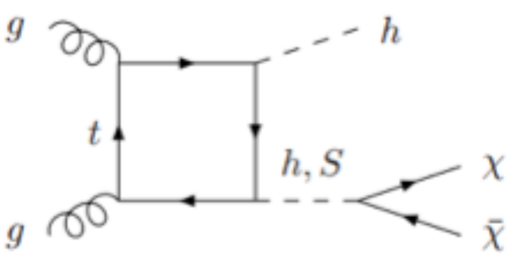
- **Scalar Mediator Model for Mono-H**

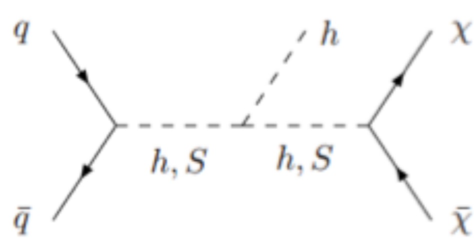
Potential :

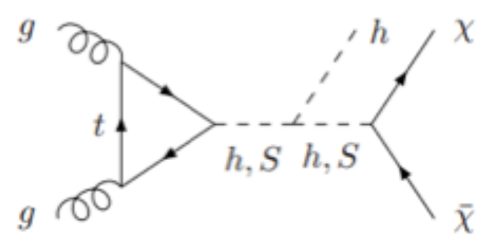
$$V \supset a|H|^2 S + b|H|^2 S^2 + \lambda_h |H|^4$$

After SSB

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







SIMPLIFIED BENCHMARK MODELS

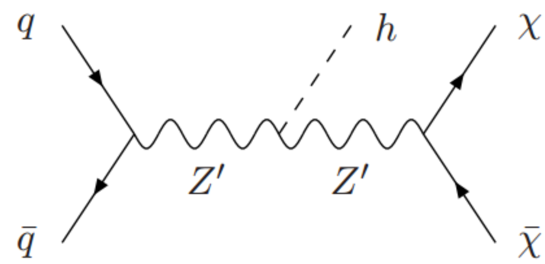
- **Baryonic Z'**

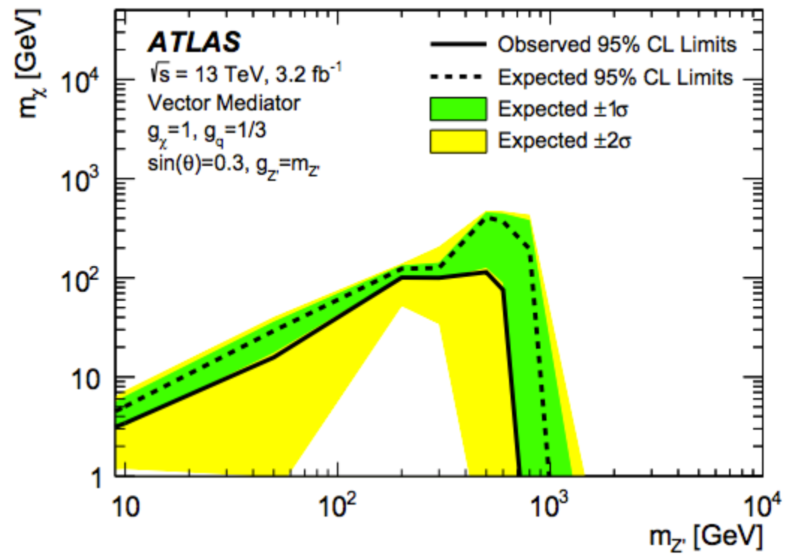
$$g_q \bar{q}\gamma^\mu q Z'_\mu + g_\chi \bar{\chi}\gamma^\mu \chi Z'_\mu$$

After SSB

$$-g_{hZ'Z'} h Z'_\mu Z'^\mu$$

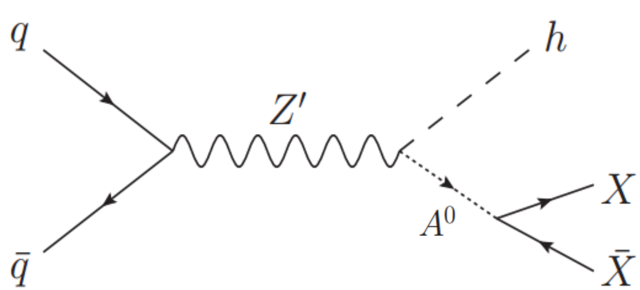
$$g_{hZ'Z'} = \frac{m_{Z'}^2 \sin\theta}{v_B}$$

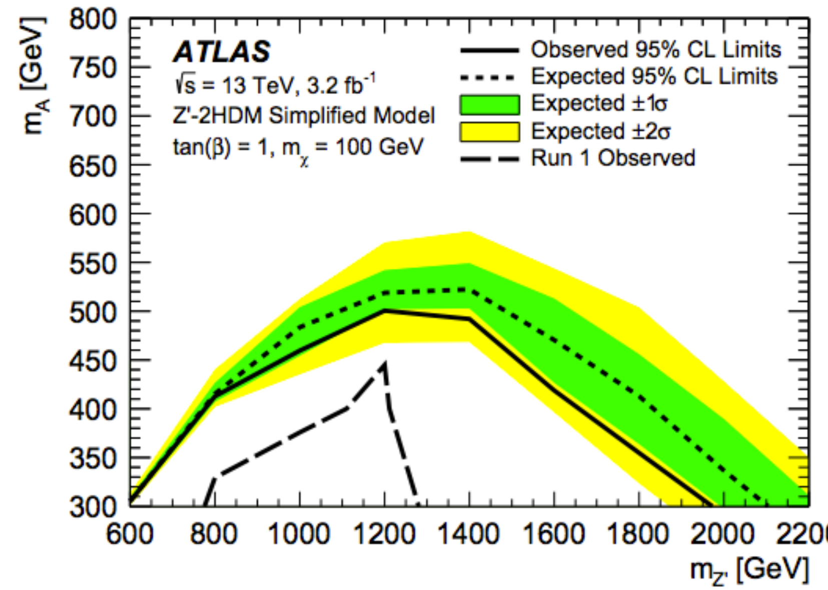




- **Z'-2HDM**

$$-\mathcal{L} \supset y_u Q \tilde{\Phi}_u \bar{u} + y_d Q \Phi_d \bar{d} + y_e L \Phi_d \bar{e} + \text{h.c.}$$





POTENCIAL IMPROVEMENT IN THE MONO-H (H->bb) SEARCH
Multi Variate Analisis on Scalar Mediator Model

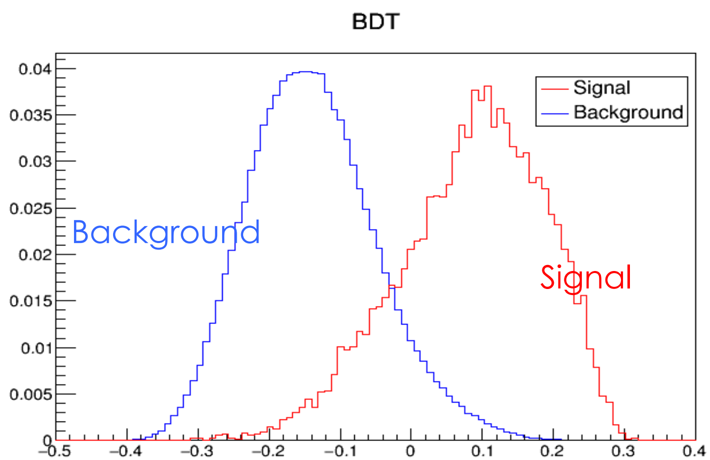
- **Multi Variate Analisis:** Use several observables from the events and their correlations to form ONE combined variable and use it in order to discriminate between "signal" and "background".
- **Plan:** Study a multivariate approach for mono-Higgs search and compare the potential improvement to the current cut-based analysis

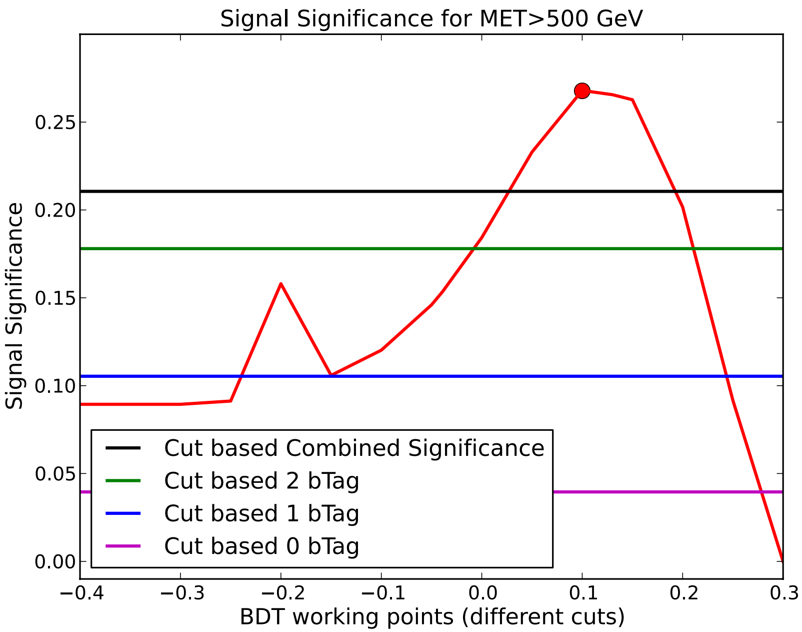
Boosted Desition Tree (BDT) on Merged Regime (MET > 500 GeV)

In the training: Kinematic variables for the bb system, variables related to AntiQCD cuts, number of b-tagged jets, fat jets, additional jets, small radius jets and missing transverse energy and momentum

ALL MET RANGE

The BDT method was trained for all the MET range.

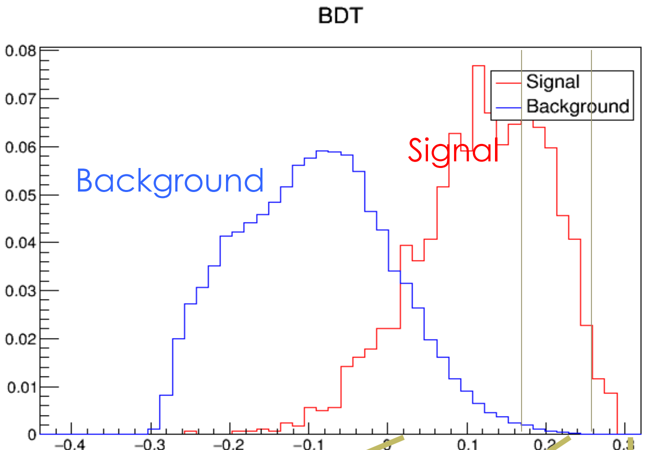




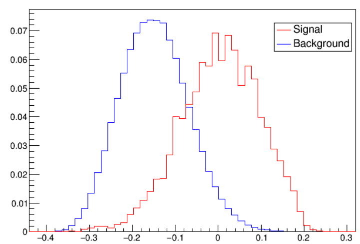
Significance improvement is +28%

Boosted Desition Tree (BDT) on Merged Regime For Scalar Mediator Model

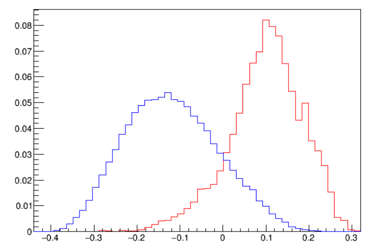
The number of b-tag jets is a variable considered in the BDT training ⟶ We can define BDT categories that can be related to 0, 1 and 2 btags.



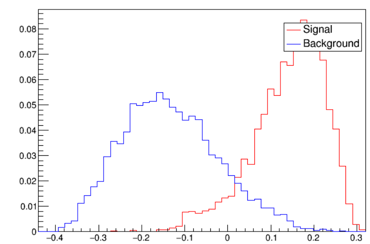
- Low BDT
- 0 b tag

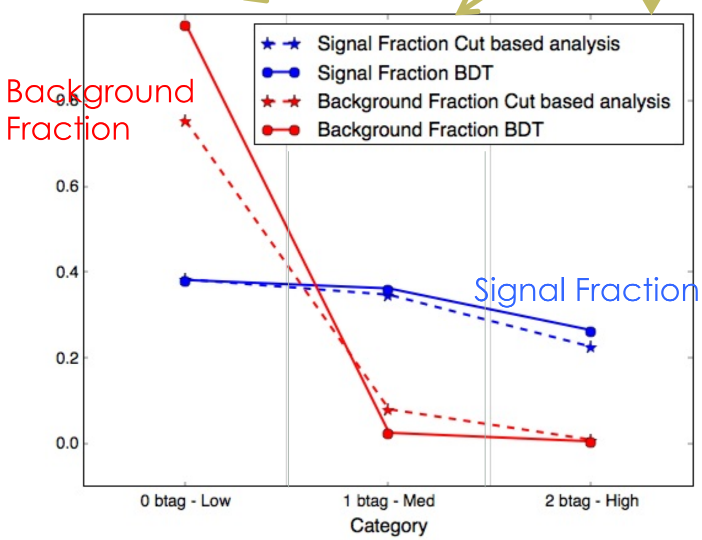


- Medium BDT
- 1 b tag



- High BDT
- 2 b tag





Better separation power.

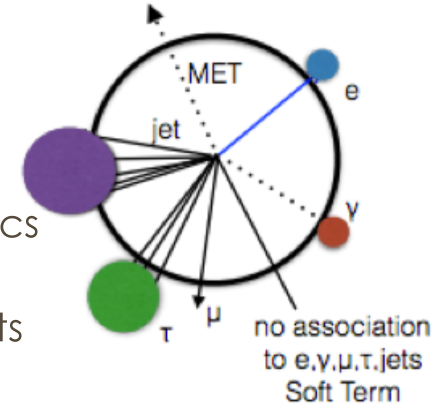
	Cuts Based Analysis			BDT		
	0 btag	1btag	2 btag	Low	Medium	High
Signal Significance	0.039	0.105	0.178	0.034	0.177	0.23
Combined Sign.	0.2104368			0.2929455		

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

MISSING TRANSVERSE ENERGY SIGNIFICANCE IMPROVEMENT

- **Missing Transverse Energy:** Fundamental piece for many of the ATLAS analysis (DM searches)

$$\sum_i \vec{p}_i = 0 \quad \xrightarrow{\text{When something is missing}} \quad \vec{E}_T^{\text{miss}} = -\sum_{\text{observable}} \vec{p}_i$$



Two effects could imply an imbalance in the total transverse momentum

- **Non-interacting particles** (True MET): SM particles (neutrinos), new physics (DM, SUSY, etc.)
- **Fake detection** (Fake MET): Objects misreconstruction, detector effects (dead regions).

Reconstructed Met = Calorimeter Objects + Muons + Soft Terms

Calorimeter Objects: Electrons, Photons, Taus, Jets

Soft Terms: Every signal which can't be clearly identified/calibrated.

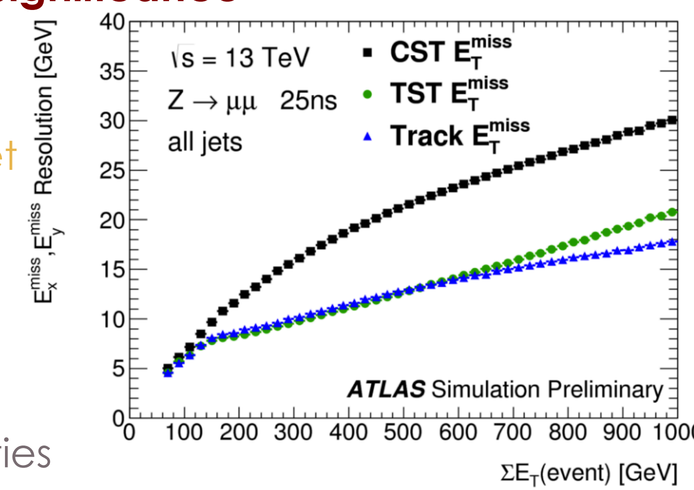
Missing Transverse Energy (MET) Significance

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.

Currently defined as:

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

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



DEFINITION OF 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.

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

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.

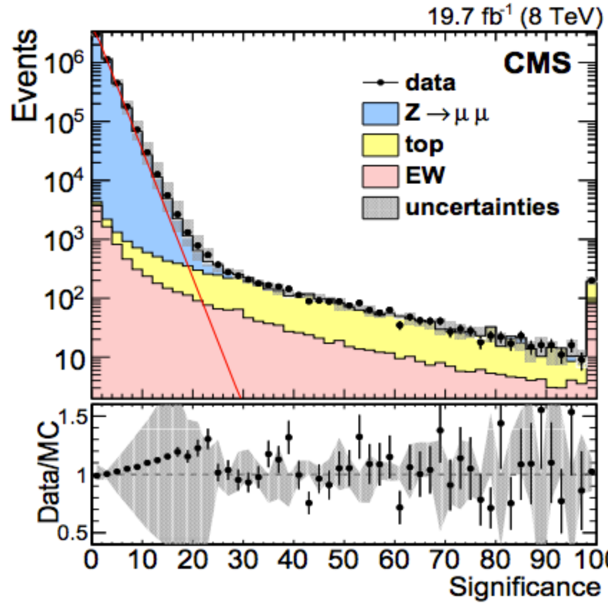
If we assume that …

- The sum of all the truth transverse momentum is equal to zero
- Gaussian uncertainties distributions

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

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

- Plan: Implement this definition in ATLAS and evaluate if it will actually improve analysis needing MET Significance



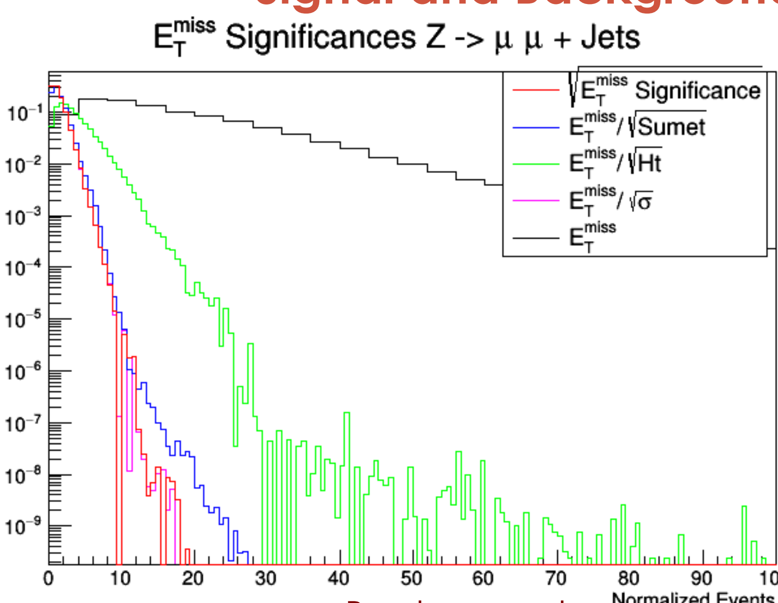
MET SIGNIFICANCES SEPARATION POWER

- Comparison of the separation power between
  - Background: Z-> $\mu\ \mu$ + jet ⟶ No genuine MET
  - Signal: ZZ-> $\mu\ \mu\ \nu\ \nu$ + jet ⟶ Neutrinos!

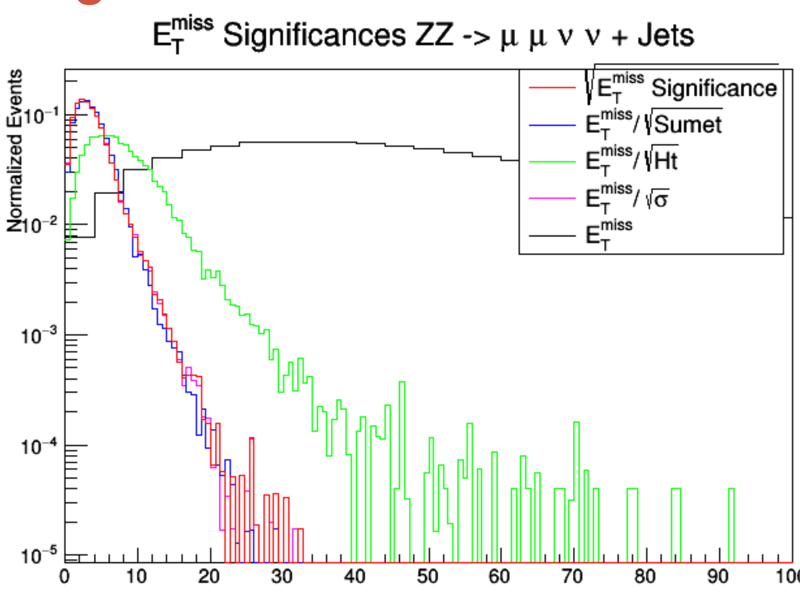
- Cutting on…
  - MET
  - MET significance
  - MET/sqrt(Sumet)
  - MET/sqrt(Ht)

- Propagationg resolutions of…
  - Jets
  - Muons
  - Electrons
  - Soft Terms

Signal and Background Met Significance Distributions



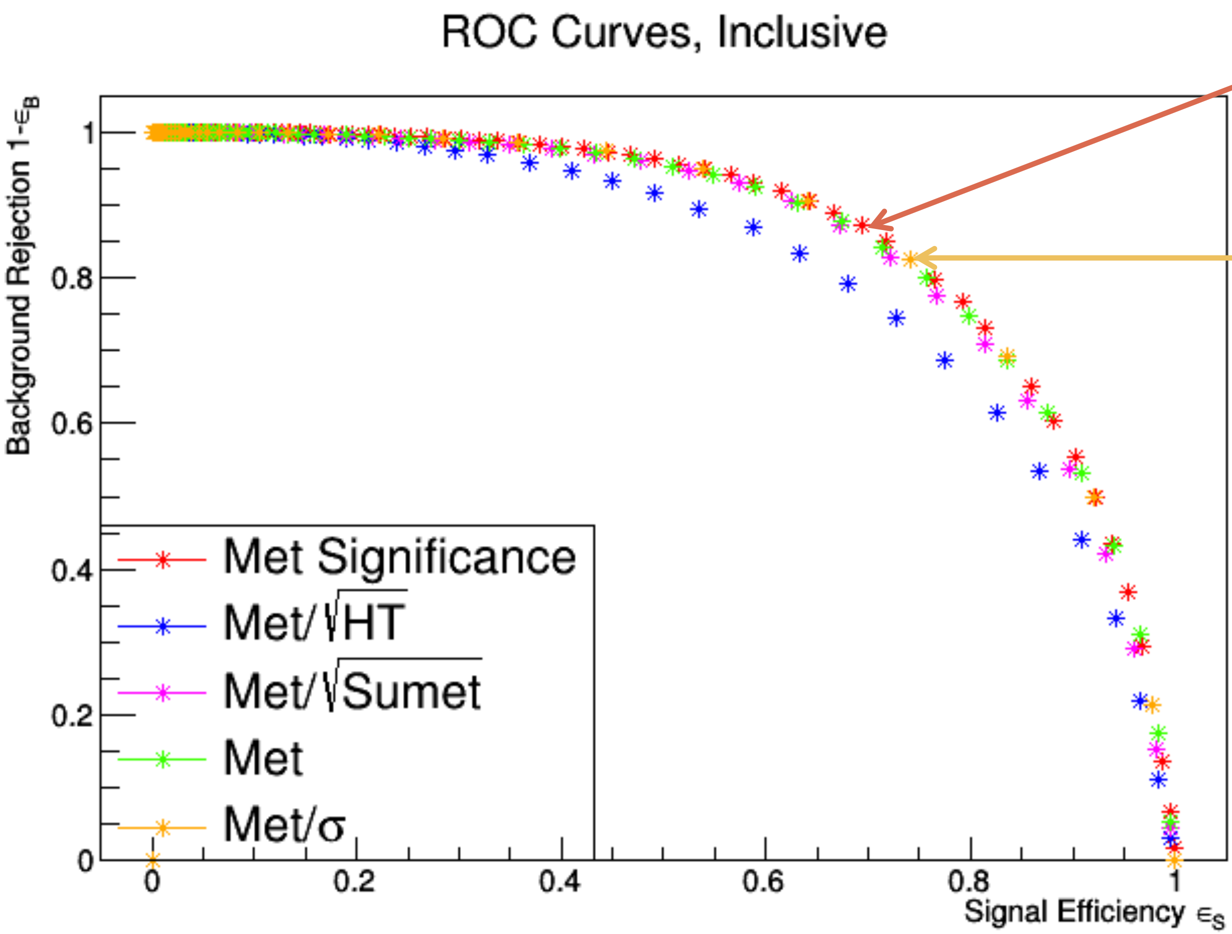
Background



Signal

MET SIGNIFICANCES SEPARATION POWER

Performance comparison: ROC Curve

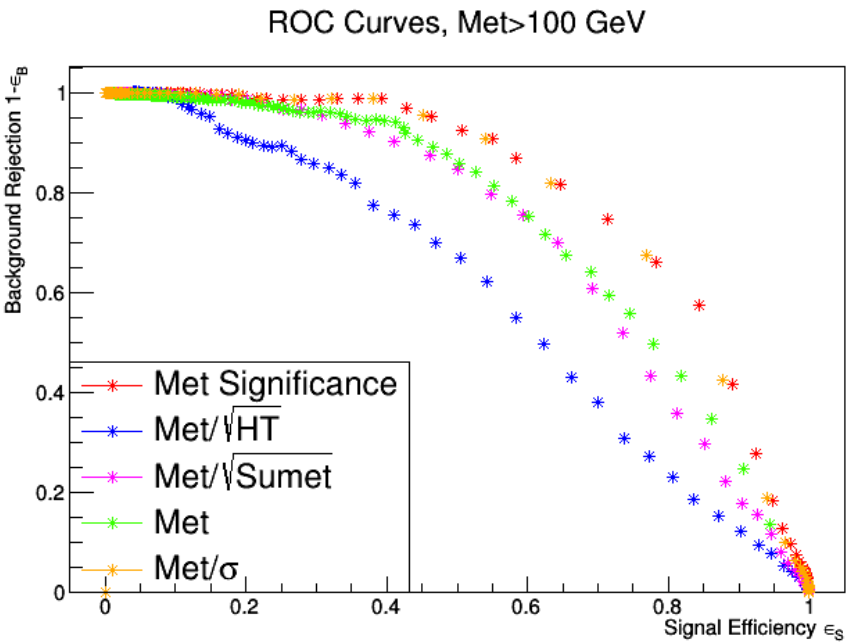


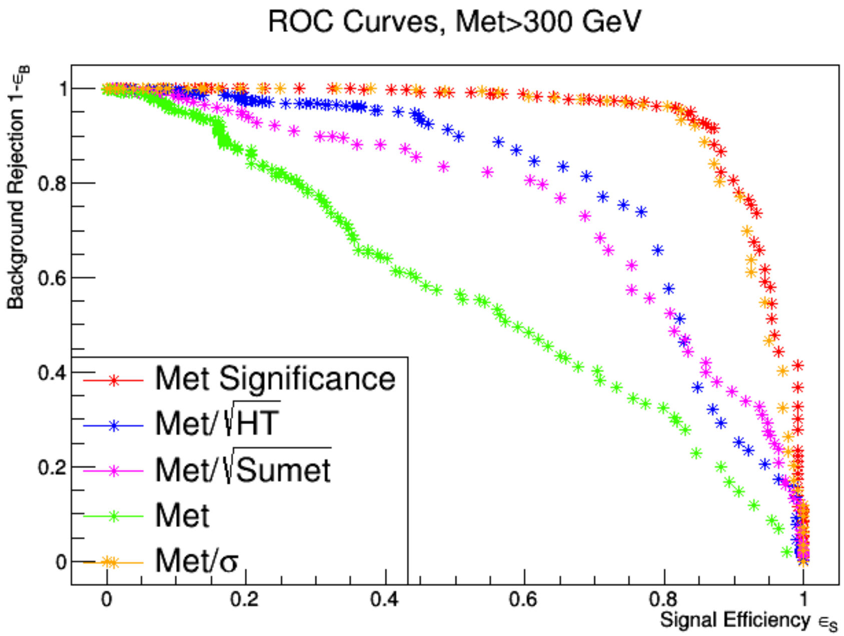
Object based Met Significance

Object based Met Significance without correlation factor

MET SIGNIFICANCES SEPARATION POWER

Performance comparison ROC Curve → Cutting on MET





Better performance for high MET events

SUMMARY

- Mono-H(H->bb) is a very atracting search in the LHC
  - Clean channel
  - No-ISR process
  - Dominant dacay mode of Higgs
- Mediator mass exclusion in Mono-H simplified Models
  - Z' mediator mass excluded < 900 GeV
  - Pseudoescalar A mass < 500 GeV excluded for mZ'>1 TeV in the Z'-2HDM
- A MVA approach for the Scalar mediator in the merge regime has a potential improvement of ~+40 % in signal significance
- Met Significance for a Z->mumu topology shows a similar performance compared w.r.t other Met Significances definitions.
- The performance of a object based MET significance is very encouraging mainly for high MET.

