Search for the Standard Model Higgs boson in $H \to WW$ decays with two associated jets using Boosted Decision Trees

Suche nach dem Standard-Modell Higgs-Boson in $H \to WW$ Zerfällen mit zwei assoziierten Jets mit Hilfe von Boosted Decision Trees

Masterarbeit an der Fakultät für Physik

der

Ludwig-Maximilians-Universität München

vorgelegt von

Philipp Heimpel

geboren am 16. Dezember 1987 in Tettnang

11. April 2013

Gutachterin: Prof. Dr. Dorothee Schaile

The Zen of Python

Beautiful is better than ugly.
Explicit is better than implicit.
Simple is better than complex.
Complex is better than complicated.
Flat is better than nested.
Sparse is better than dense.
Readability counts.
Special cases aren't special enough to break the rules.
Although practicality beats purity.
Errors should never pass silently.
Unless explicitly silenced.
In the face of ambiguity, refuse the temptation to guess.
There should be one - and preferably only one - obvious way to do it.
Although that way may not be obvious at first unless you're Dutch.
Now is better than never.
Although never is often better than *right* now.
If the implementation is hard to explain, it's a bad idea.
If the implementation is easy to explain, it may be a good idea.
Namespaces are one honking great idea - let's do more of those!

Tim Peters

Abstract

This thesis describes a search for a Standard Model Higgs boson at $m_H = 125\,\text{GeV}$ in the decay channel $H \to W^+W^- \to \bar{l}\nu_l l\bar{\nu}_l$ with two associated jets. A dataset with an integrated luminosity of $\mathcal{L} = 12.9\,\text{fb}^{-1}$, collected by the ATLAS-Experiment at the Large Hadron Collider, where protons collide at a center of mass energy of $\sqrt{s} = 8\,\text{TeV}$ was used.
Boosted Decision Trees (BDTs) are used to separate the Higgs boson signal produced in vector boson fusion from the background, which is mainly top quark pair production. The goal is to increase the significance $S/\sqrt{B}$ and to obtain similar results as the cut-based analysis while maintaining good data to Monte Carlo agreement. This is done by finding the best set of variables and parameters for the BDTs.

Zusammenfassung

In dieser Arbeit werden Studien zur Suche nach dem Standard-Modell Higgs-Boson mit einer Masse von $m_H = 125\,\text{GeV}$ vorgestellt. Dabei wird der Higgs-Boson Zerfallskanal $H \to W^+W^- \to \bar{l}\nu_l l\bar{\nu}_l$ in Assoziation mit zwei Jets genauer untersucht. Die verwendeten Daten mit einer integrierten Luminosität von $\mathcal{L} = 12.9\,\text{fb}^{-1}$ wurden am ATLAS-Experiment am LHC gewonnen und stammen aus Proton-Proton-Kollisionen bei einer Schwerpunktsenergie von $\sqrt{s} = 8\,\text{TeV}$.
Der Schwerpunkt der Studien ist der Vektorboson Fusionskanal und die damit verbundene Reduzierung des top Quark Paar Untergrundes mit Hilfe von BDTs. Ziel ist es die Signifikanz $S/\sqrt{B}$ zu erhöhen um ähnliche Ergebnisse wie die schnittbasierte Analyse zu erhalten. Dabei wird darauf geachtet, ein