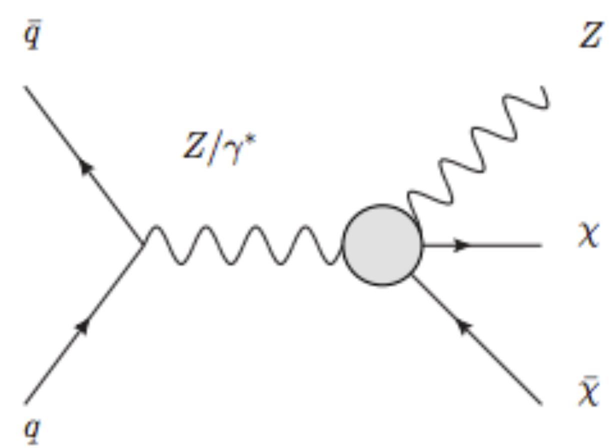
THANKS

BACKUP SLIDES

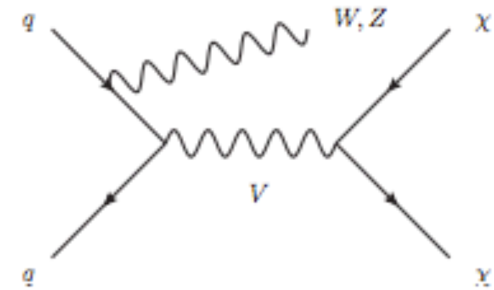
Mono-X channels

- X = Vector Boson or Higgs are covered.
  - Highlights are on 13 TeV analysis (all ATLAS at this moment)
- Mono-photon
  - Low background.
- Mono-Z or mono-W
  - Z can be emitted from mediator in t-channel.
  - Hadronic decay mode → larger cross section
  - Leptonic decay mode → cleaner signature.
- Mono-H
  - No ISR (Initial State Radiation) Higgs.
  - H can be emitted from mediator in s-channel.
  - H -> bb decay mode → larger cross section.
  - H -> γγ decay mode → clean signature.
- VVxx (HHxx) contact interaction is unique.
- Other mono-X:
  - Mono-jet: Andreas Korn's talk.
  - Mono-heavy quark(s): Alberto Zucchetta's talk.

Mono-photon	8 TeV	ATLAS: arXiv:1411.1559[hep-ex] CMS: arXiv:1410.8812
	13 TeV	ATLAS: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/EXOT-2015-05/
Mono-Z/W (hadr)	8TeV	ATLAS: arXiv:1309.4017[hep-ex] CMS: CMS PAS EXO-12-055
	13 TeV	ATLAS: ATLAS-CONF-2015-080
Mono-Z(ll)	8 TeV	ATLAS: arXiv: 1404.0051[hep-ex] CMS: arXiv: 1511.09375
Mono-W(lv)	8 TeV	ATLAS: arXiv:1407.7495[hep-ex] CMS: arXiv:1408.2745[hep-ex]
Mono-H(bb)	8 TeV	ATLAS: arXiv:1510.0621[hep-ex]
	13 TeV	ATLAS: ATLAS-CONF-2016-019
Mono-H(gamgam)	8 TeV	ATLAS: arXiv:1506.01081[hep-ex]
	13 TeV	ATLAS: ATLAS-CONF-2016-011



Simplified model parameters in Run2



- Based on the Dark Matter Forum recommendation (arXiv:1507.00966 [hep-ex]).
- Dark matter: Dirac particles.
- Mediator: Vector, Axialvector, Scalar or Pseudoscalar particles.
- Mediator width: minimal width = sum of contributions from DM and quarks lighter than a half of the mediator mass.
- S-channel coupling constants:
  - Coupling to DM: $g_{DM} = 1.0$
  - Coupling to SM: universal to all quarks.
    - Vector and Axialvector: $g_{SM} = 0.25$ (larger values are constrained by dijet searches, also to keep the mediator width narrow).
    - Scalar and Pseudoscalar: $g_{SM} = 1.0$
- T-channel couplings: $g_{DM} = g_{SM} = 0.1 - 7$

13 TeV

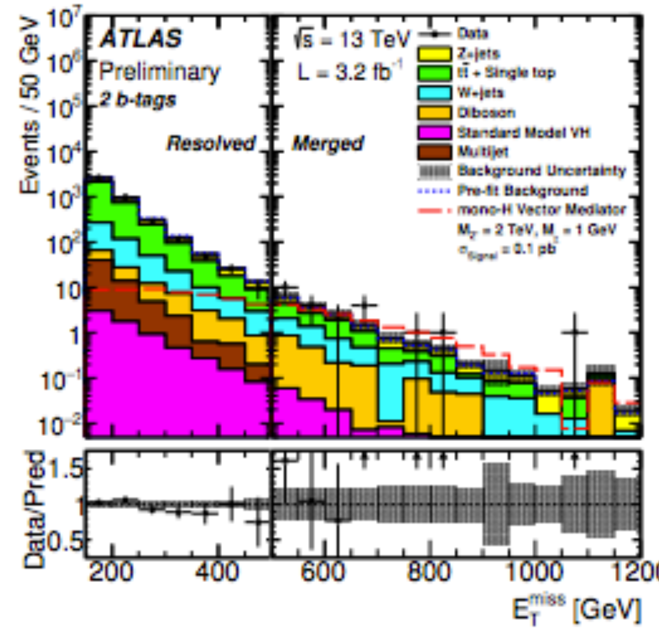
Mono-H(bb) Results

NEW

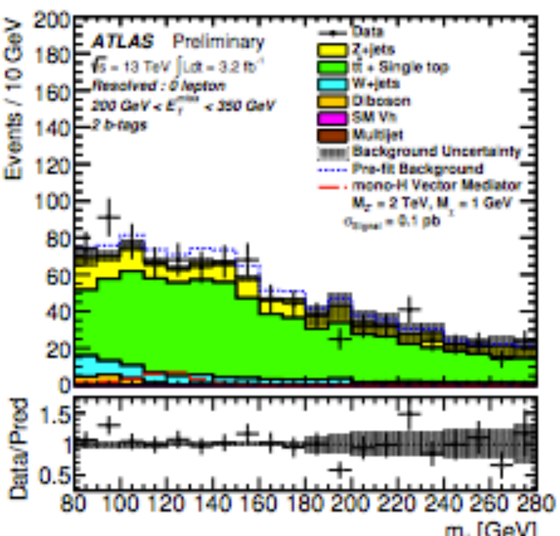
- ATLAS 13 TeV, 3.2 fb^{-1}; 4 signal regions:

E_T^{miss} (GeV)	Resolved 150–200	Resolved 200–350	Resolved 350–500	Merged >500
$Z + jets$	258.52 ± 26.81	171.24 ± 13.13	14.63 ± 1.21	3.80 ± 0.44
$W + jets$	94.78 ± 27.79	70.14 ± 21.67	7.51 ± 2.42	2.48 ± 0.71
$t\bar{t}$ & Single top	1444.38 ± 44.39	656.02 ± 24.51	30.76 ± 1.41	4.83 ± 0.88
Multijet	21.38 ± 9.96	10.89 ± 5.08	0.58 ± 0.27	–
Diboson	17.84 ± 1.62	18.73 ± 0.98	2.53 ± 0.22	1.20 ± 0.12
$SMVh$	2.77 ± 1.30	2.78 ± 1.40	0.46 ± 0.23	0.15 ± 0.08
Tot. Bkg.	1839.68 ± 33.12	929.80 ± 19.63	56.47 ± 2.08	12.47 ± 1.27
Data	1830	942	56	20
Exp. Signal	80.15 ± 7.95	244.53 ± 17.76	160.58 ± 11.56	149.28 ± 33.67

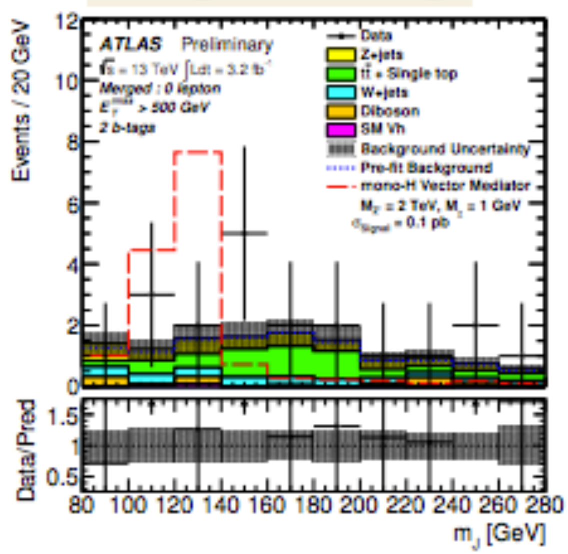
MET



Resolved, 2 b-tag



Merged, 2 b-tag



- Stat error 20.5%
- Systematic error 10.3%
- Main background:
  - Z+jets, W+jets, ttbar
  - Estimated from 1and 2 lepton CR

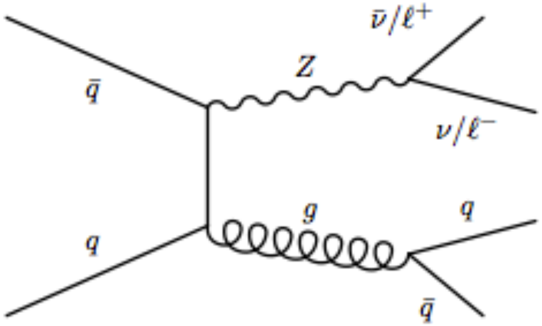
Backup: Mono-H(bb) Event Selection

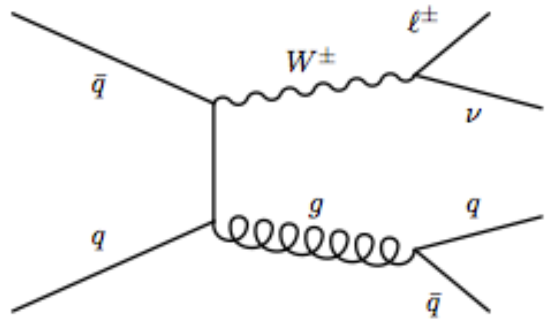
13 TeV

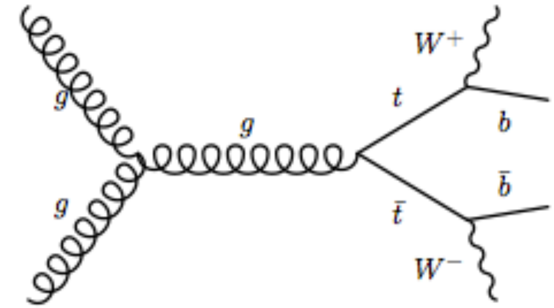
- 13 TeV, 3.2 fb^{-1}, MET trigger.
- MET > 150 GeV, and track based MET: p-MET > 30 GeV
- Lepton veto (no isolated electron or muon with pT>7GeV)
- H candidate:
  - Two small-R jets (j_1 and j_2) in resolved region (MET<500GeV)
    - Leading jet p_T > 45 GeV
  - One large-R jet in merged region (MET>500GeV)
  - 1 or 2 b-tagged jet(s).
- Resolved region : cuts to suppress multi-jets background
  - min[Δφ(MET,jets)] > 20 deg: No jets near MET.
  - Δφ(MET,p-MET) < 90 deg: MET and track MET align.
  - Δφ(MET,Higgs) > 120 deg: MET and H go back-to-back.
  - $\Delta\varphi(j_1,j_2)$ < 140 deg: Two jets are not back-to-back.

Dominant backgrounds are:

$Z \to \nu\nu$+jets: | $W \to \ell\nu$+jets: | $t\bar{t}$+production:







- Two lepton region to estimate $Z \to \ell\ell$+jets: which is kinematically similar to the desired estimate of $Z \to \nu\nu$+jets
- Single muon region to estimate W+jets and $t\bar{t}$: number of b-tags naturally separates them
- Try to have the control regions close to the signal region

→ Apply similar cuts

- Less dominant backgrounds: single-top, diboson, SM VH $b\bar{b}$, multijet

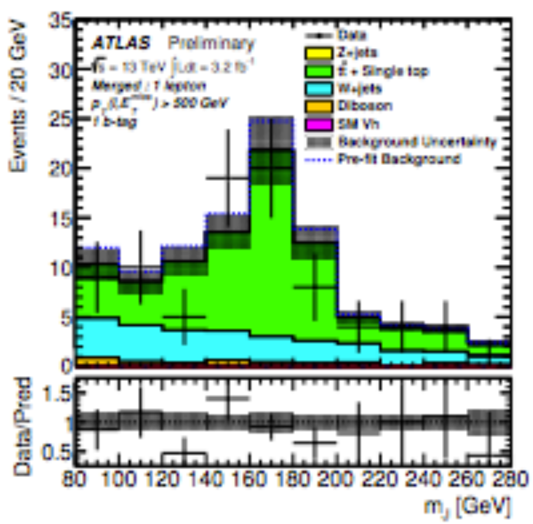
# of leptons \ # of b-tagged jets	0	1	2
0	CR, Z+jets	SR	SR
1	CR, W+jets	CR, W+jets, $t\bar{t}$	CR, W+jets, $t\bar{t}$
2	CR, Z+jets	CR, Z+jets	CR, Z+jets

13 TeV

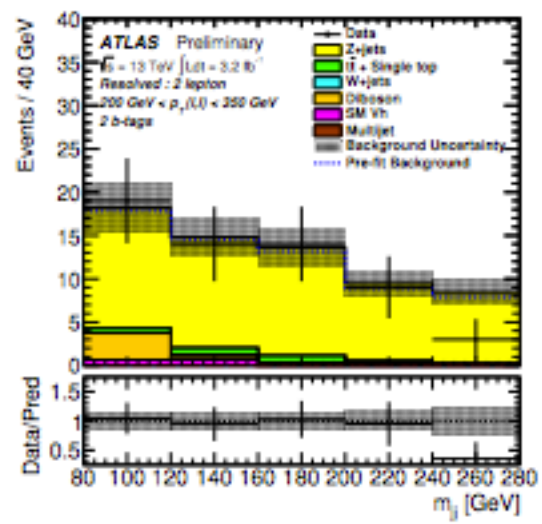
Backup: Mono-H(bb) Background Estimation

- W+jets and ttbar with missing lepton: One-muon control region.
- Z+jets with missing leptons: Two-lepton control region
- Multi-jet background (resolved region): data-driven method. Derived from multi-jet dominant region.

One-muon CR, Merged, 1 b-tag



Two-lepton CR, Resolved, 2 b-tag

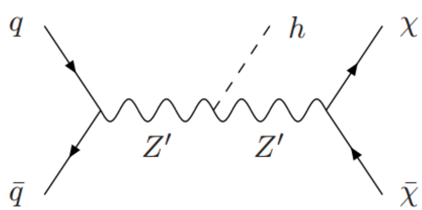


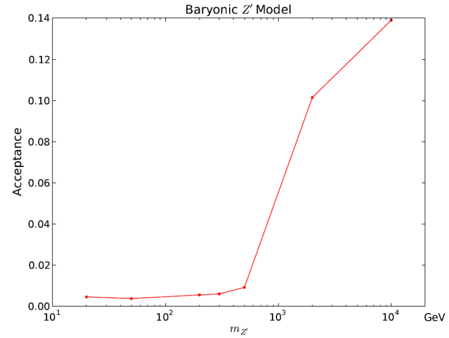
SIMPLIFIED BENCHMARK MODELS

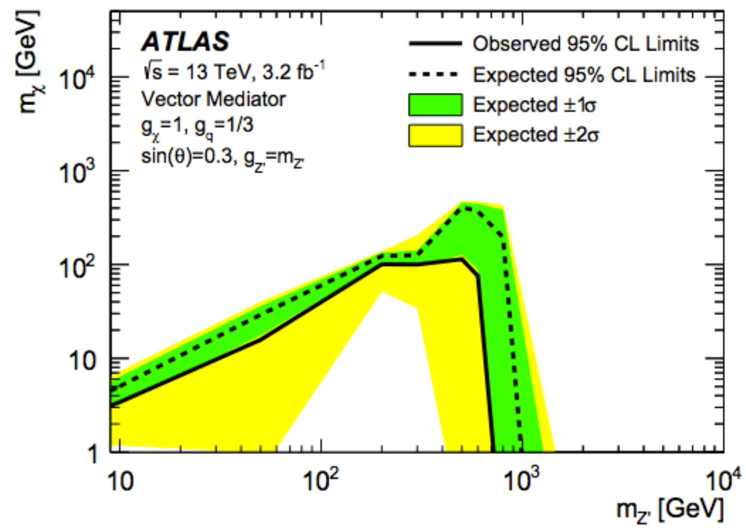
- **Baryonic Z'**

$$\mathscr{L} \supset g_q \bar{q}\gamma^\mu q Z'_\mu + \begin{cases} i g_\chi \chi^\dagger \overleftrightarrow{\partial^\mu} \chi Z'_\mu + g_\chi^2 |\chi|^2 Z'_\mu Z'^\mu & \text{scalar} \\ g_\chi \bar{\chi}\gamma^\mu \chi Z'_\mu & \text{fermion} \end{cases}$$

Acceptance mX=1. Gev for Met > 500. GeV



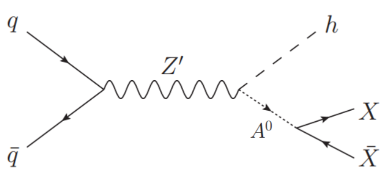


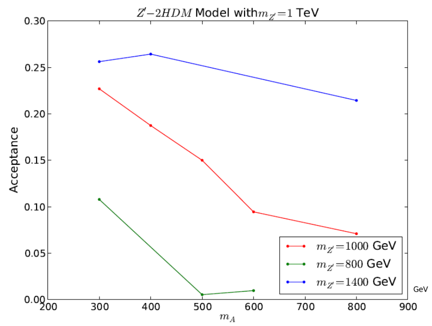


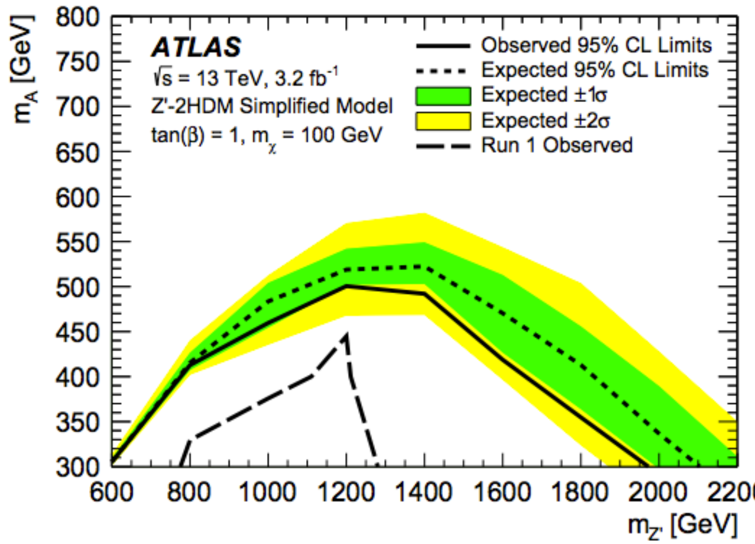
- **Z'-2HDM**

$$-\mathcal{L} \supset y_u Q \tilde{\Phi}_u \bar{u} + y_d Q \Phi_d \bar{d} + y_e L \Phi_d \bar{e} + \text{h.c.}$$

Acceptance mX=1. Gev for Met > 500. GeV







SIMPLIFIED BENCHMARK MODELS

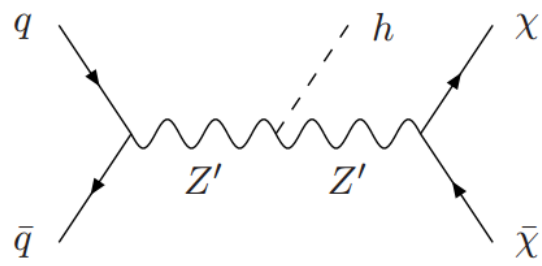
- **Baryonic Z'**

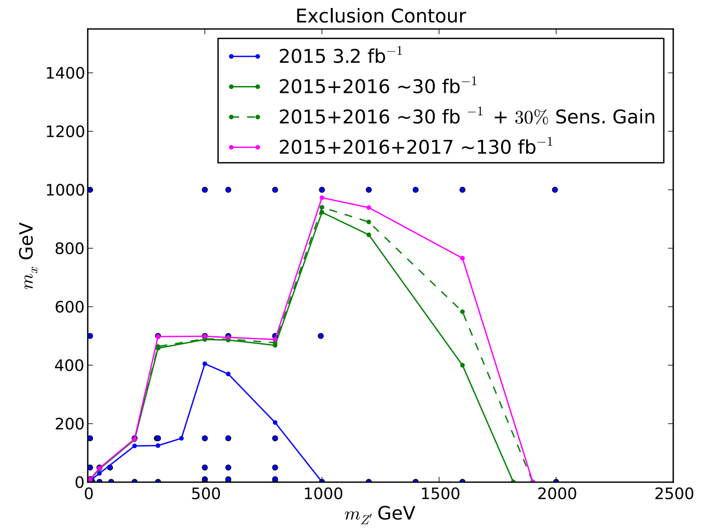
$$g_q \bar{q}\gamma^\mu q Z'_\mu + g_\chi \bar{\chi}\gamma^\mu \chi Z'_\mu$$

After SSB

$$-g_{hZ'Z'} h Z'_\mu Z'^\mu$$

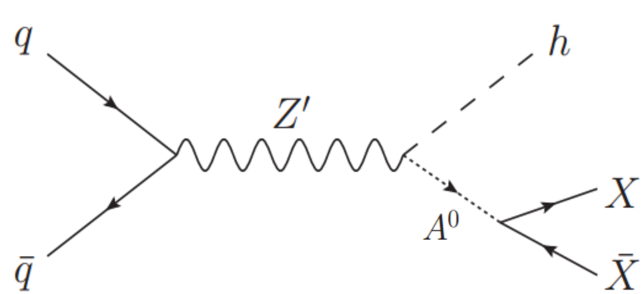
$$g_{hZ'Z'} = \frac{m_{Z'}^2 \sin\theta}{v_B}$$

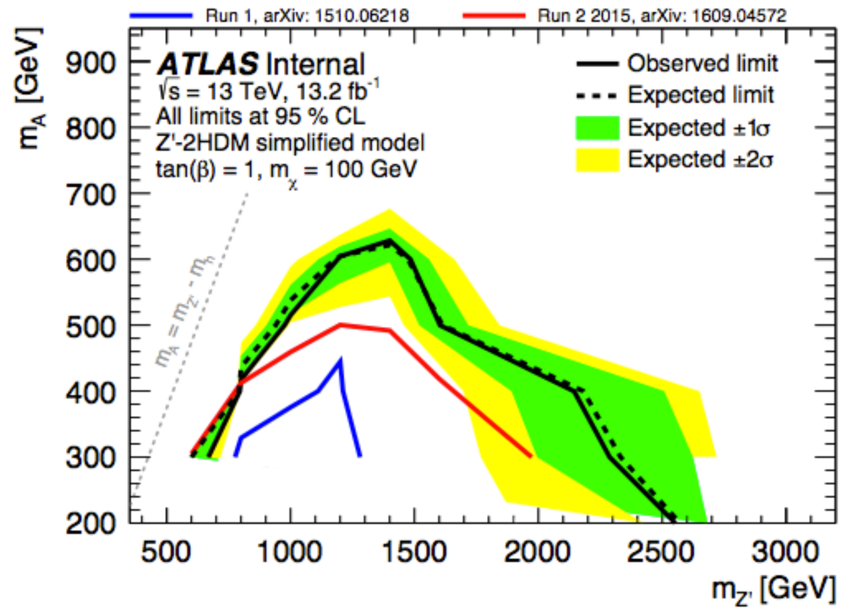




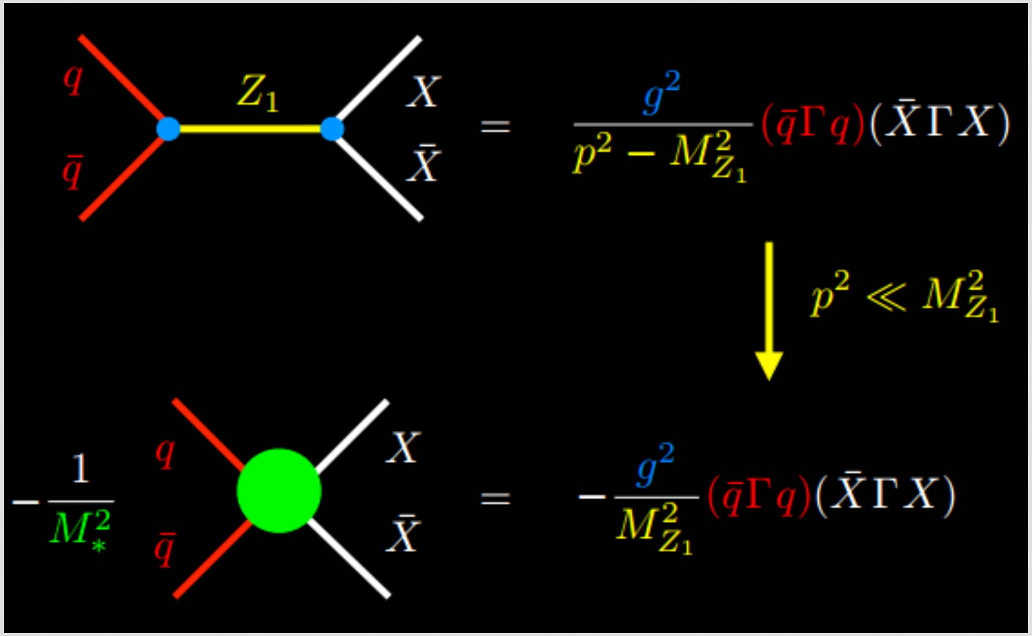
- **Z'-2HDM**

$$-\mathcal{L} \supset y_u Q\tilde{\Phi}_u \bar{u} + y_d Q\Phi_d \bar{d} + y_e L\Phi_d \bar{e} + \text{h.c.}$$





EFT



q
$\bar{q}$
Z_1
X
$\bar{X}$
$= \frac{g^2}{p^2 - M_{Z_1}^2}(\bar{q}\Gamma q)(\bar{X}\Gamma X)$
$p^2 \ll M_{Z_1}^2$
$-\frac{1}{M_*^2}$
q
$\bar{q}$
X
$\bar{X}$
$= -\frac{g^2}{M_{Z_1}^2}(\bar{q}\Gamma q)(\bar{X}\Gamma X)$

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.

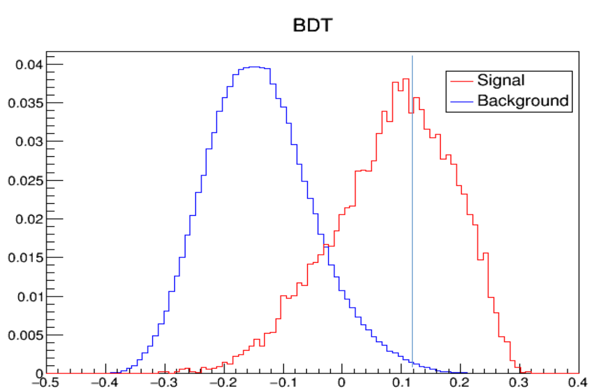
BDT ON THE SCALAR MEDIATOR S.M

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.

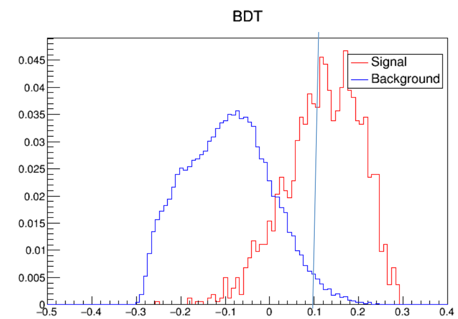
- **Classifier Output Distribution**



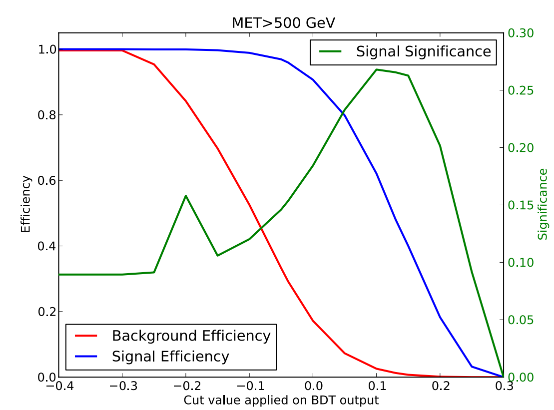
MET > 500 GeV

We are interested in the high MET category .

- **Classifier Output Distribution**

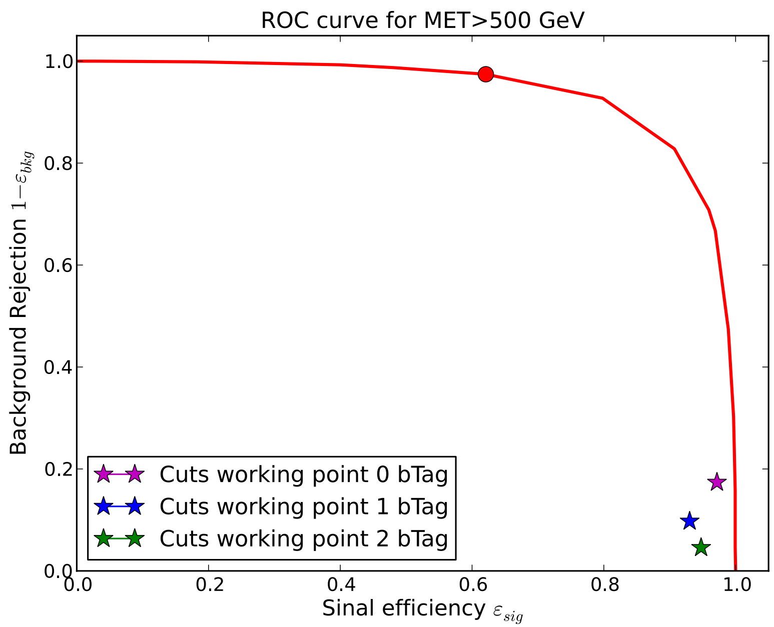


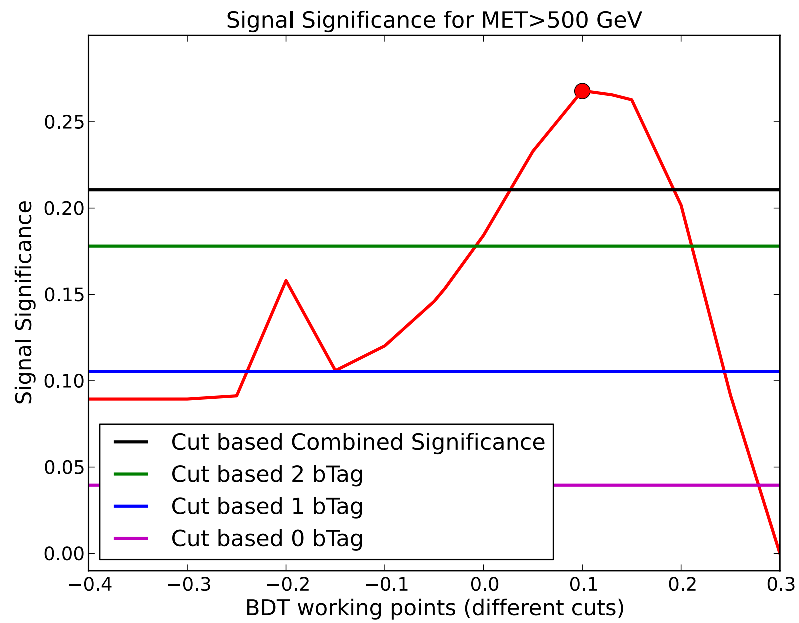
- **Classifier Cut efficiencies**



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

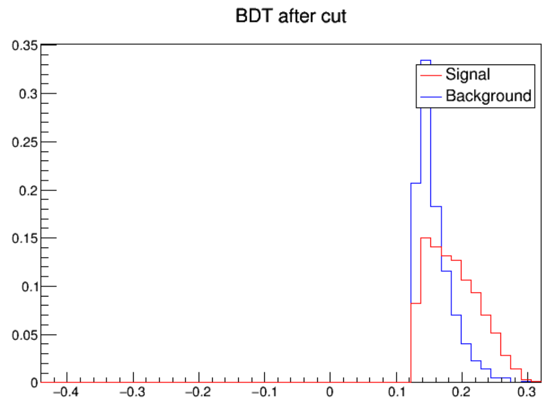
$$S = \sqrt{S_{0-btag}^2 + S_{1-btag}^2 + S_{2-btag}^2}$$

Significance improvement is +28%

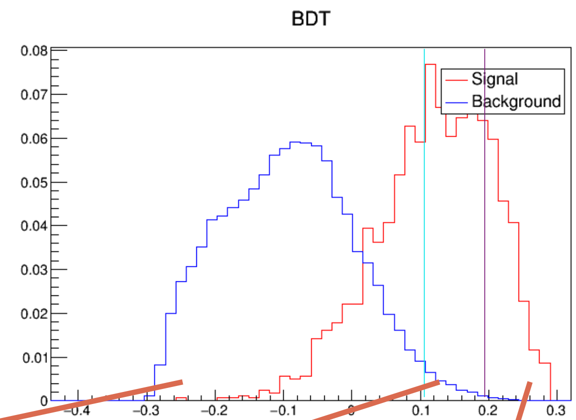
B-TAG JETS VS BDT CATEGORIES

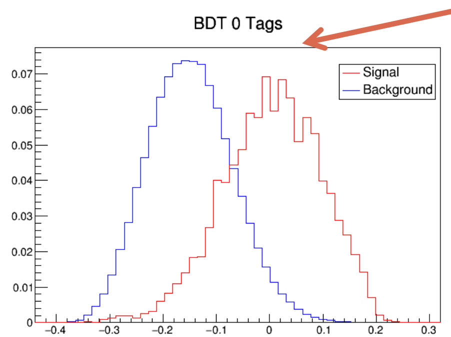
The number of b-tag jets is a variable considered in the BDT training.

Then, instead of having a single sharp cut on BDT

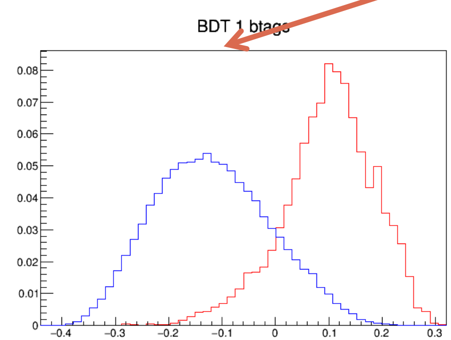


We can define BDT categories that can be related to 0, 1 and 2 btags.

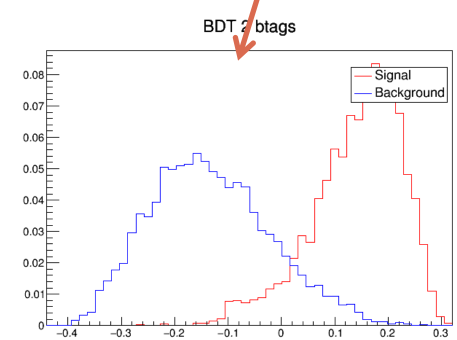




✧ Low BDT category



✧ Medium BDT category



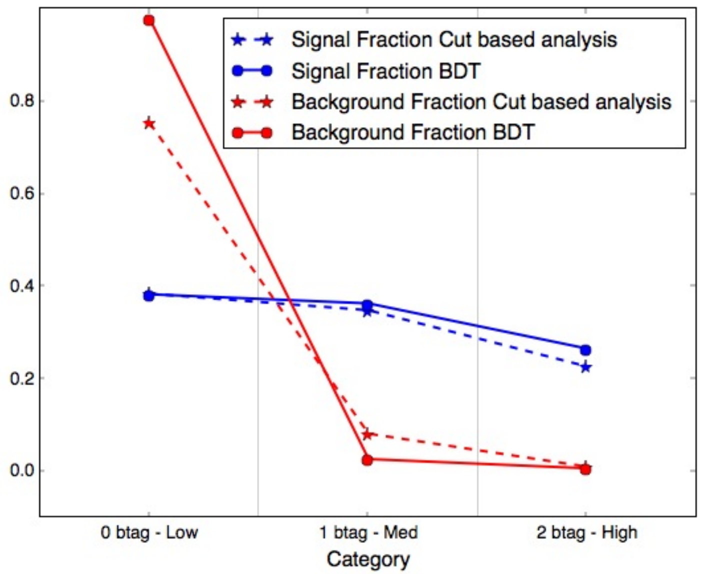
✧ High BDT category

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

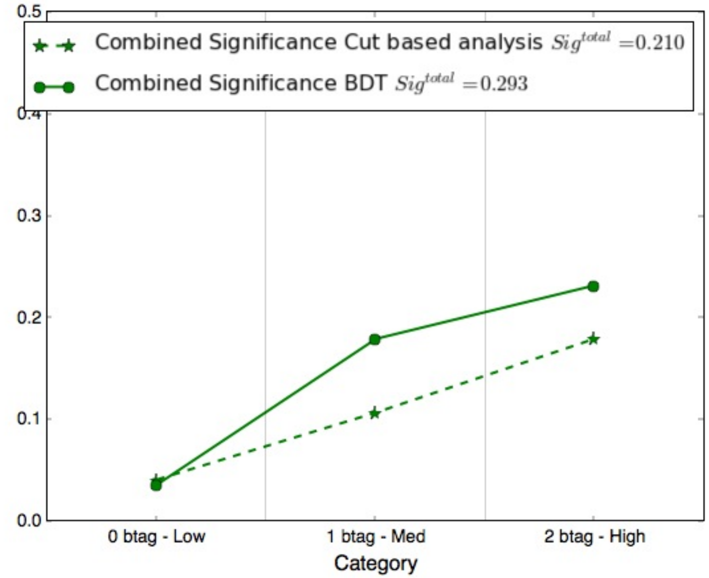
COMPARISON BETWEEN BDT CLASSIFIER AND CUTS BASED ANALYSIS FOR MET >500 GEV

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.

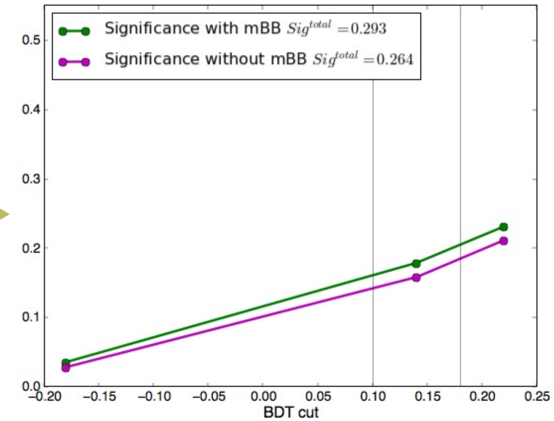


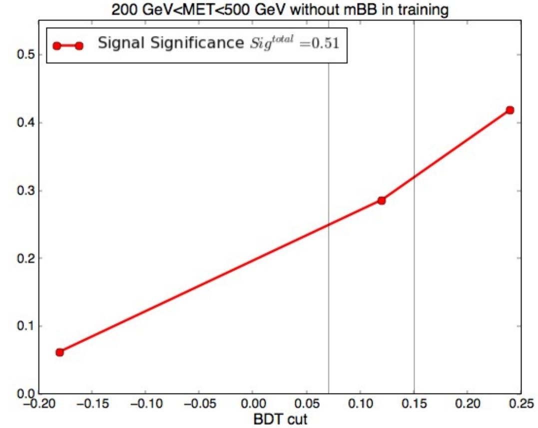
	Cuts Based Analysis			BDT			
	0 btag	1btag	2 btag	Low	Medium	High	
Signal Events Expected*	0.105569	0.0956614	0.0618382	0.105227	0.0997611	0.0726001	
Background Events Expected *	7.04099	0.730347	0.058999	9.1536	0.215417	0.0267651	
Purity	0.01499348813	0.130980753	1.04812285	0.01149569568	0.4631069043	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.

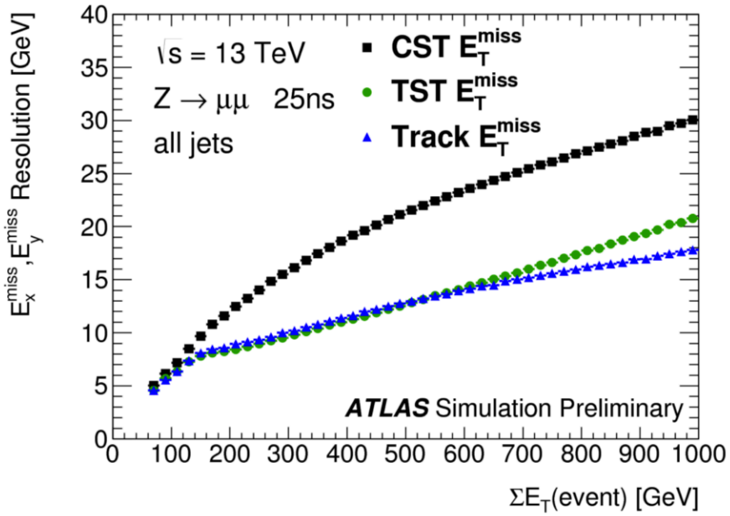
- A high value of S is an indication that the observed Met is not well explained by resolution smearing alone, suggestions that the event may contain unseen objects such as neutrinos or more exotic weakly interacting particles.

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.

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

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 $\overrightarrow{\epsilon_i} = \overrightarrow{E_{T_i}} - \overrightarrow{e_{T_i}}$ has a gaussian probability density function

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

…the likelihood for two objects is given by:

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

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

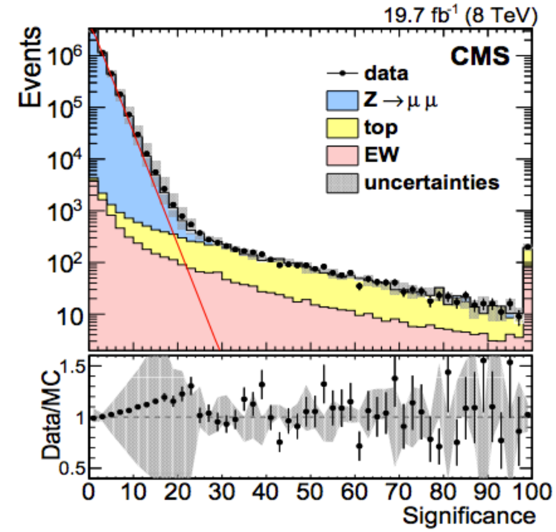
OBJECT BASED MET SIGNIFICANCE

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

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

- For each object contributing to the MET, the covariance matrix is rotated into a coordinate system with the x axis parallel to the total Met axis.

$$U_i = \begin{pmatrix} \sigma^2_{E_{T_i}} & 0 \\ 0 & E^2_{T_i}\sigma^2_{\Phi_i} \end{pmatrix} \longrightarrow \mathbf{V} = \sum_i V_i = \begin{pmatrix} \sigma^2_{\parallel} & \sigma^2_{\parallel\perp} \\ \sigma^2_{\parallel\perp} & \sigma^2_{\perp} \end{pmatrix}$$

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

Met Significance Simplification

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

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

- Where the correlation coefficient is:

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

NOTE: If $\rho^2 = 1$

In this case the definition becomes:

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

For $\rho^2 \geq 0.9$

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.

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

- Assuming gaussian uncertainties distributions:

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

- 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:

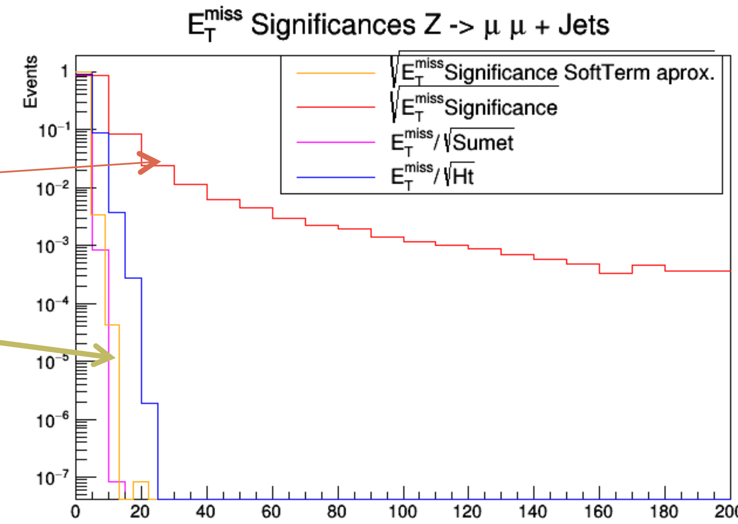
$$S \sim \frac{\left(\sum_i \overrightarrow{E_{T_i}}\right)^2}{\sigma_{\parallel}^2}$$

For $\rho^2 \geq 0.9$

Object based Met Significance

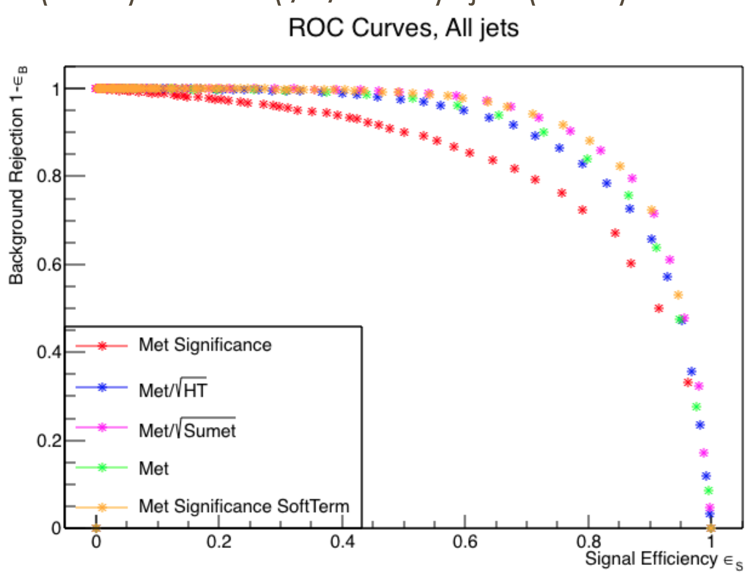
- **Considering** **jet resolutions**
- **Z-> μ μ + jets (no genuine Met)**
  - The Met error is under estimated (due to the missing if the other objects resolutions)so we obtain a over estimated distribution.
  - First approximation for the resolution of the Soft terms

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



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

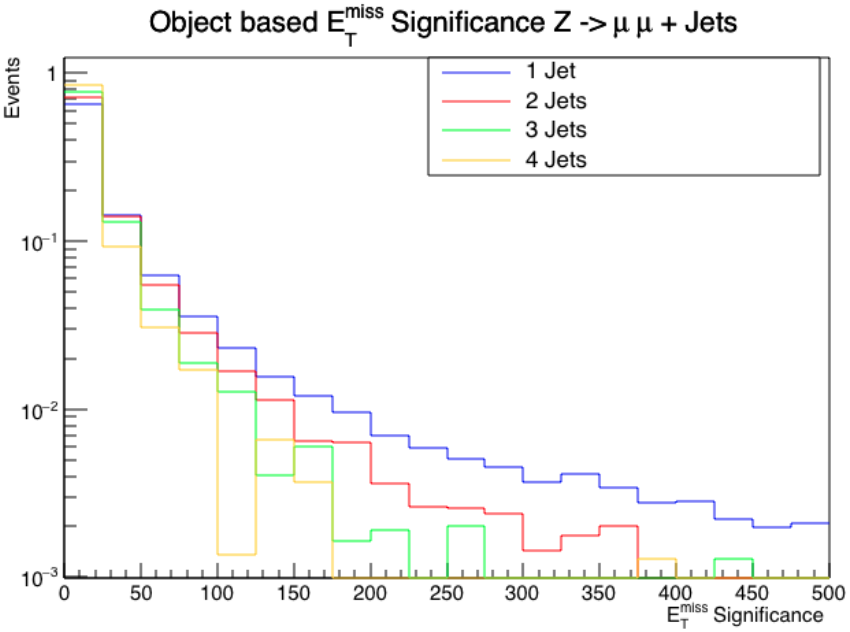
FIRST MET SIGNIFICANCE IMPLEMENTATION

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

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

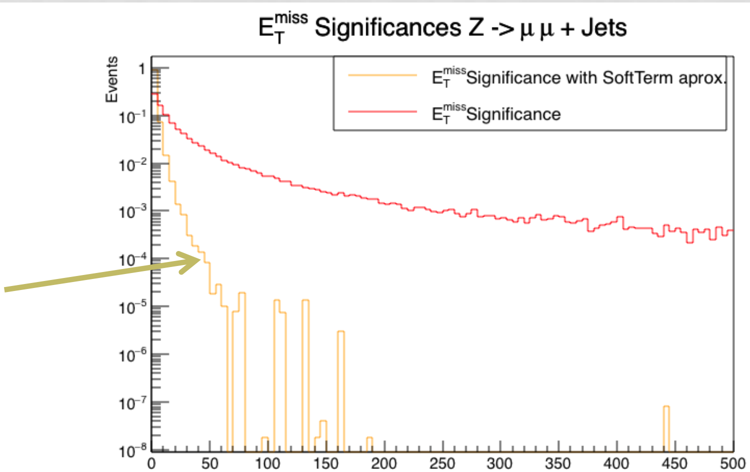
The Met error in the denominator of the Significance is under estimated (due to the missing if the other objects resolutions)so we obtain a over estimated significance distribution.

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



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)

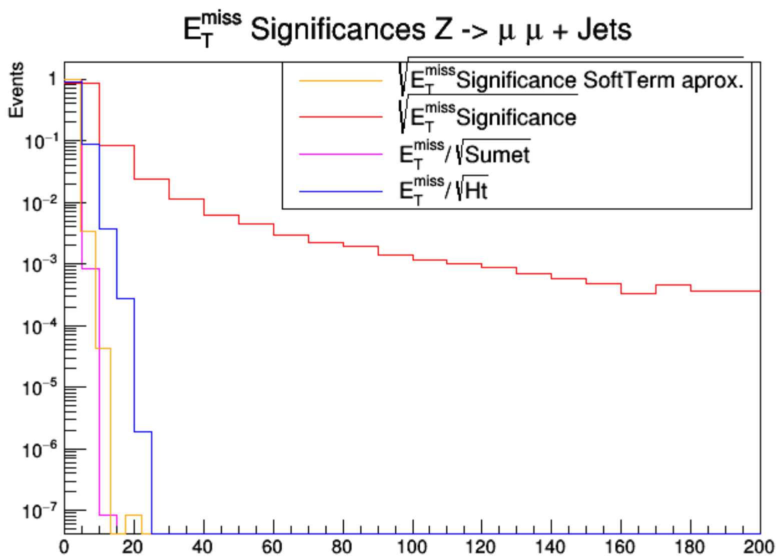


MET SIGNIFICANCES SEPARATION POWER

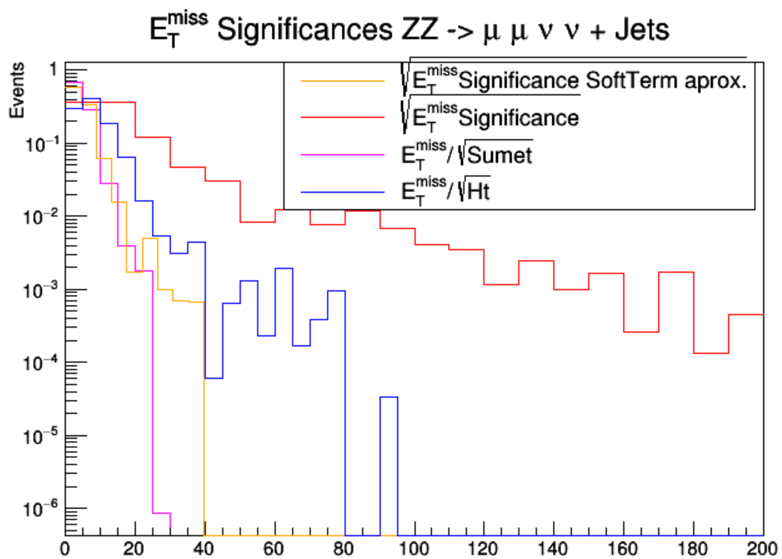
- Comparison of the separation power between
  - Z($\mu\mu$)+jet (BKG) and ZZ($\mu\mu\nu\nu$)+jet (SGN)
  - Cutting on MET - cutting on MET significance - cutting on MET/sqrt(HT) or cutting on MET/ sqrt(Sumet)

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



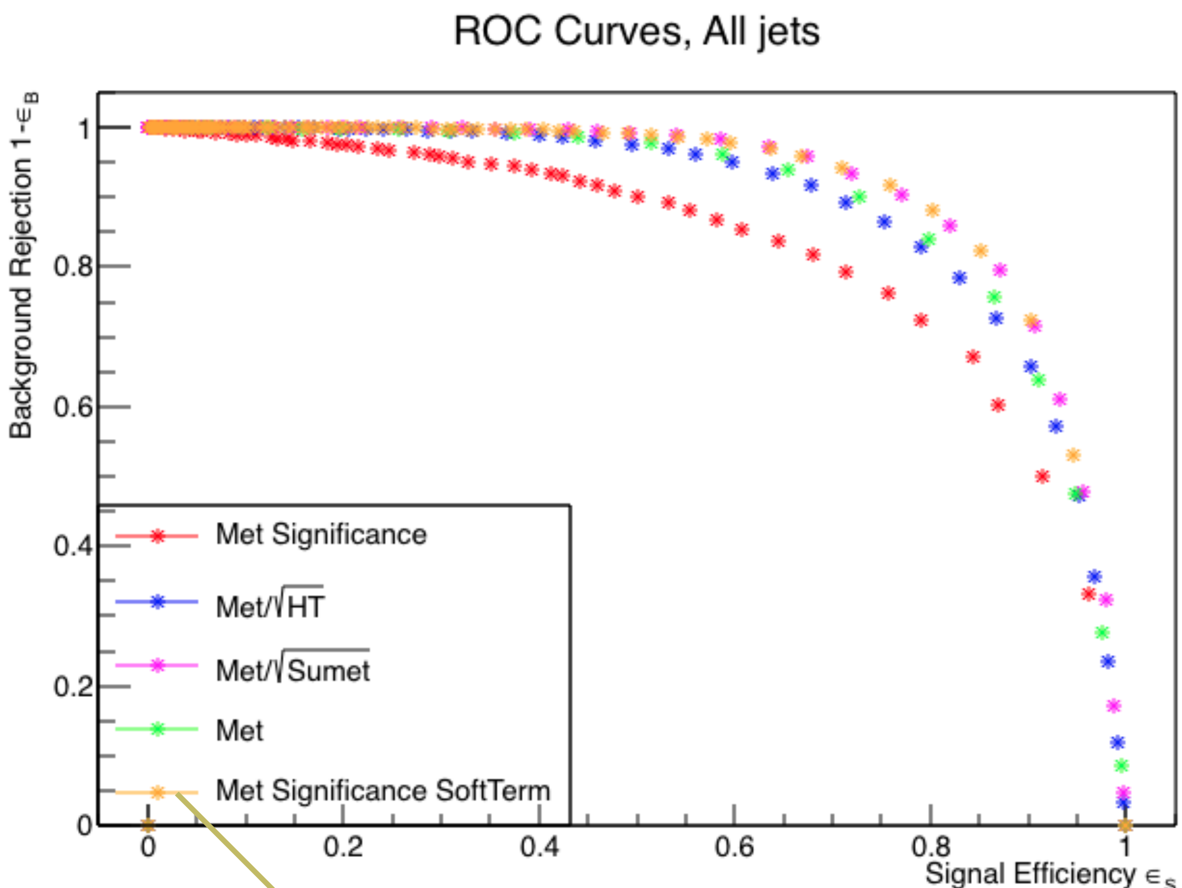
Background



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)

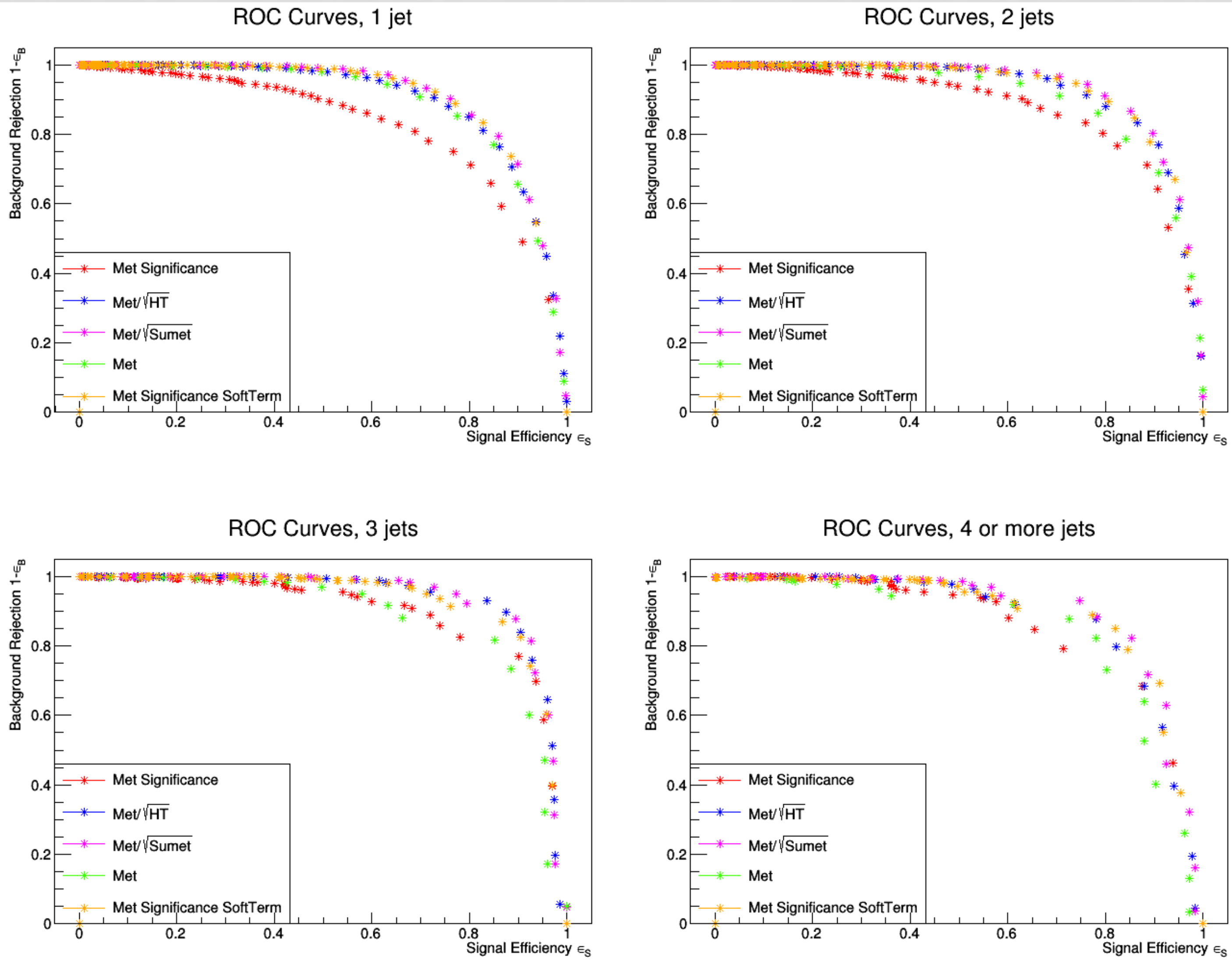
We need to properly implement the Soft Term resolutions!

…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

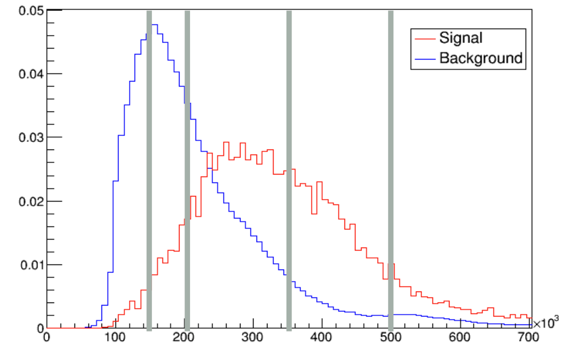
- First test: Build the V in an easy environment, start selecting Z->mumu+1 jet. Why?
- Assuming contribution from resolution from jets bigger than the one from muon, we can neglect muons resolution for the moment.
- Selecting events with only 1 jet should be easier to calculate V
- The only terms participating 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.

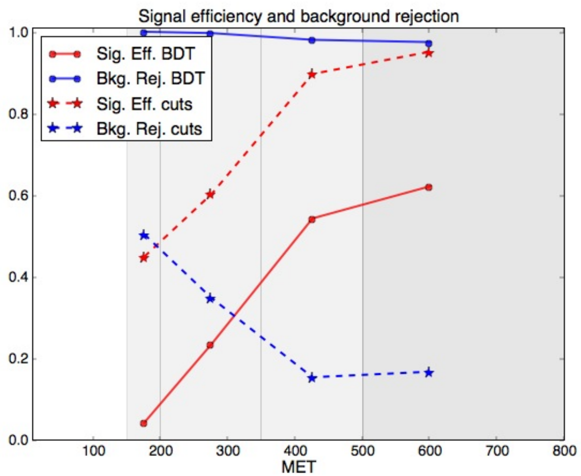
- Once the first MET significance is calculated, planning to compare the separation power between Z($\mu\mu$)+jet (BKG) and ZZ($\mu\mu\nu\nu$)+jet (SGN)
  - cutting on MET - cutting on MET significance - cutting on MET/sqrt(HT) or cutting on MET/sqrt(Sumet)
- Depending on these first tests, we will start 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.

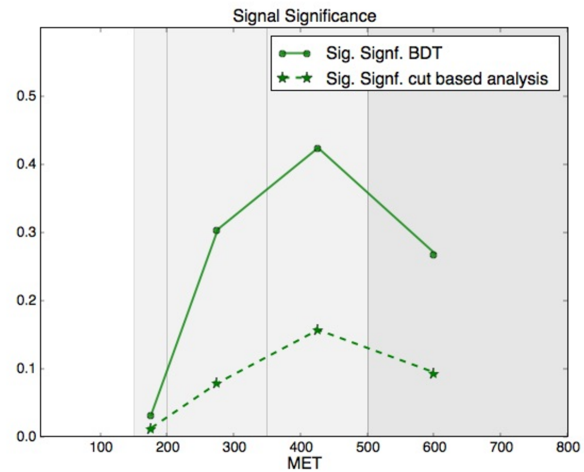
NAÏVE COMPARISON BETWEEN BDT CLASSIFIER AND CUTS BASED ANALYSIS

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**

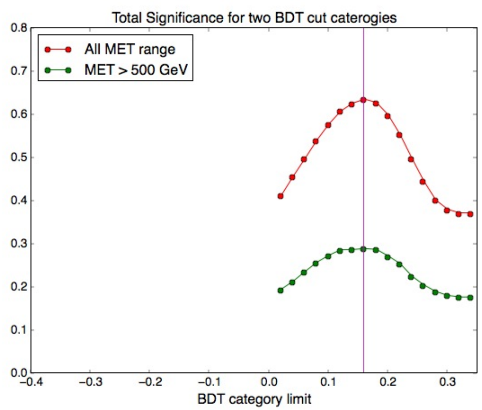
* From now on, the number of signal and background events are associated to the TestTree which correspond to half of the statistics

DEFINING BDT CATEGORIES

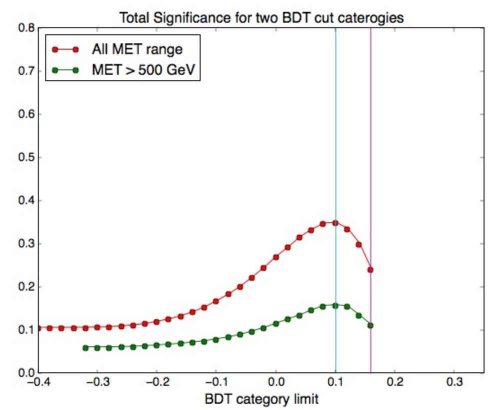
Low – medium – high BDT categories are defined by maximizing the combined significance at first order.

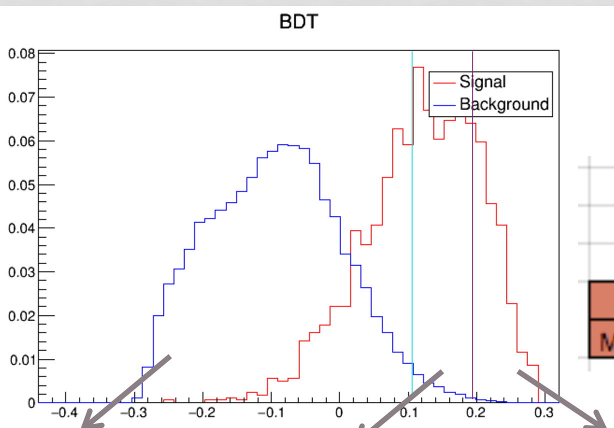
$$S = \sqrt{S^2_{LOW} + S^2_{MEDIUM} + S^2_{HIGH}}$$

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

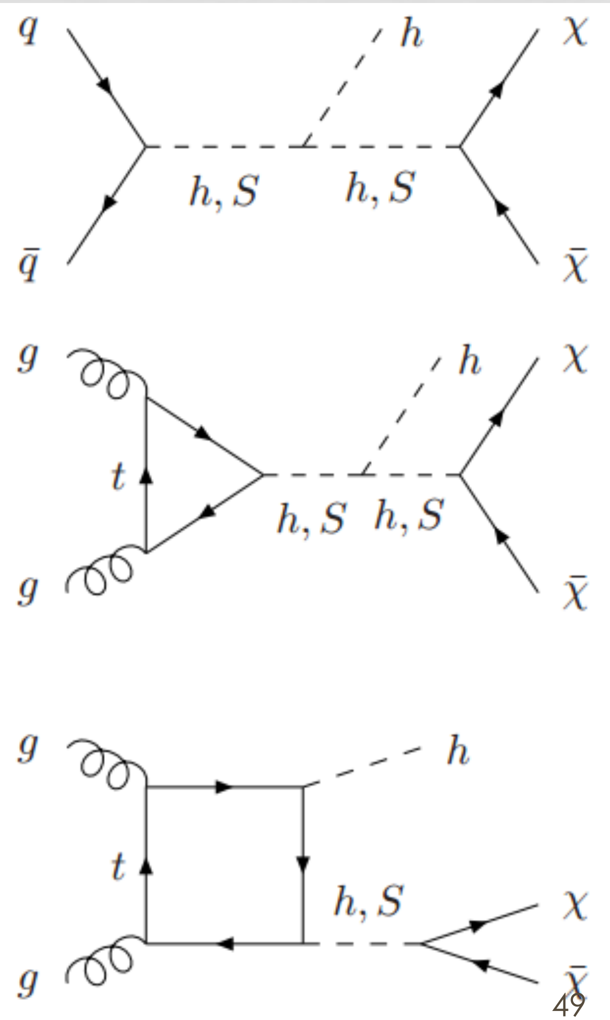
$$V \supset a|H|^2 S + b|H|^2 S^2 + \lambda_h |H|^4$$

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

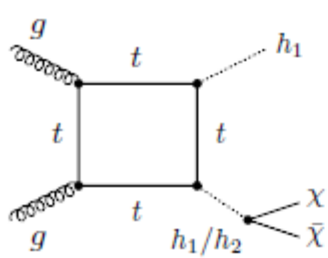
$$\mathcal{L}_{\text{fermion},H} \supset -\mu_s s^3 - \lambda_s s^4 - y_\chi \bar{\chi}\chi s - \mu_p s|H|^2 - \lambda_p s^2|H|^2$$

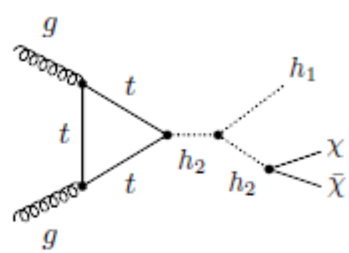
$$\langle s \rangle = 0$$

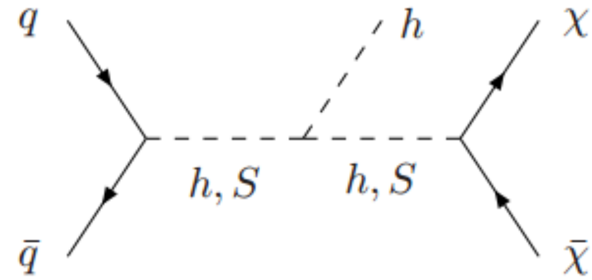
$$\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} h \\ s \end{pmatrix} \qquad \tan(2\theta) = \frac{2v\mu_p}{m_s^2 + \lambda_p v^2 - m_h^2}$$

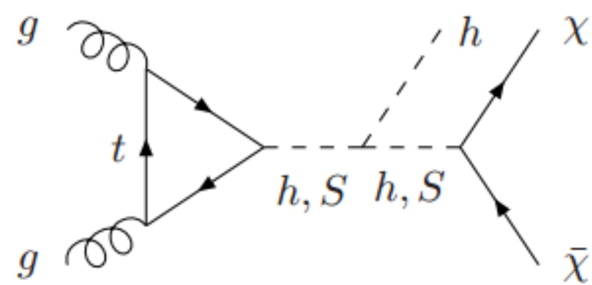
$\theta \to 0$ the dark sector is decoupled from the SM $\qquad \sin\theta \lesssim 0.4$

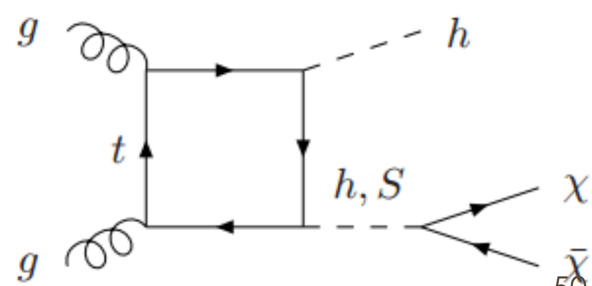
$m_{h_1} \simeq m_h$ and $m_{h_2} \simeq (m_s^2 + \lambda_p v^2)^{1/2}$











BOOSTED DESITION TREE

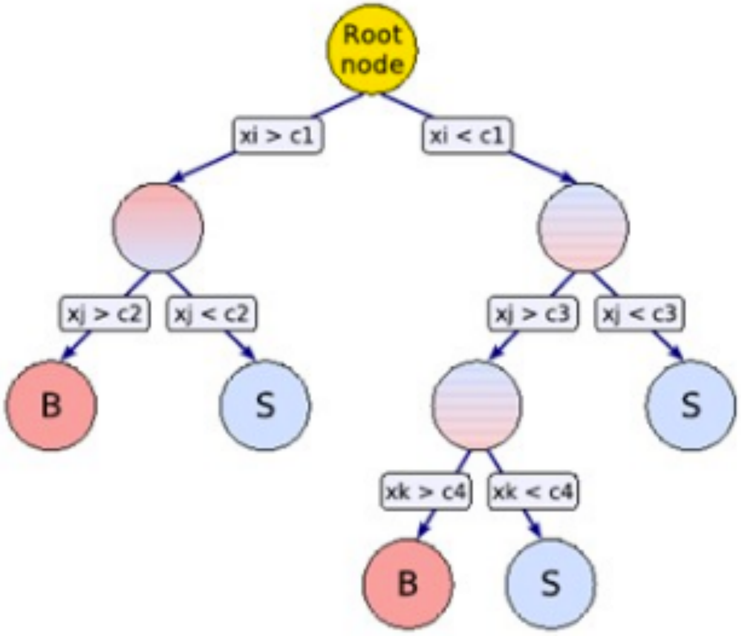
- Decision Tree

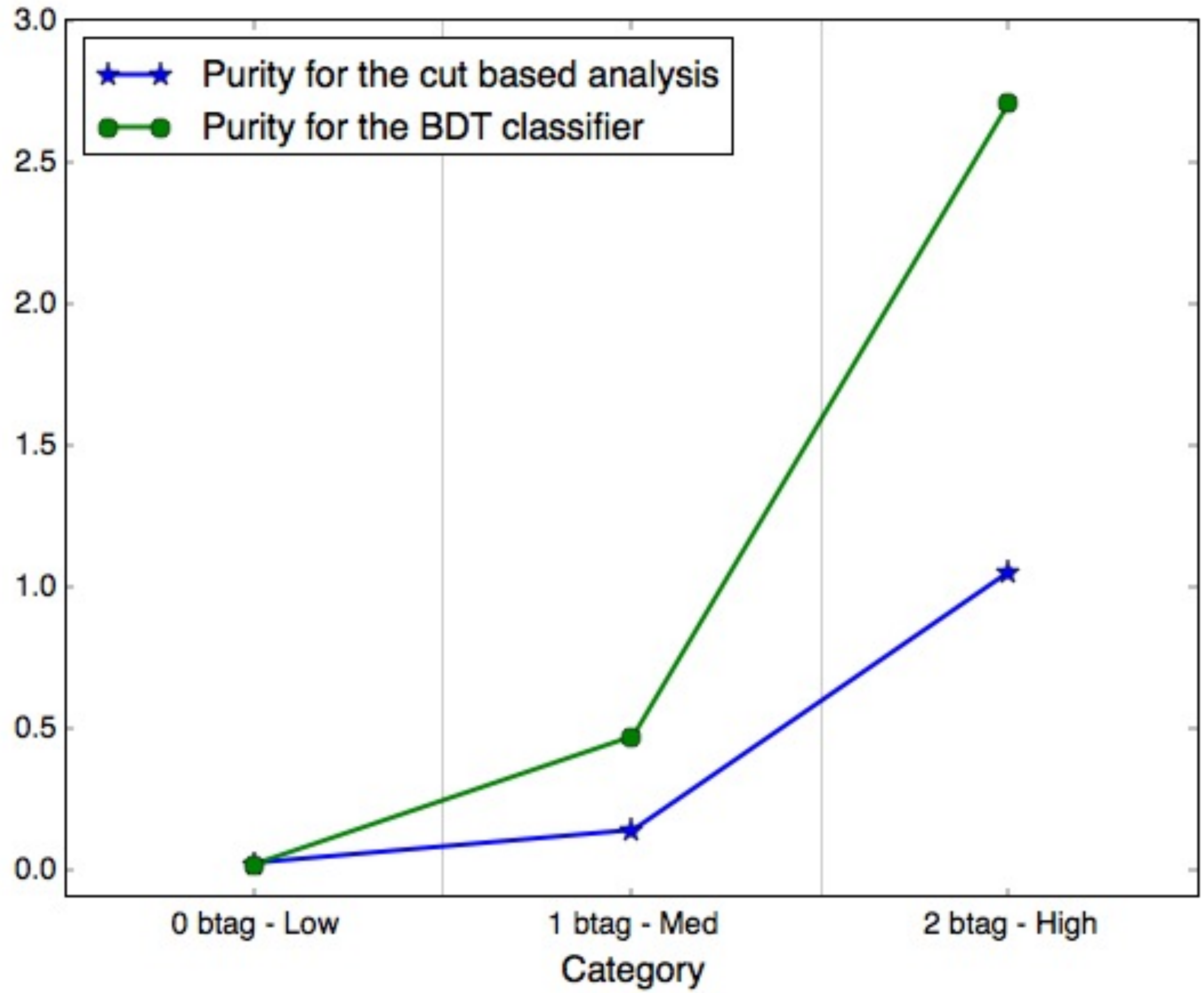
- Consecutive set of questions (nodes)
  - Only two possible answers per question
  - Each question depends on the formerly given answers
- Final veredict (leaf) is reached after a given maximum number of nodes

- Random forest

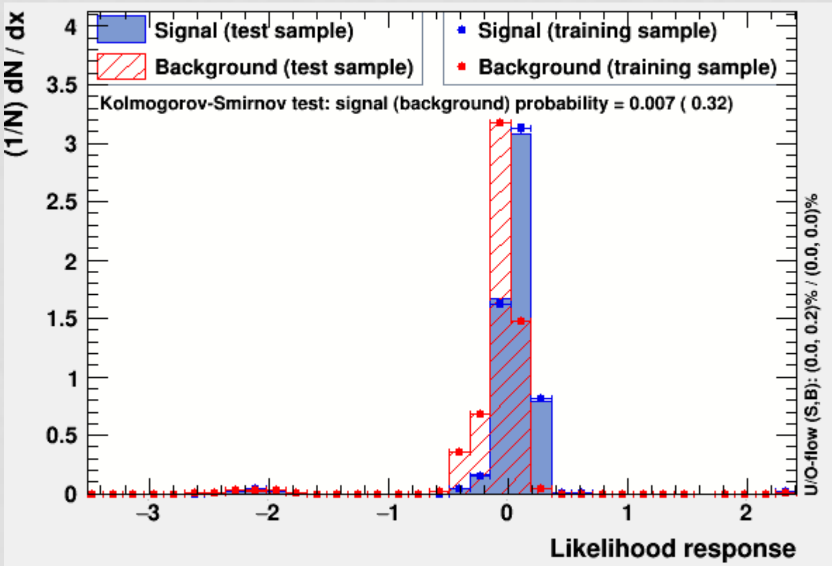
- Random Forests is an ensemble method that combines different trees
- Final output is determined by the majority vote of all trees
- Boosting
  - Misclassified events are weighted higher so that future learners concentrate on these

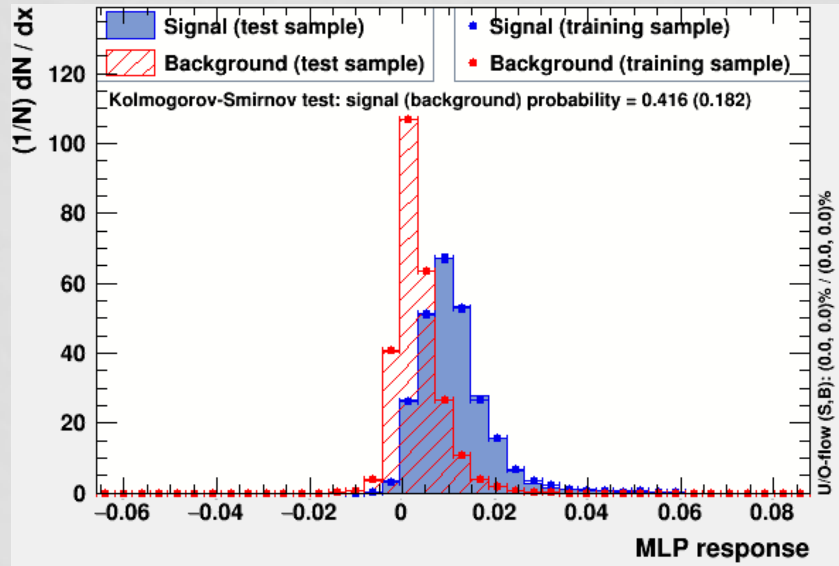
- Training a decision tree



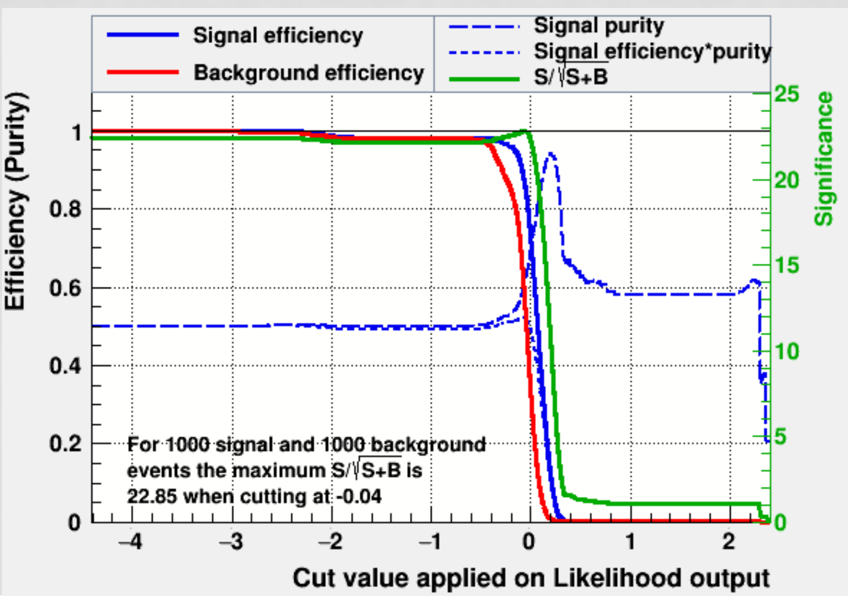

3.0
2.5
2.0
1.5
1.0
0.5
0.0
Purity for the cut based analysis
Purity for the BDT classifier
0 btag - Low
1 btag - Med
2 btag - High
Category

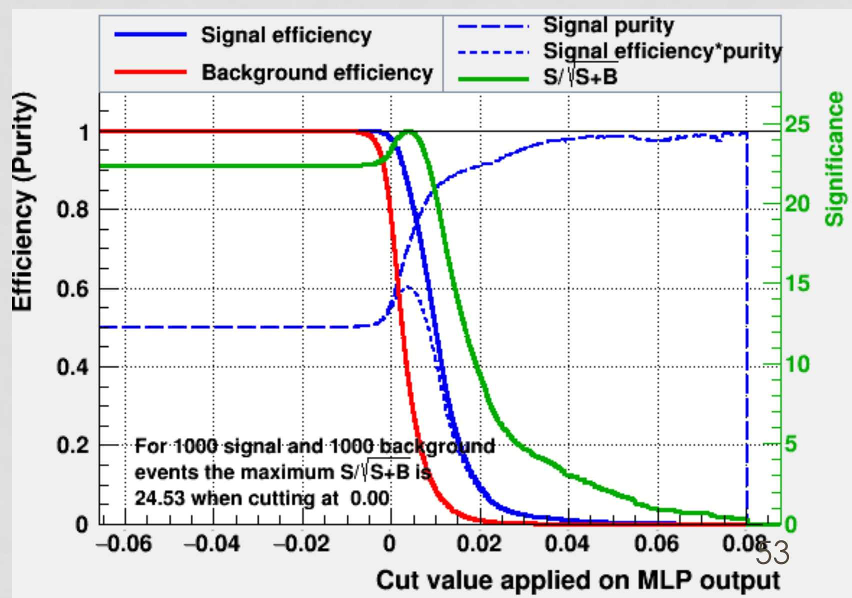
Classifier Output Distributions





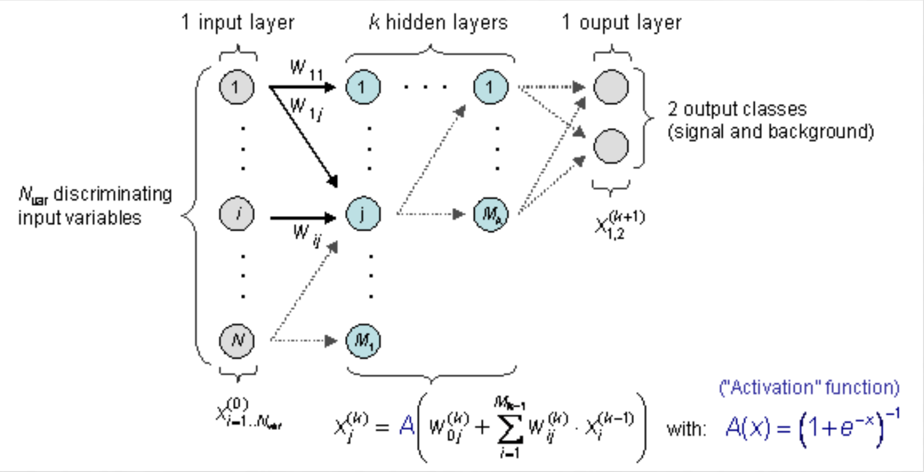
Classifier Cut Efficiencies





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

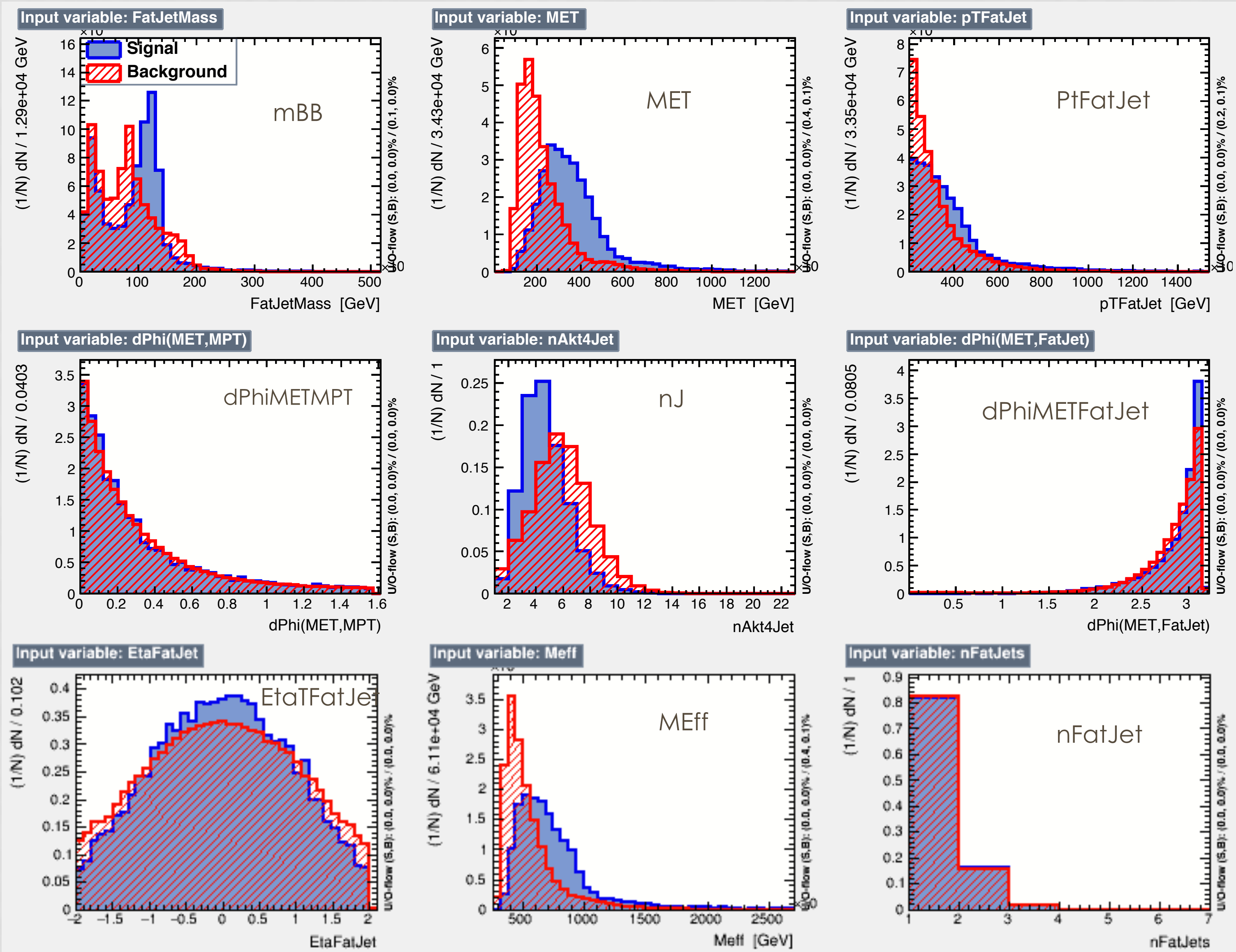
- Signal Sample

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

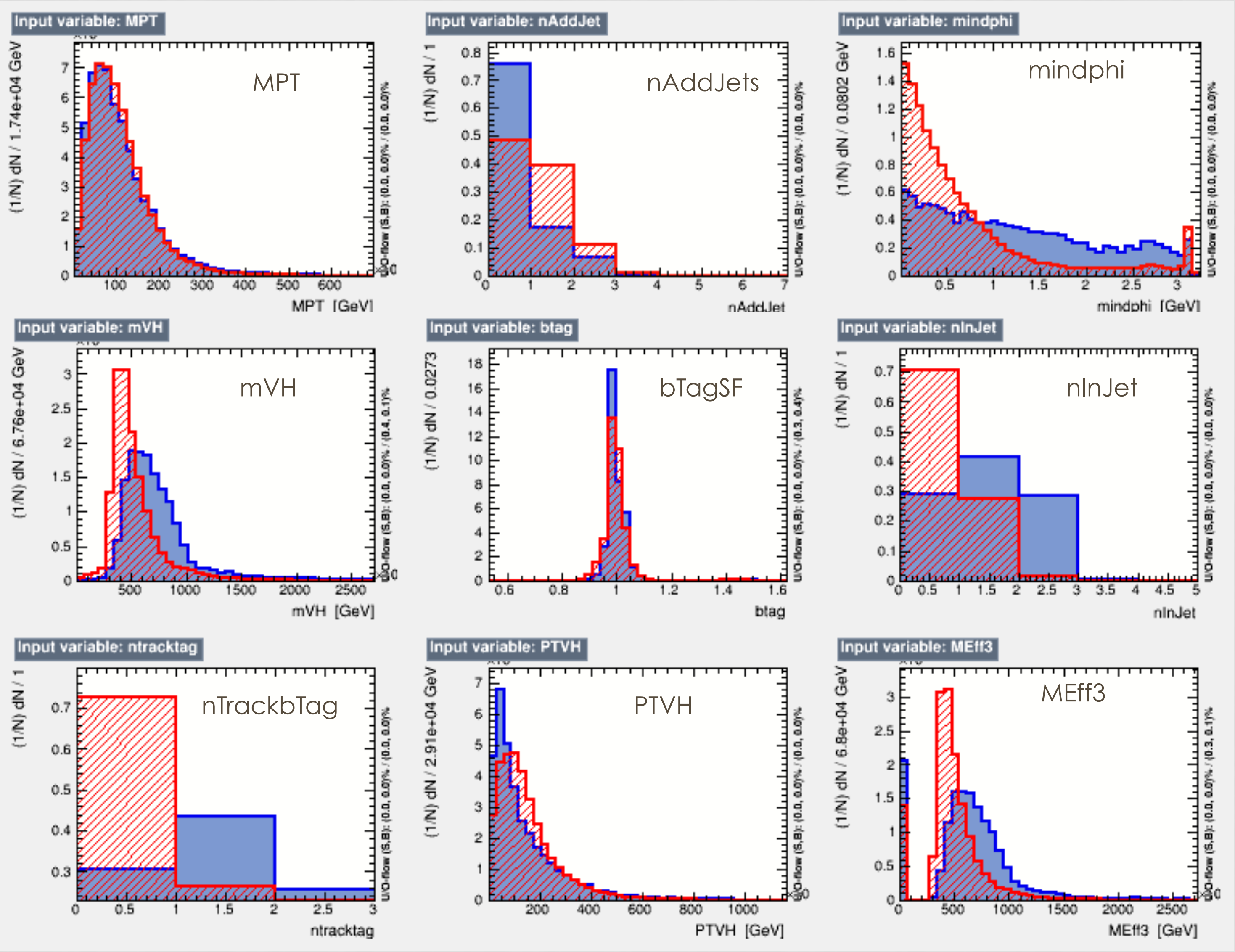
- Samples taken from CxAOD_00-14-05
- Normalizaded to 1 fb-1 of integrated luminosity

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

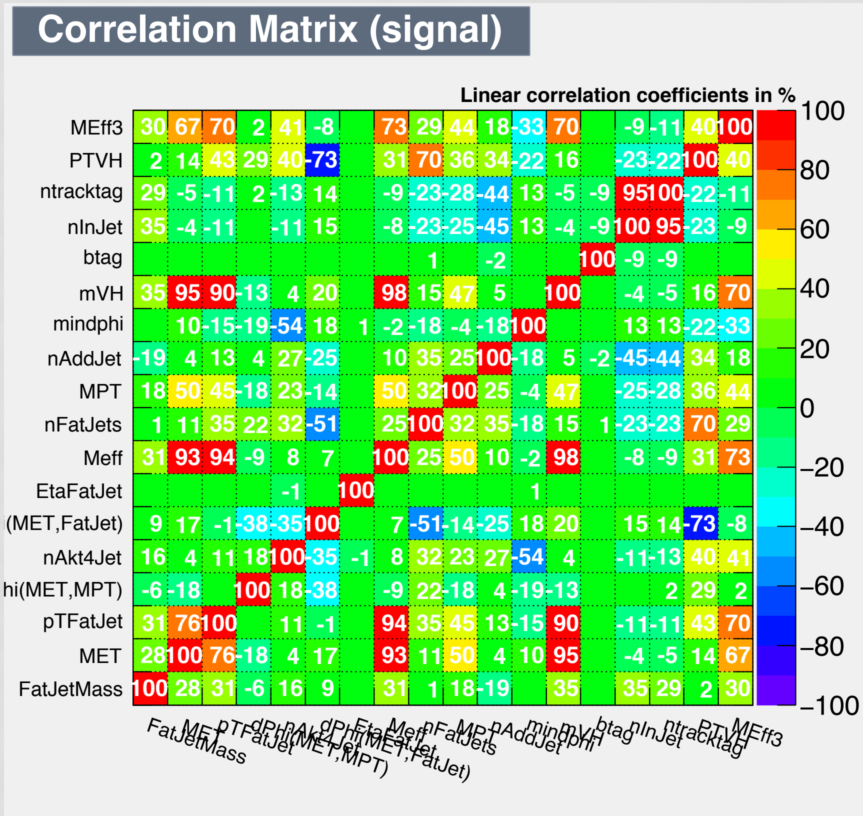
VARIABLES DISTRIBUTIONS

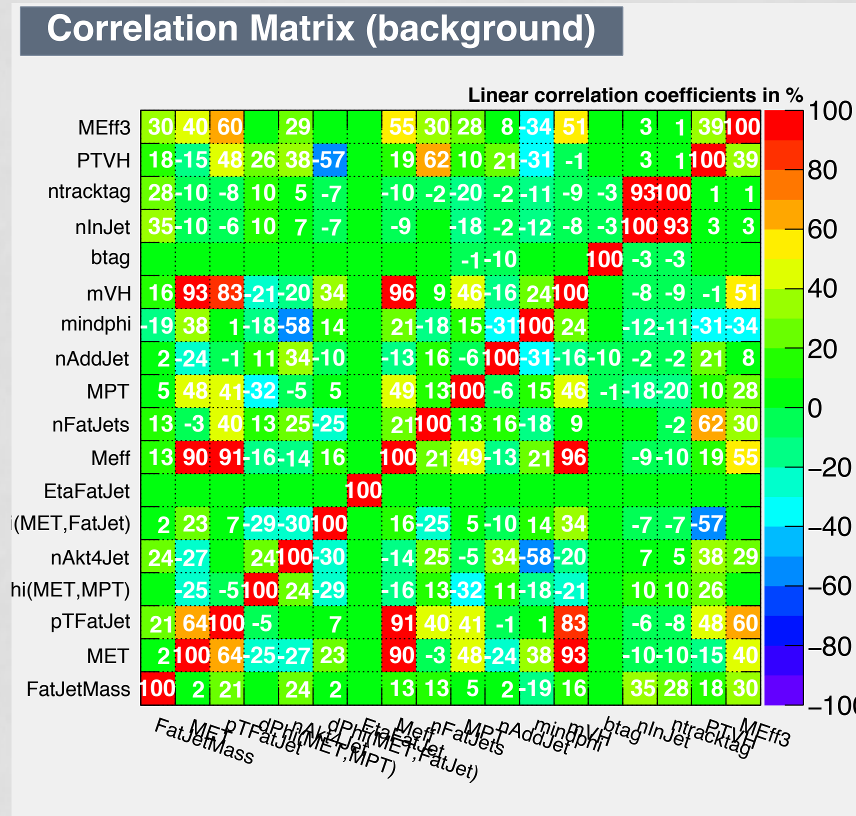


VARIABLES DISTRIBUTIONS



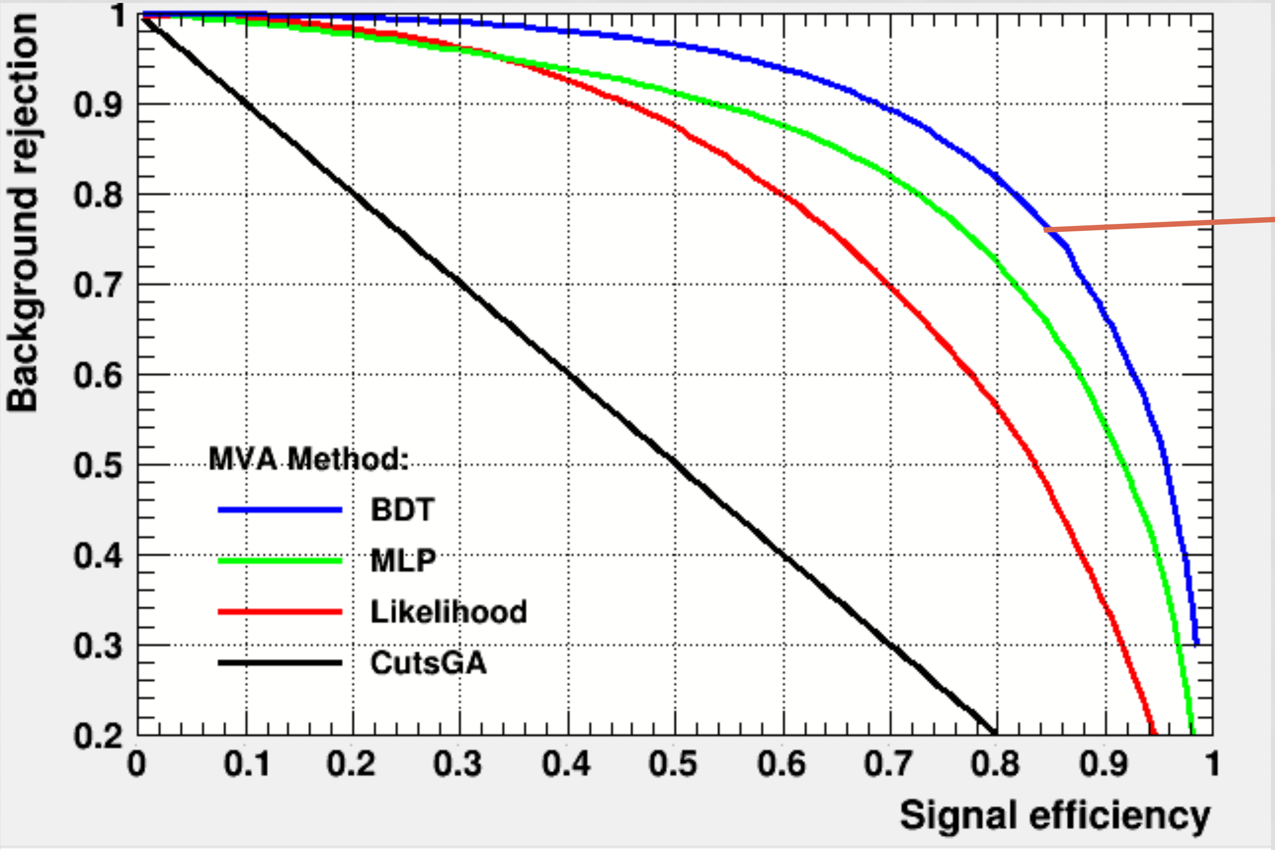
CORRELATION COEFFICIENTS





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

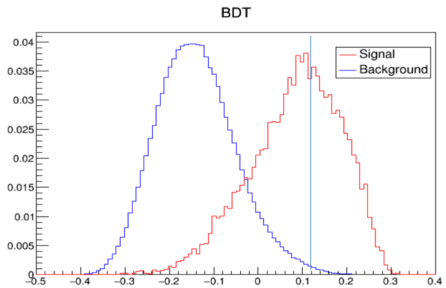
BDT

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.

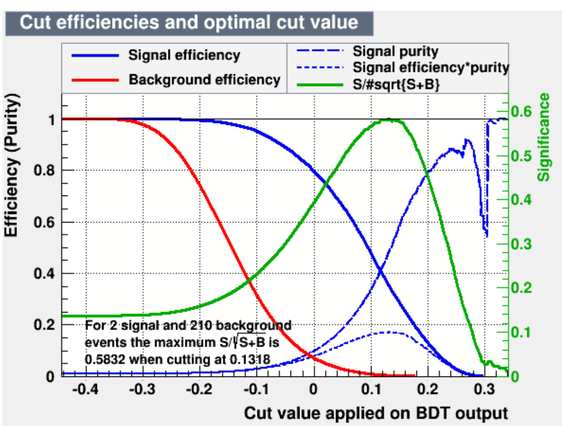
ALL MET RANGE

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**

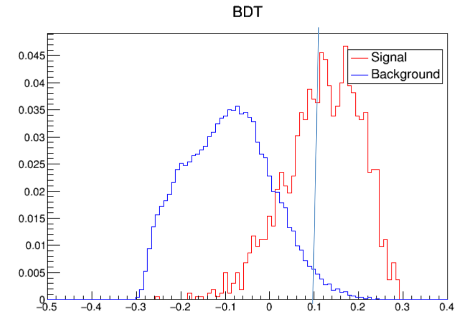


- Sign. Events: 1.98 (TestTree)
- Bckg. Events: 210.02 (TestTree)

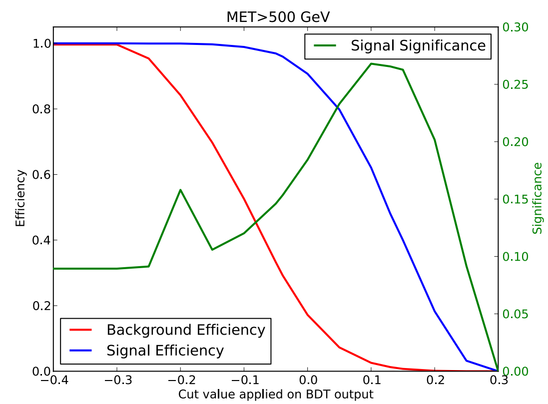
MET > 500 GeV

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

- **Classifier Output Distribution**

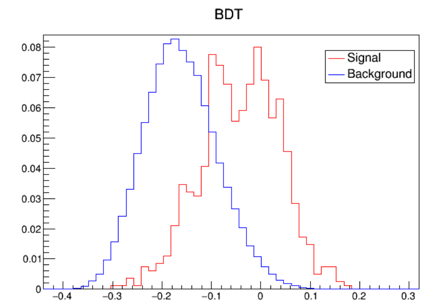


- **Classifier Cut efficiencies**

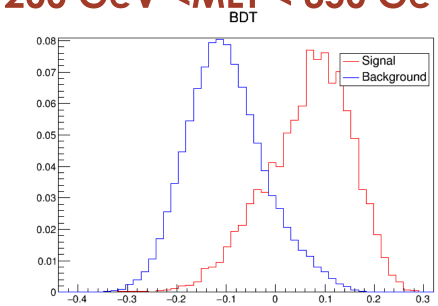


- Sign. Events: 0.28 (TestTree)
- Bckg. Events: 9.37 (TestTree)

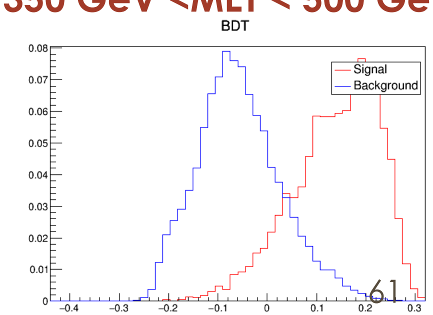
150 GeV <MET < 200 GeV



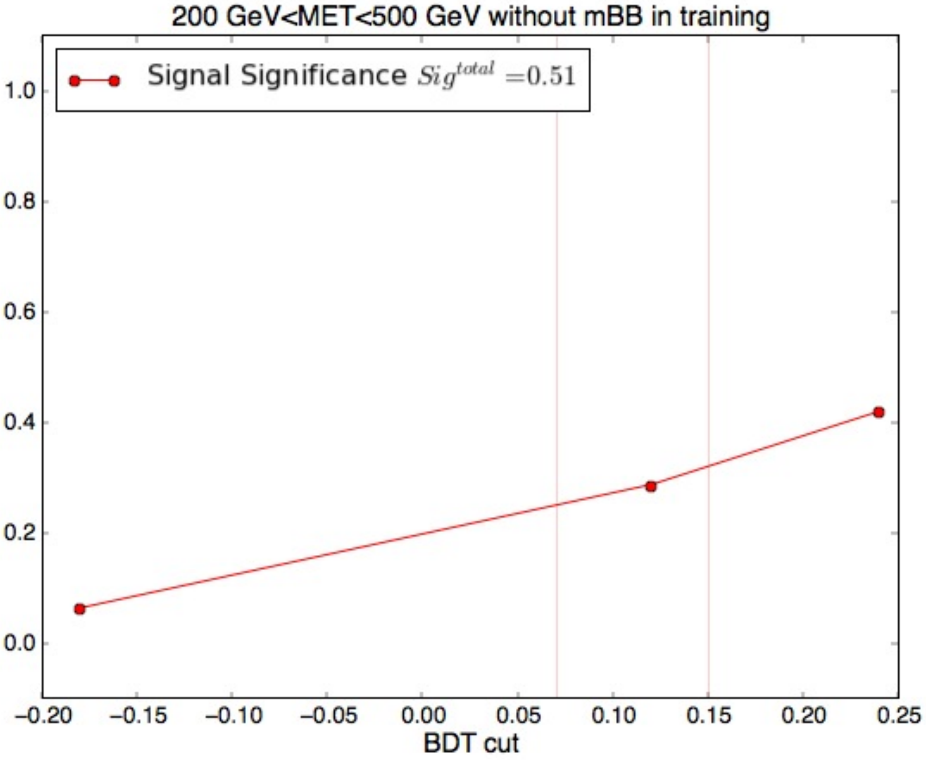
200 GeV <MET < 350 GeV

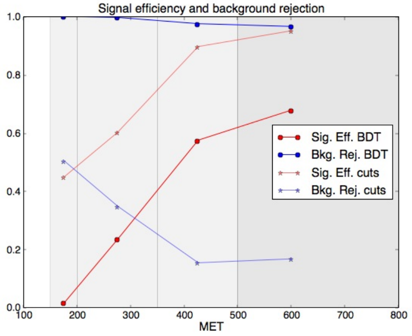


350 GeV <MET < 500 GeV



MERGED EVENTS WITH MET<500 GEV

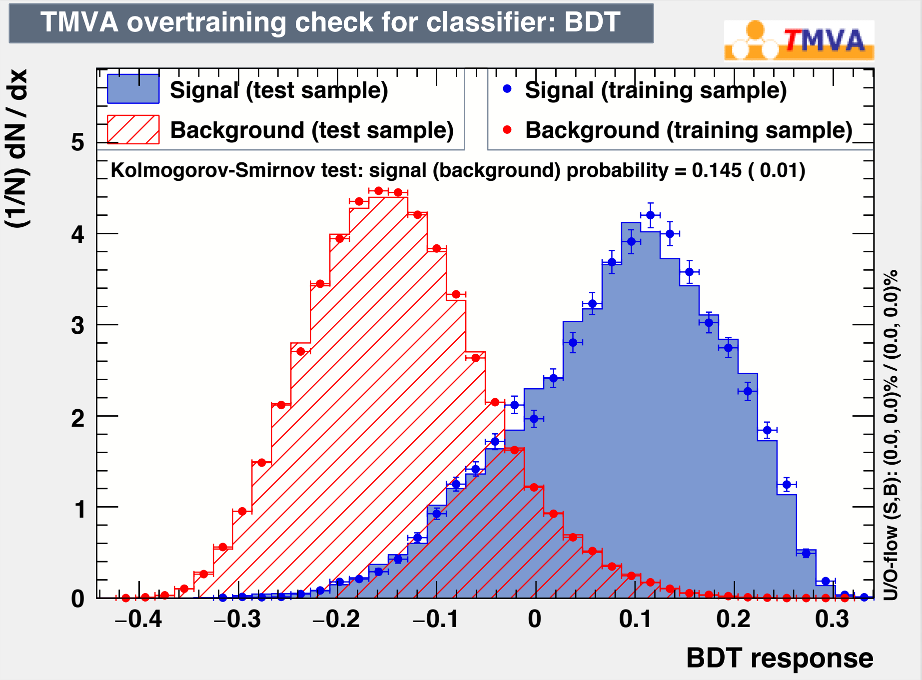




PRELIMINARY MVA ON SCALAR SM

- Variables in training

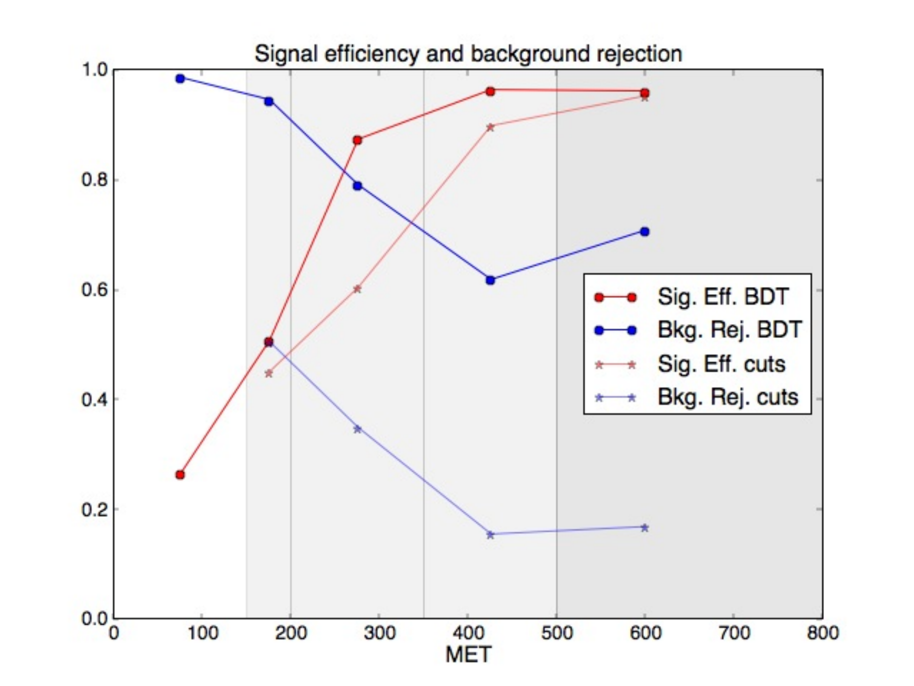
Rank	Variable	Variable Importance
1	EtaFatJet	9.03E-02
2	mindphi	8.67E-02
3	mBB	7.86E-02
4	dPhiMETMPT	7.41E-02
5	nJ	7.10E-02
6	MET	7.04E-02
7	btag	6.55E-02
8	dPhiMETFatJet	5.76E-02
9	MVH	5.12E-02
10	nTrackJetsTagsOut	4.94E-02
11	pTFatJet	4.58E-02
12	MPT	4.36E-02
13	nTrackJetsTagsIn	4.34E-02
14	MEff	4.13E-02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

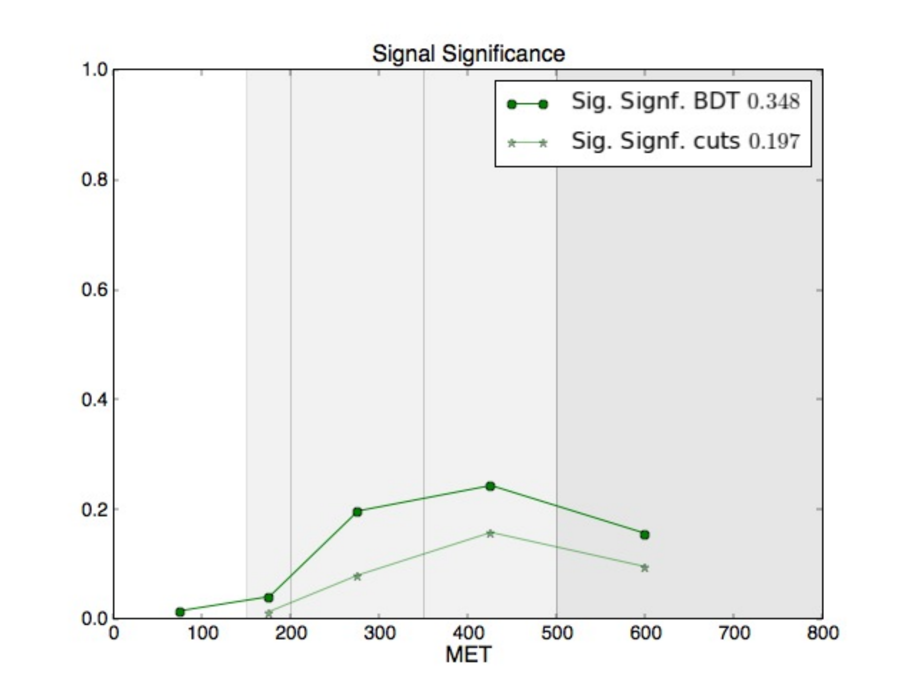


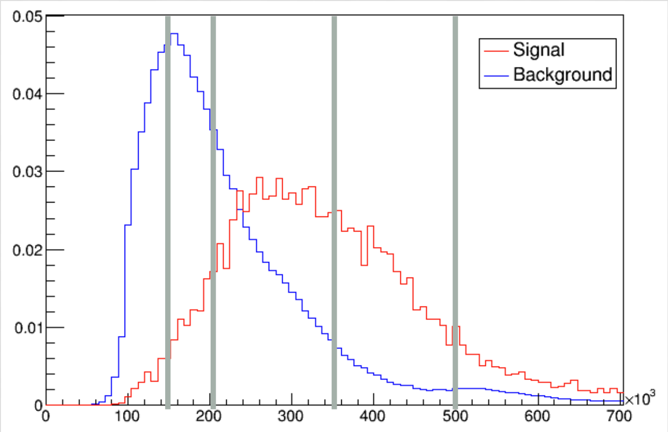
2

COMPARISON BETWEEN BDT CLASSIFIER AND CUTS BASED ANALYSIS

- MET categories





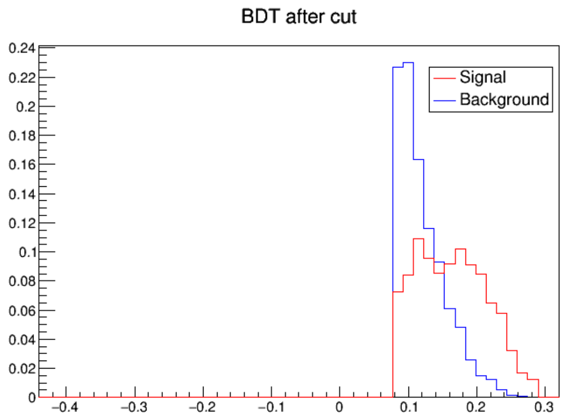


BDT cut = -0.0390 not optimized for the number of signal and background

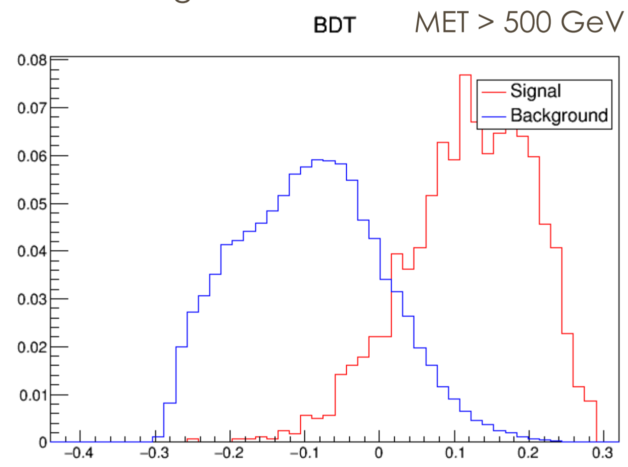
BDT CATEGORIES MET>500 GEV

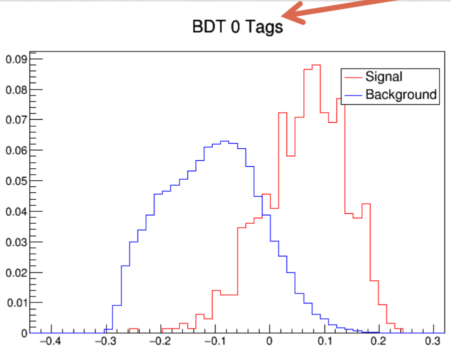
The number of b-tag jets is a variable considered in the BDT training.

Then, instead of having a single sharp cut on BDT

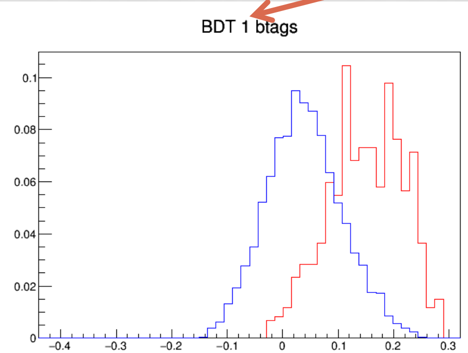


We can define BDT categories that can be related to 0, 1 and 2 btags.

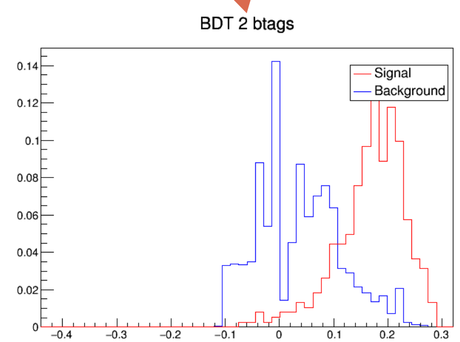




- Low BDT category



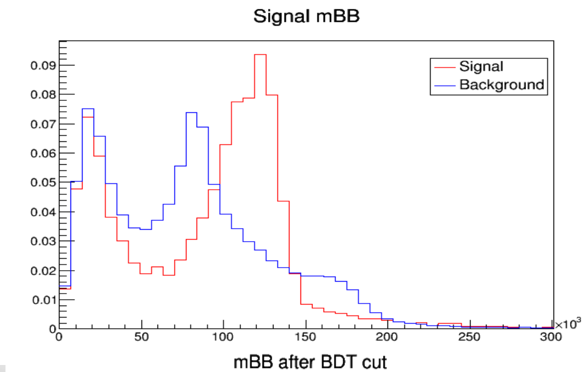
- Medium BDT category

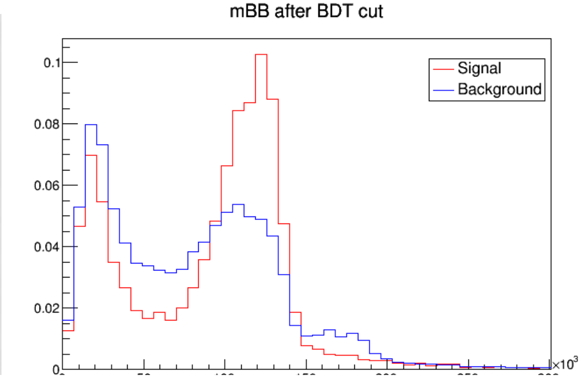


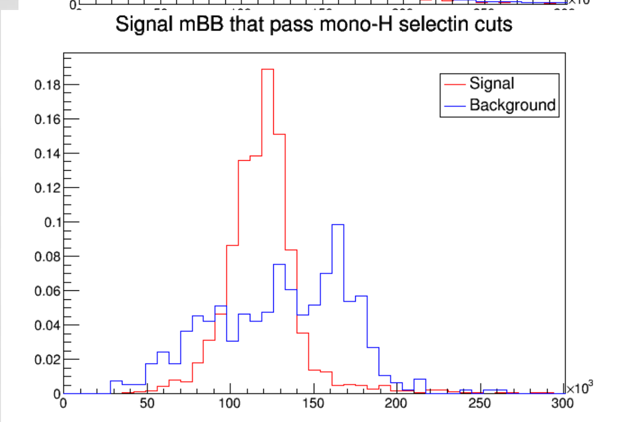
- High BDT category

Distributions of the variables

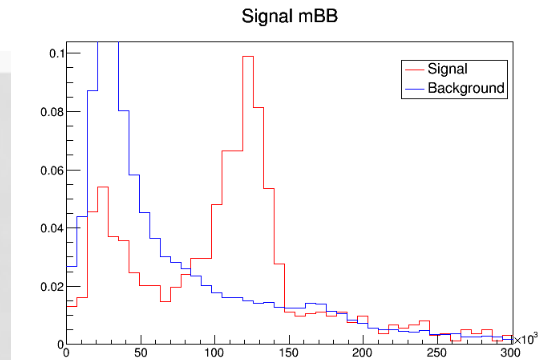
ALL MET

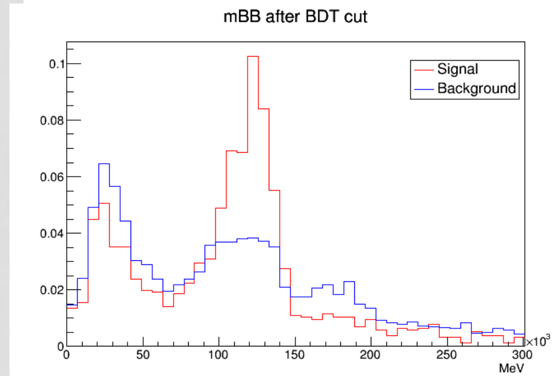


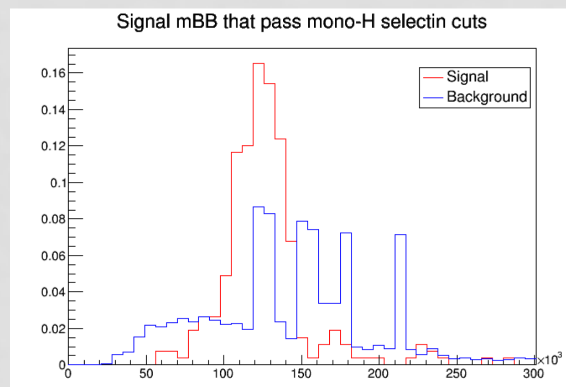




MET > 500 GeV



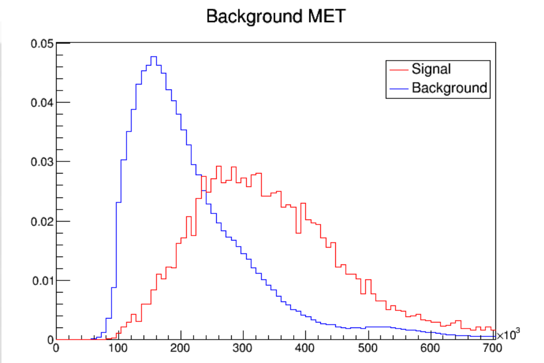


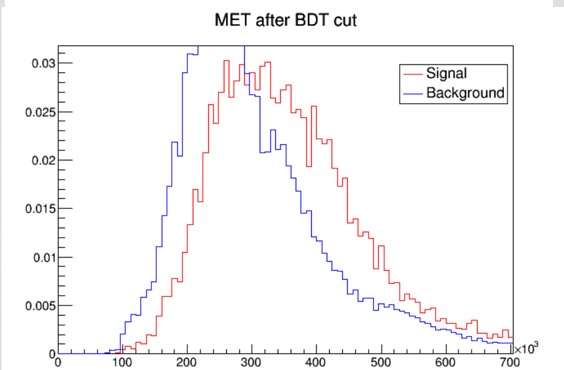


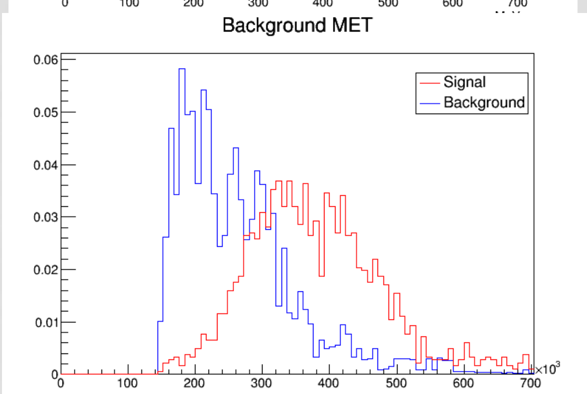
Distributions of the variables

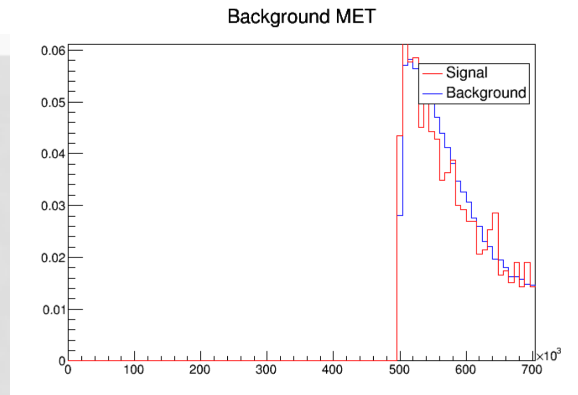
ALL MET

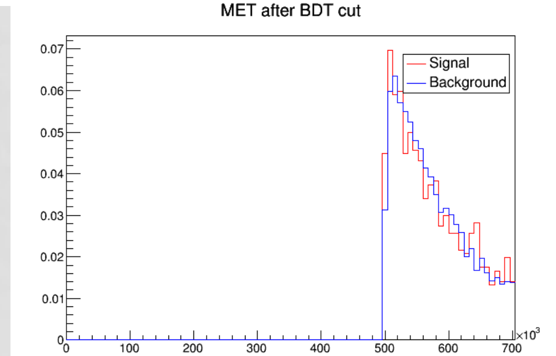
MET > 500 GeV

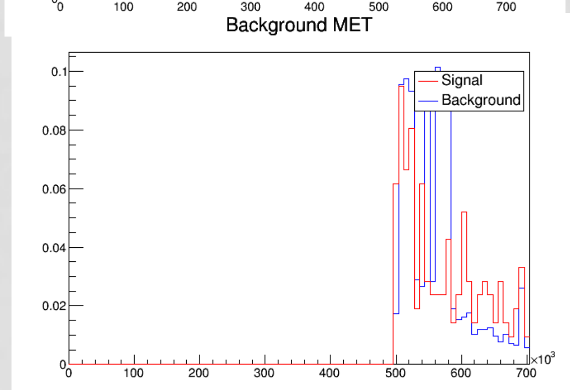






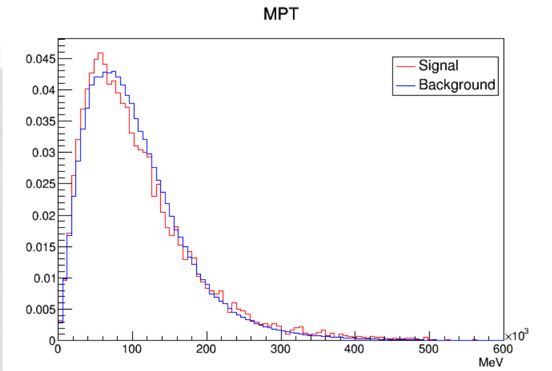


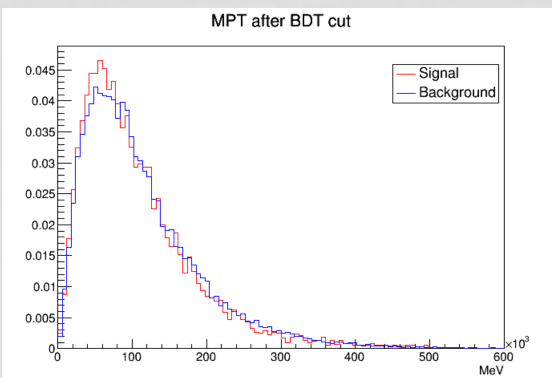


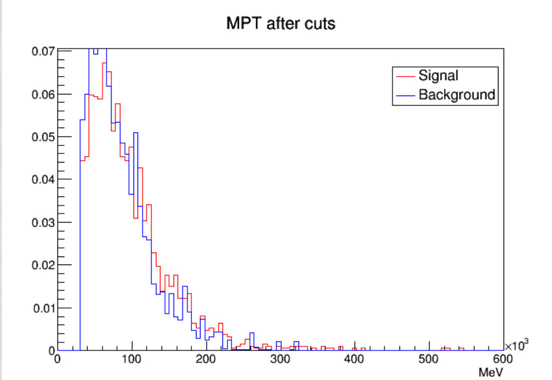


Distributions of the variables

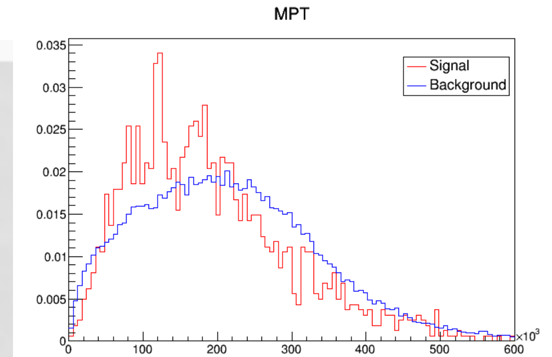
ALL MET

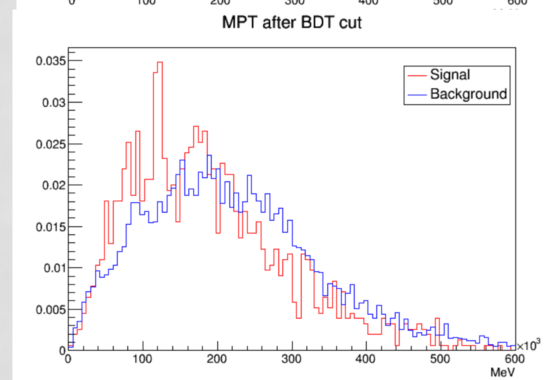


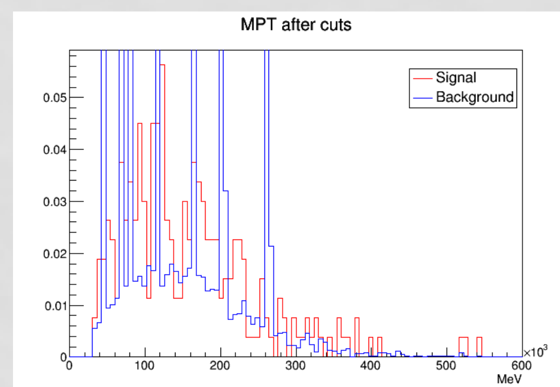




MET > 500 GeV

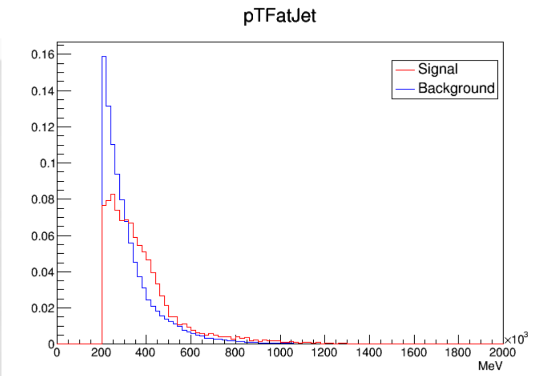


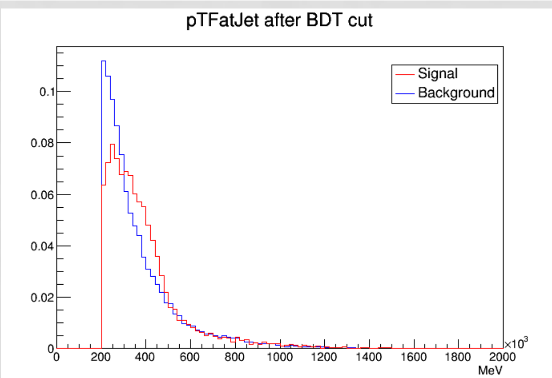


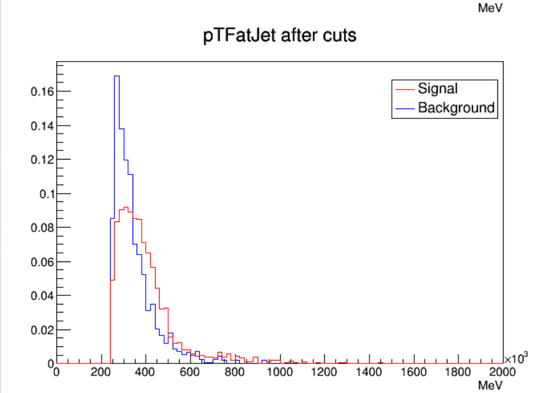


Distributions of the variables

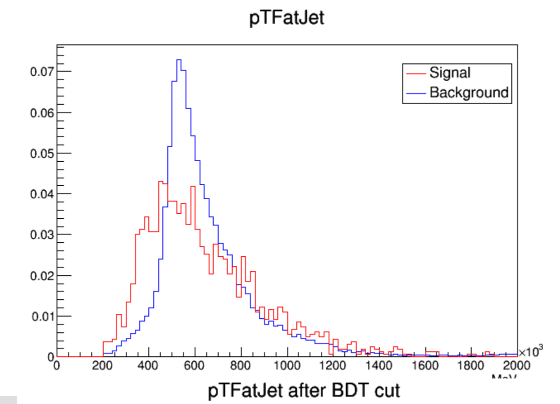
ALL MET

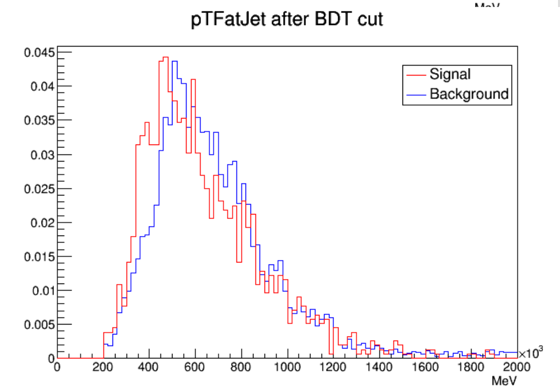


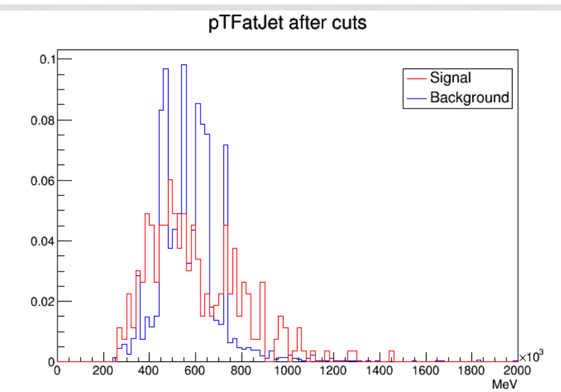




MET > 500 GeV



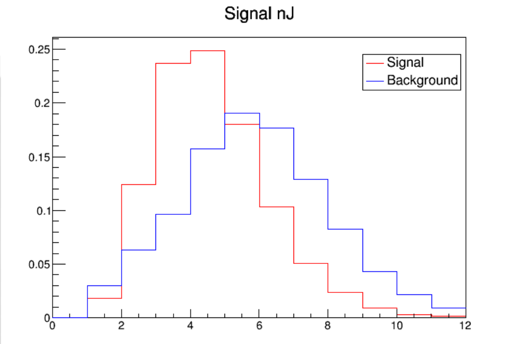


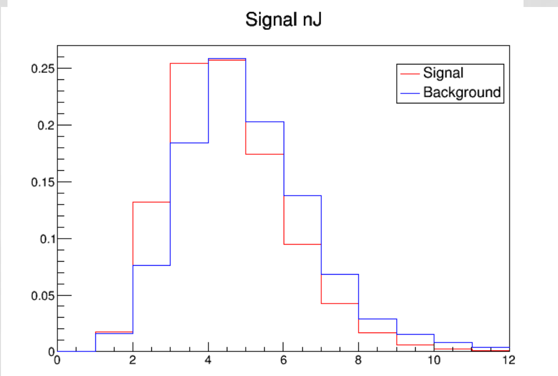


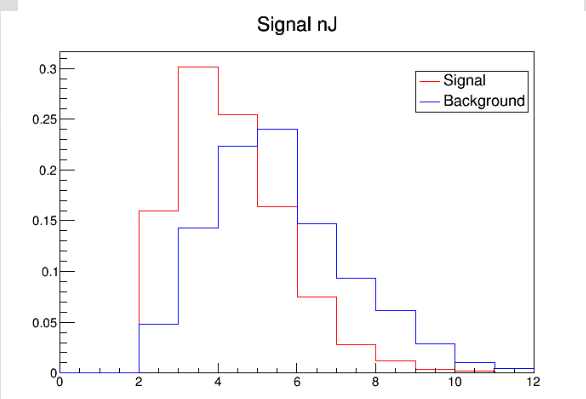
Distributions of the variables

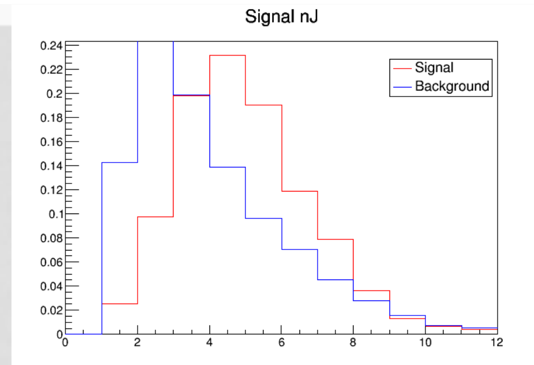
ALL MET

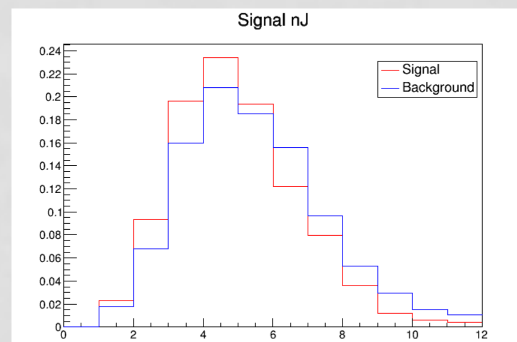
MET > 500 GeV

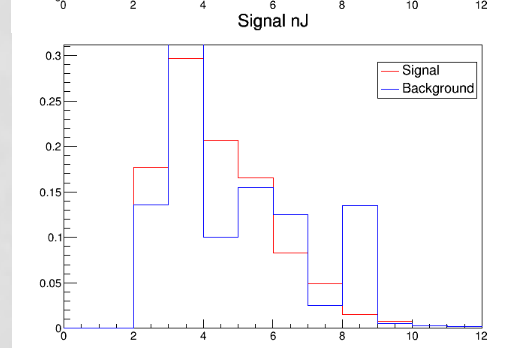






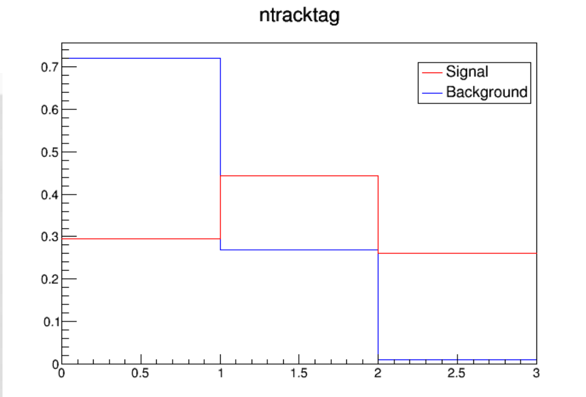


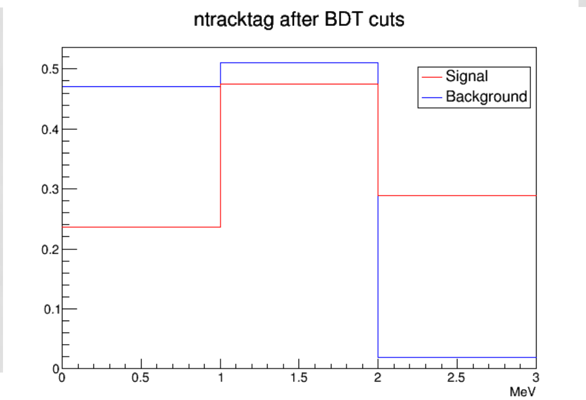


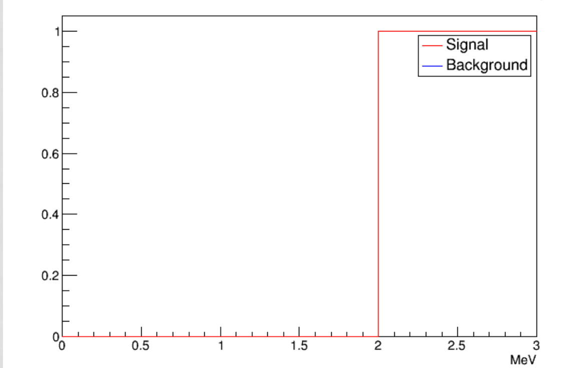


Distributions of the variables

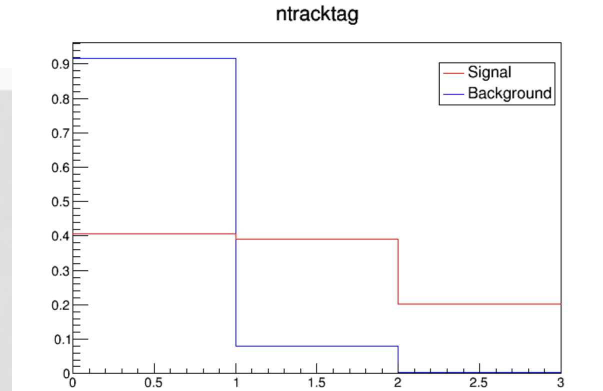
ALL MET

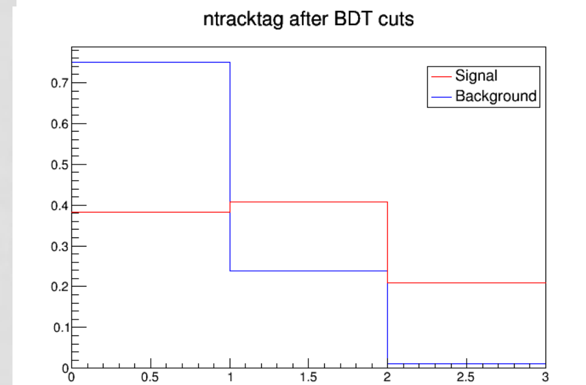


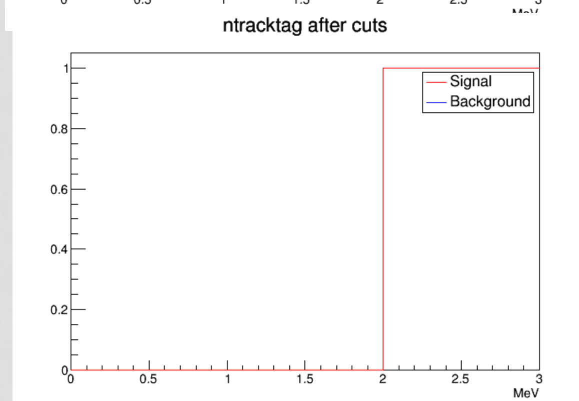




MET > 500 GeV

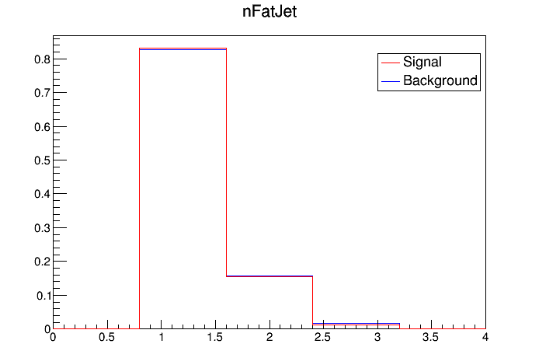


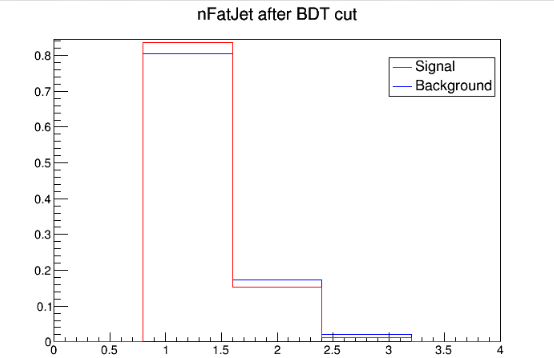


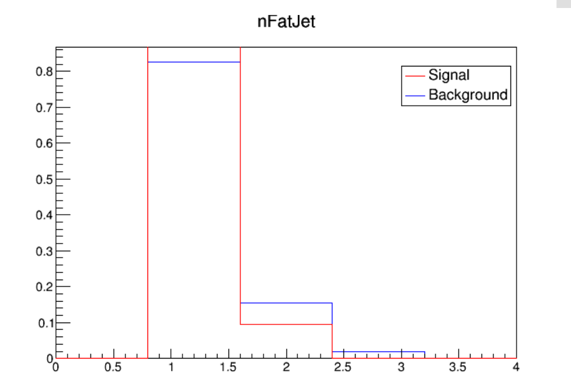


Distributions of the variables

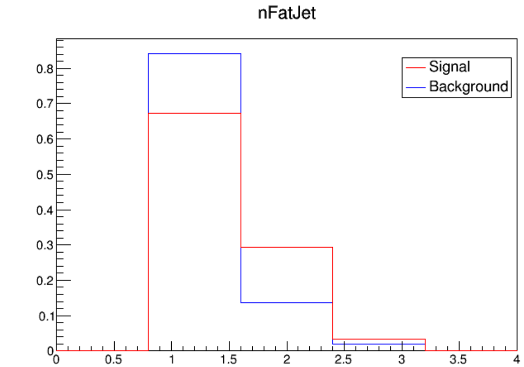
ALL MET

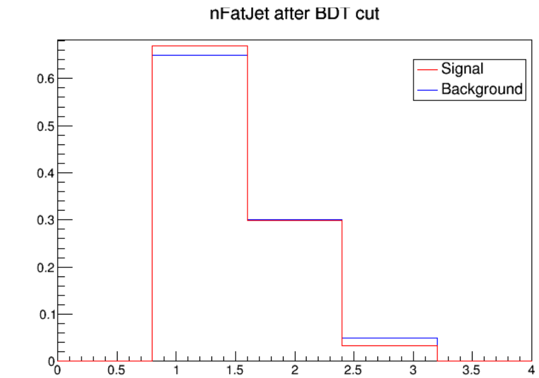


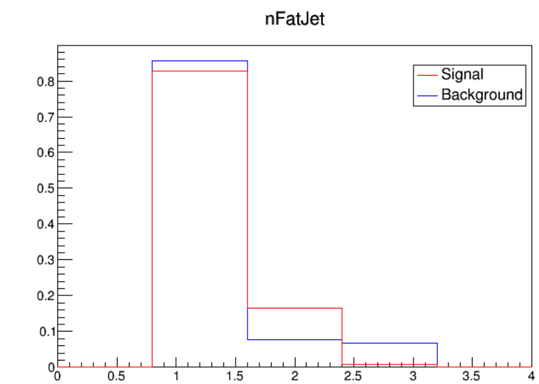




MET > 500 GeV

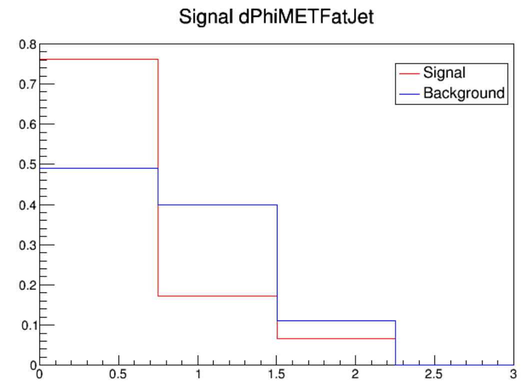


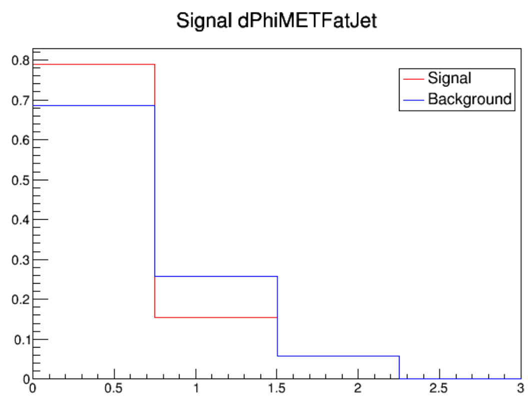


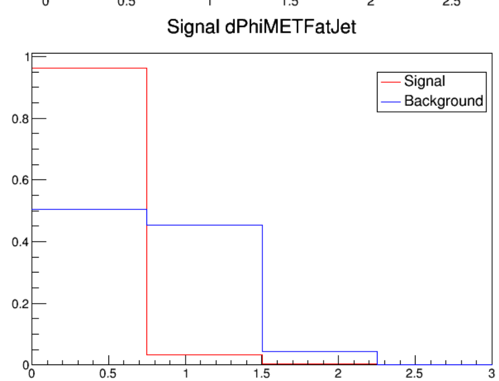


nTrackJetsTagsOut

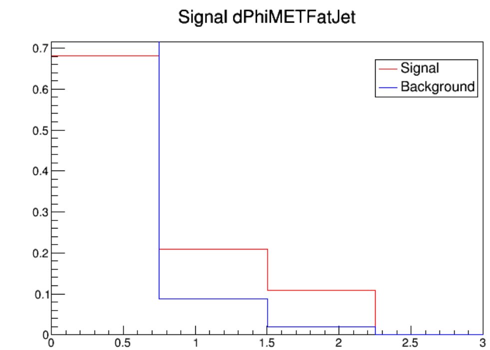
ALL MET

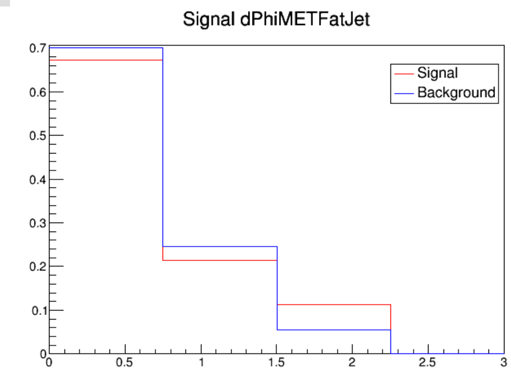


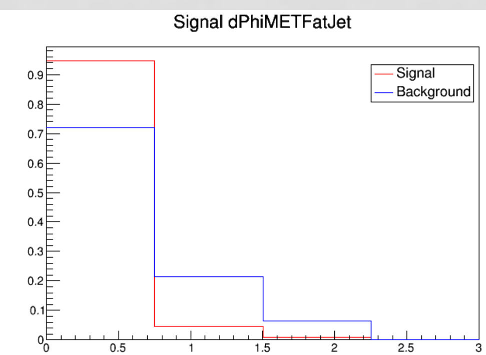




MET > 500 GeV

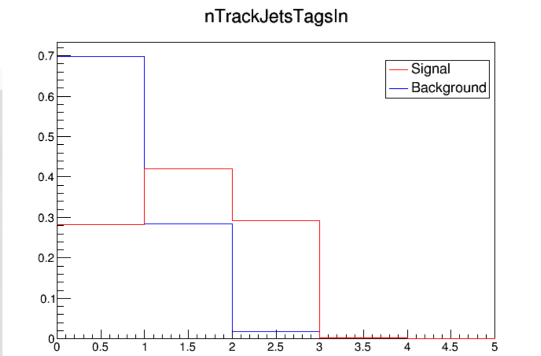


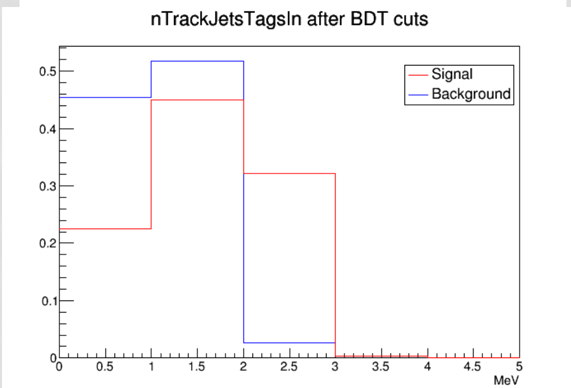


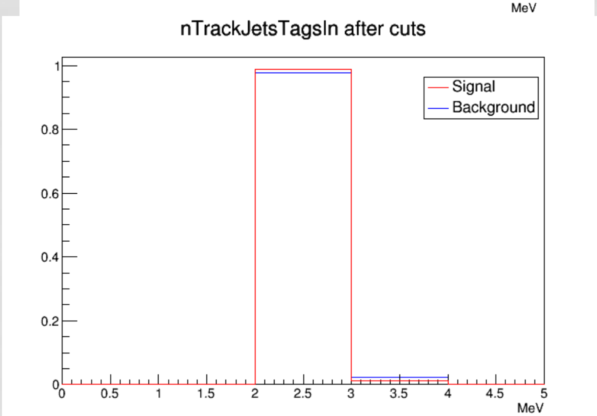


NtrackJetsTagsIn

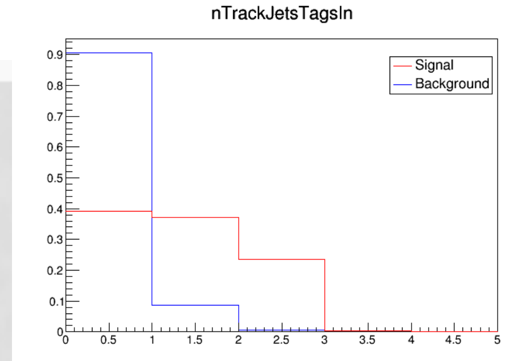
ALL MET

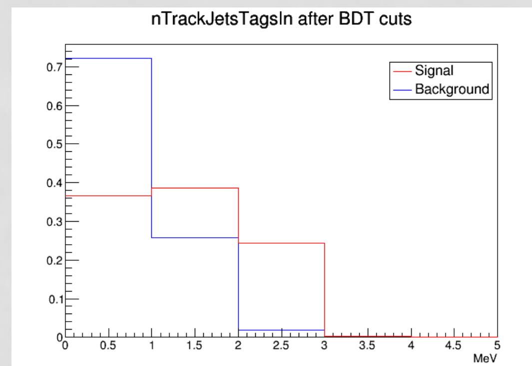


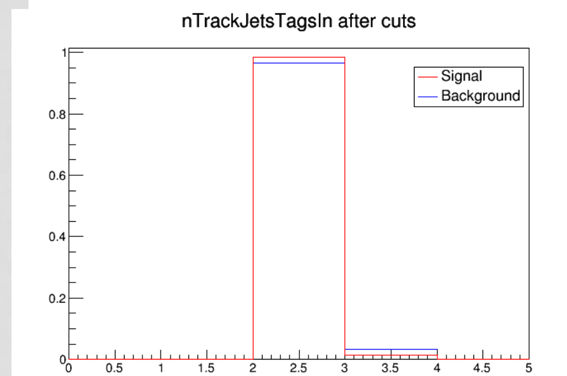




MET > 500 GeV



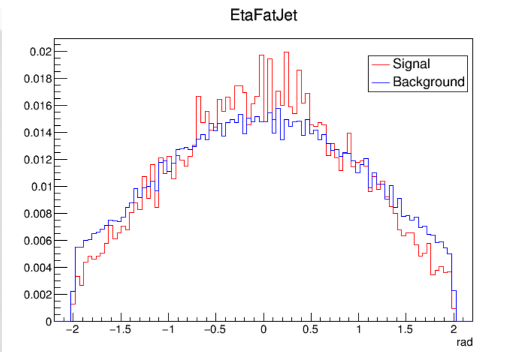


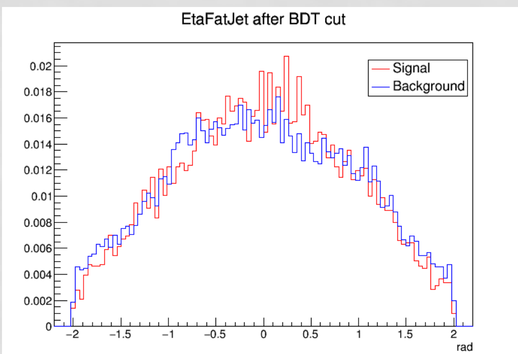


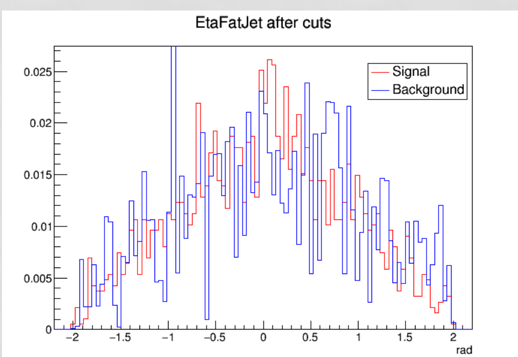
Distributions of the variables

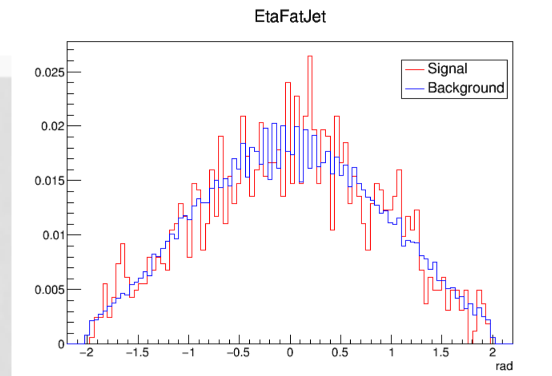
ALL MET

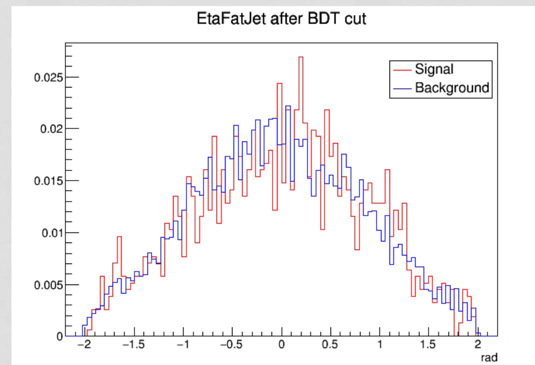
MET > 500 GeV

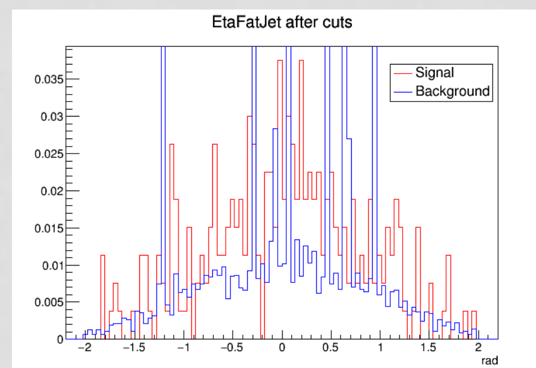








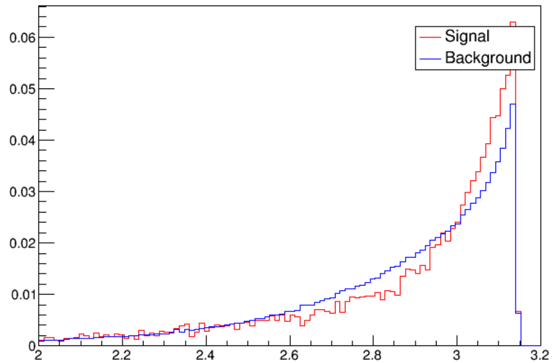




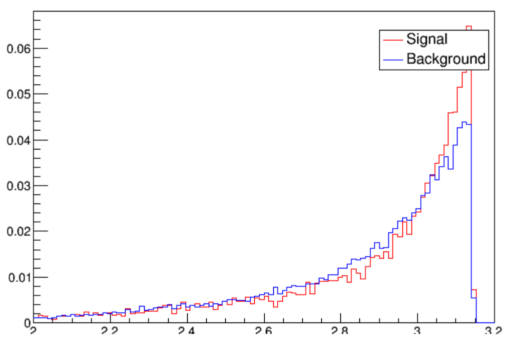
Distributions of the variables

ALL MET

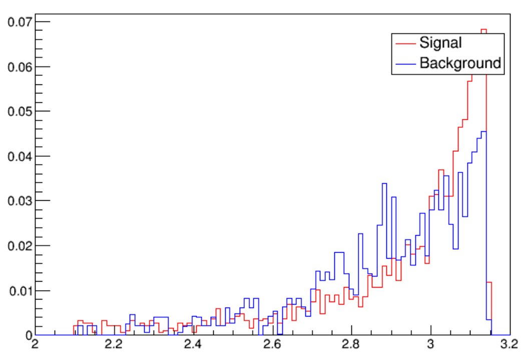
Signal dPhiMETFatJet



Signal dPhiMETFatJet

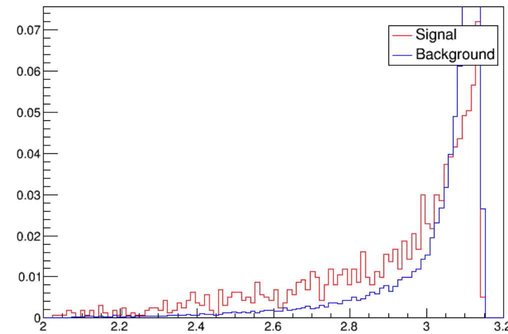


Signal dPhiMETFatJet

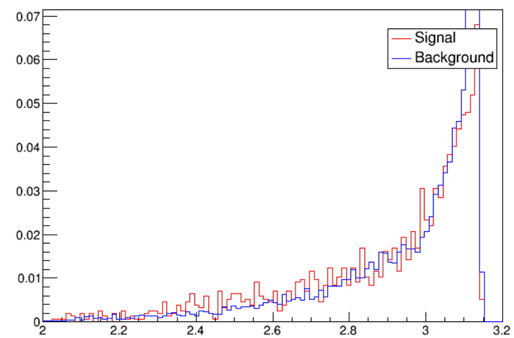


MET > 500 GeV

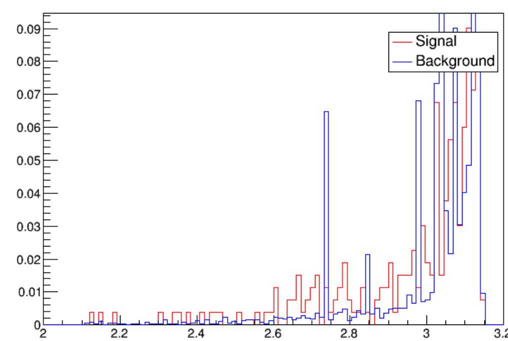
Signal dPhiMETFatJet



Signal dPhiMETFatJet

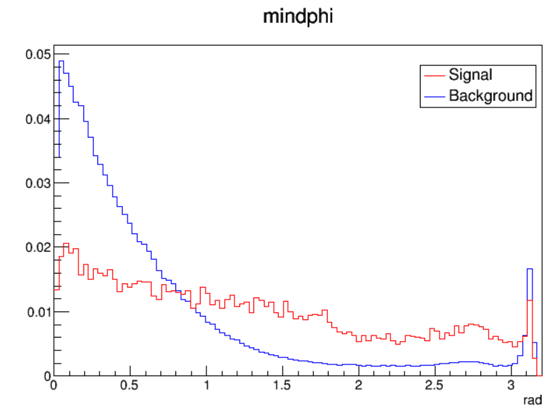


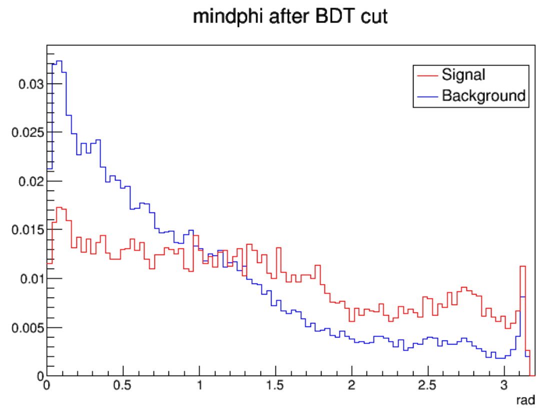
Signal dPhiMETFatJet

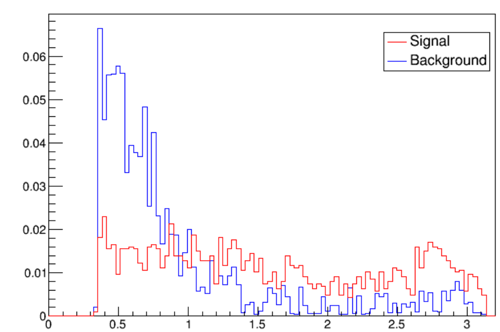


Distributions of the variables

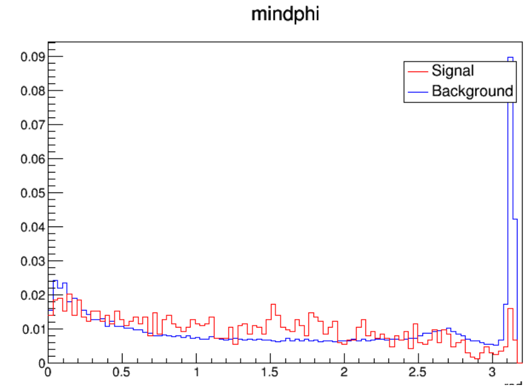
ALL MET

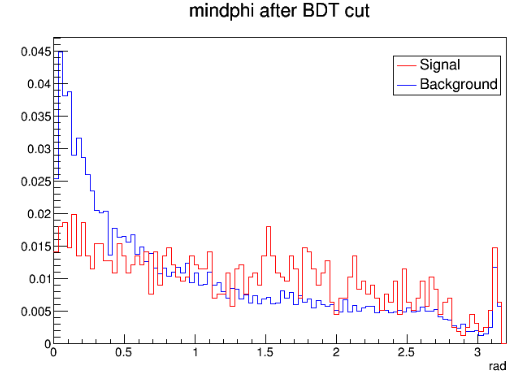


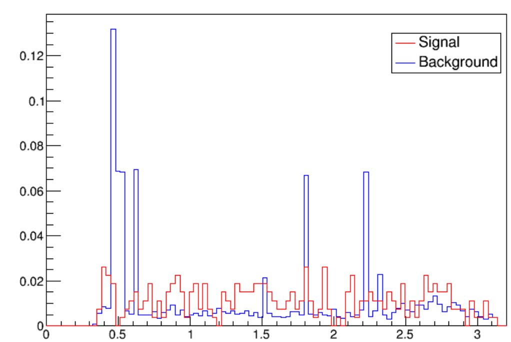




MET > 500 GeV



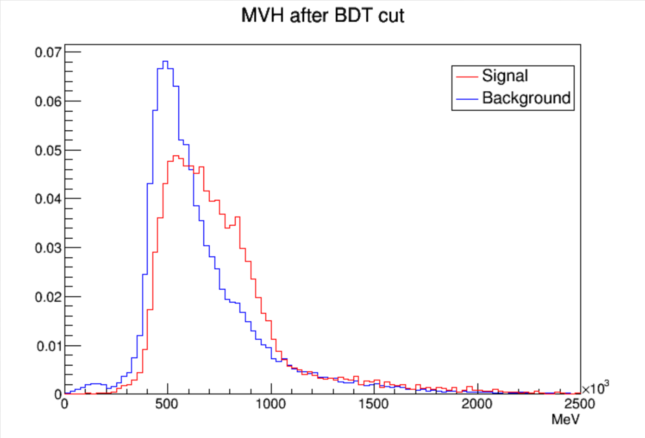


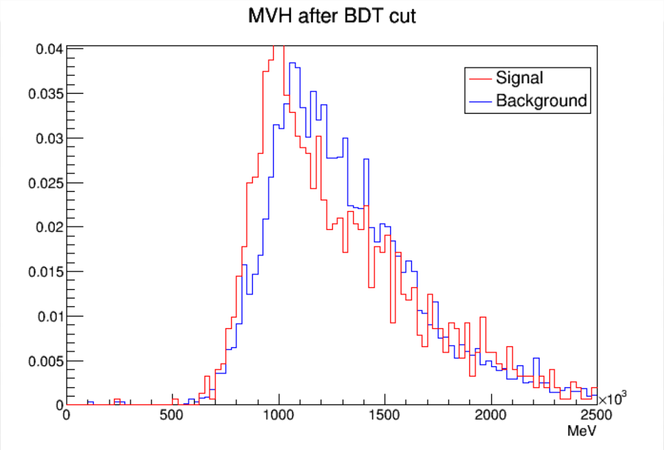


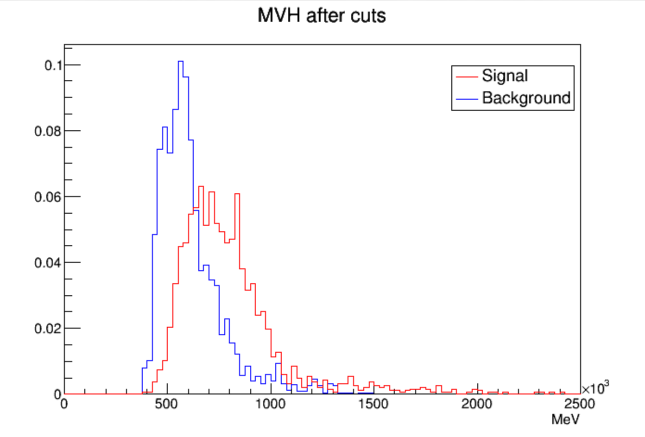
Distributions of the variables

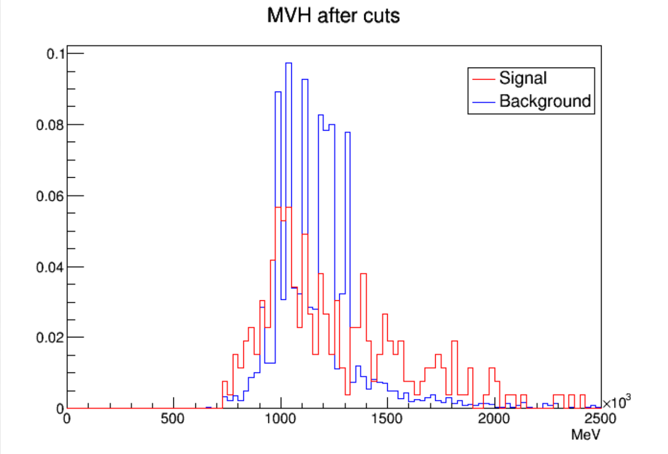
ALL MET

MET > 500 GeV



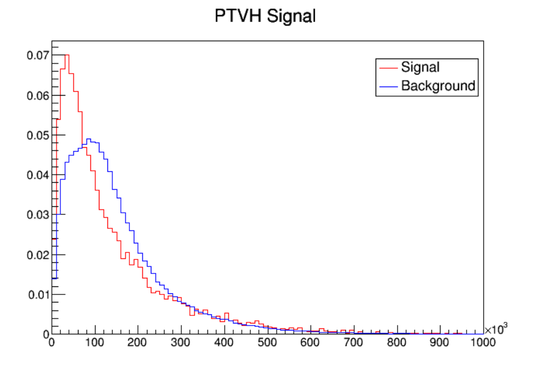


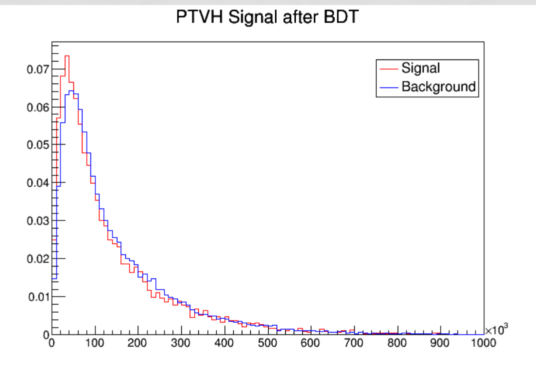


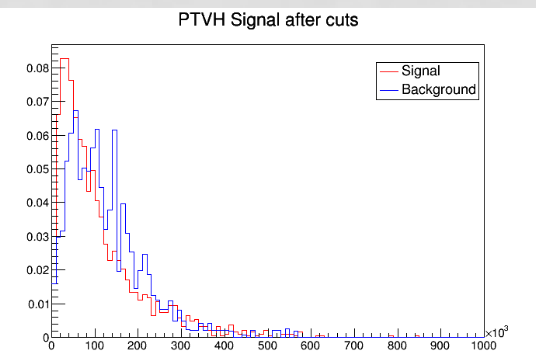


Distributions of the variables

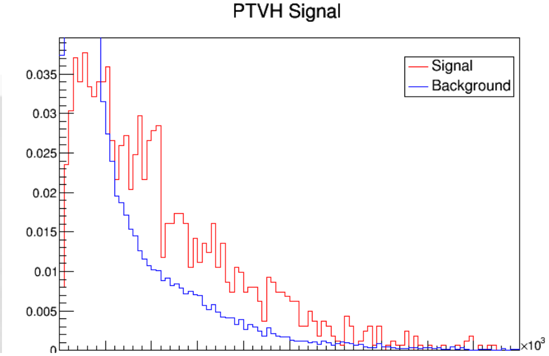
ALL MET

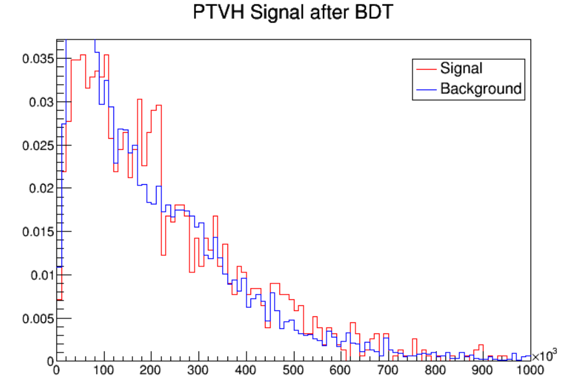


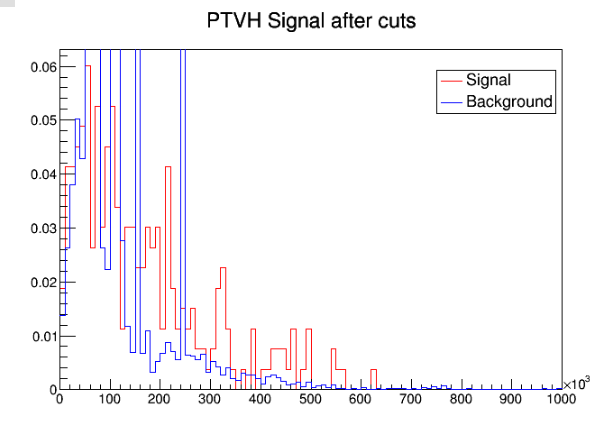




MET > 500 GeV

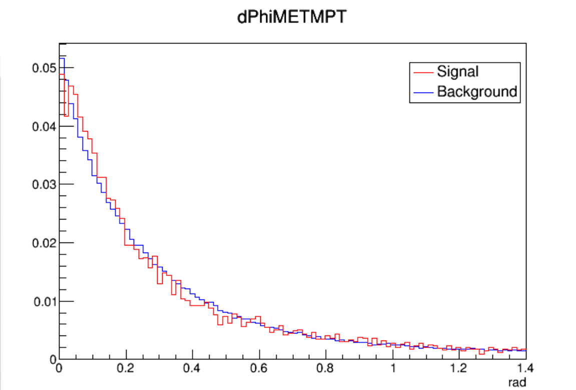


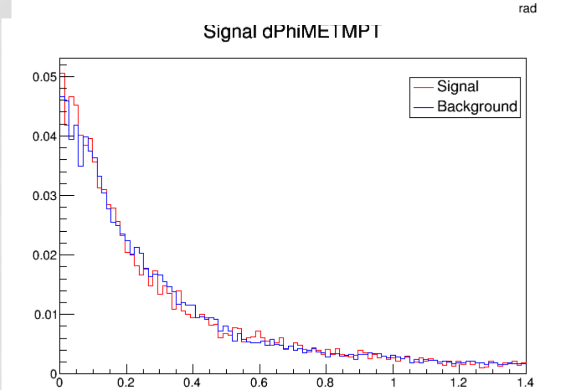


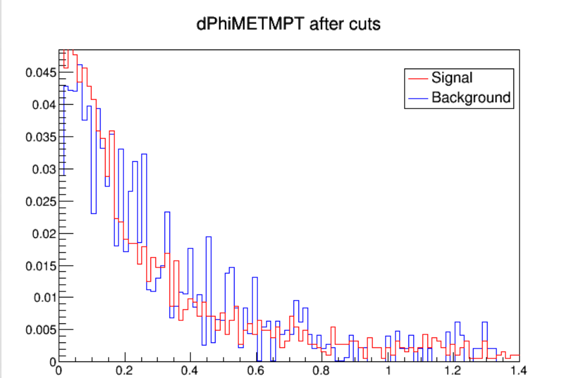


Distributions of the variables

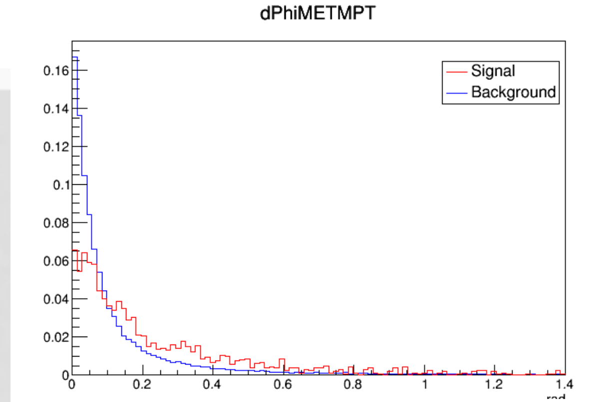
ALL MET

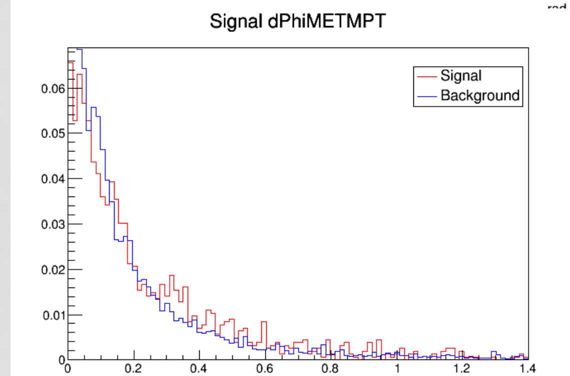


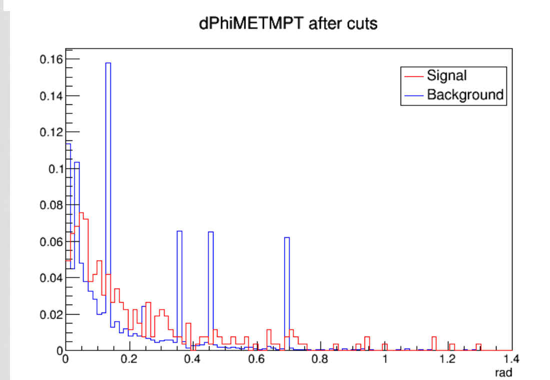




MET > 500 GeV

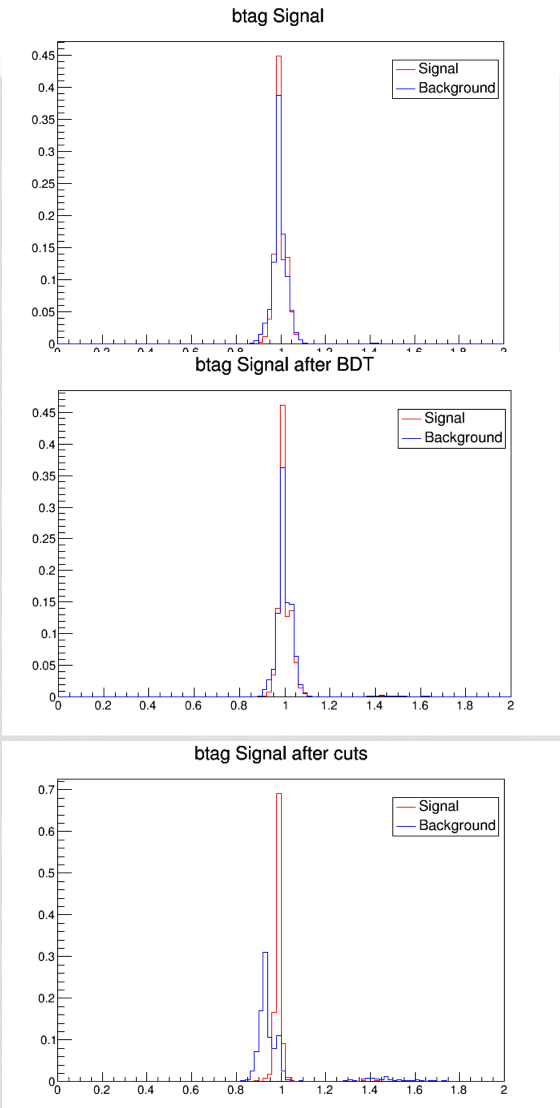




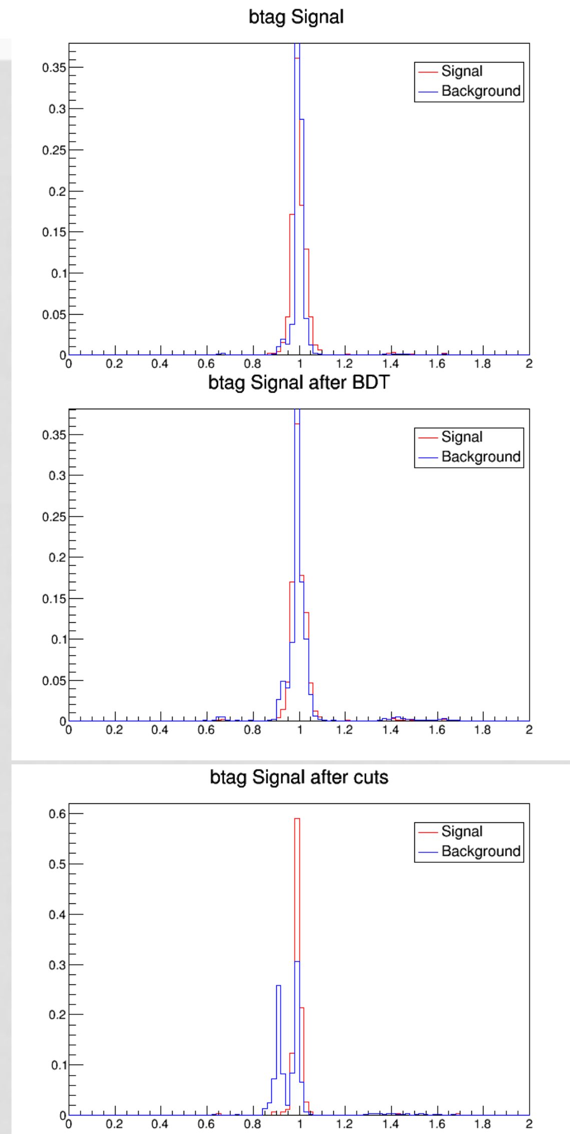


Distributions of the variables

ALL MET



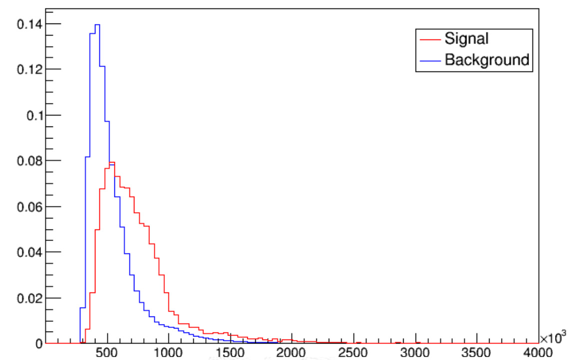
MET > 500 GeV



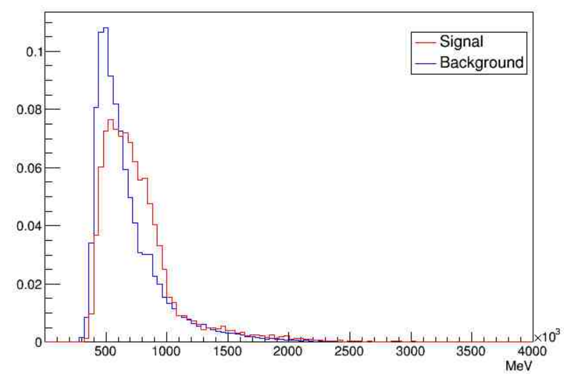
Distributions of the variables

ALL MET

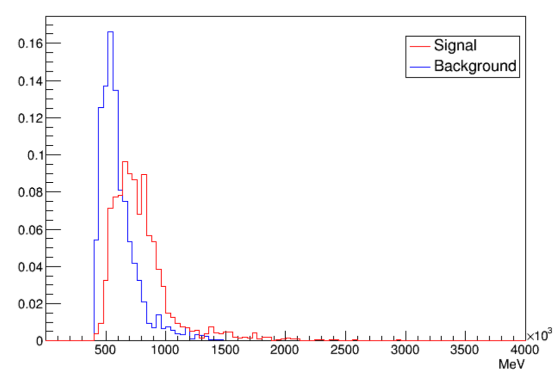
MEff



MEff after BDT cut

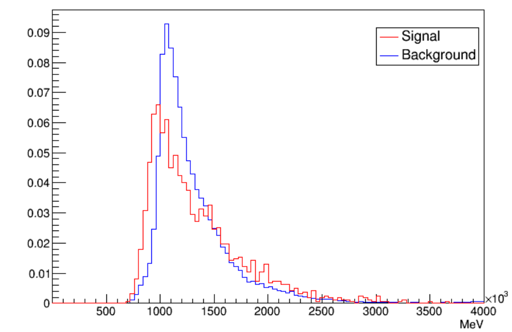


MEff after cuts

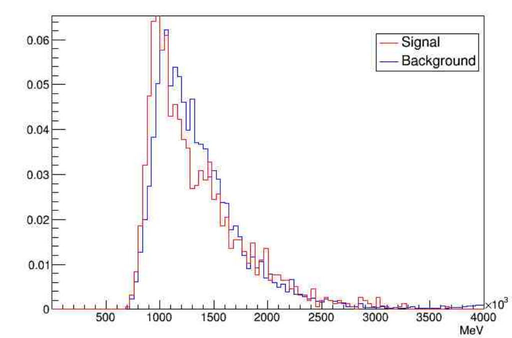


MET > 500 GeV

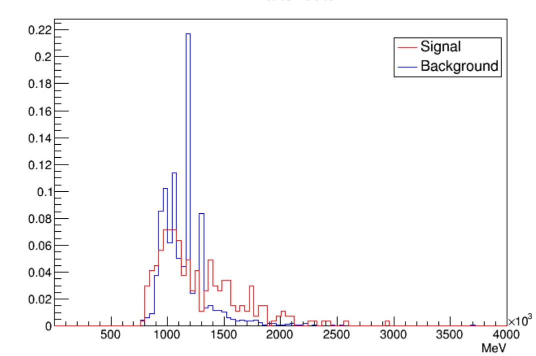
MEff



MEff after BDT cut

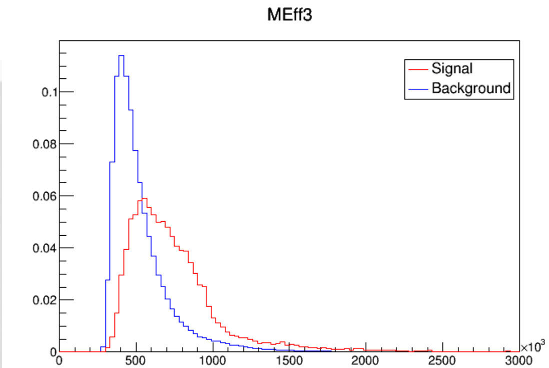


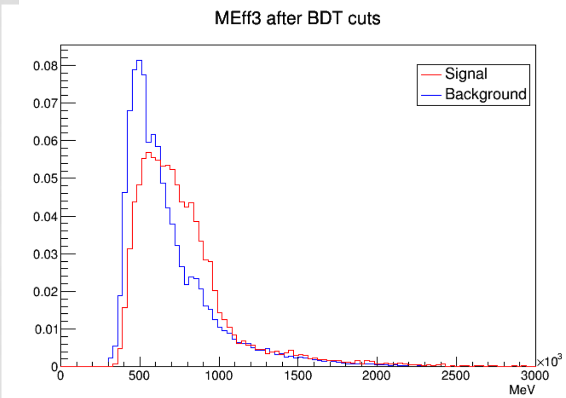
MEff after cuts

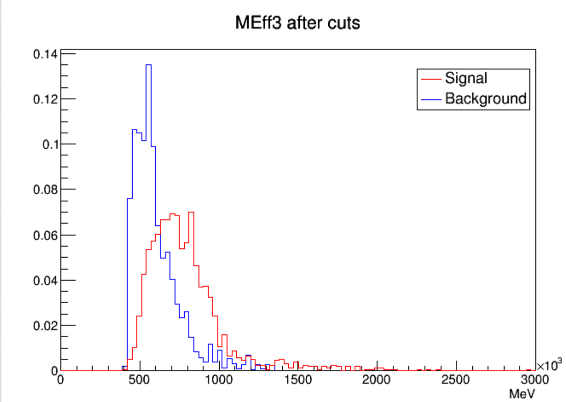


Distributions of the variables

ALL MET







MET > 500 GeV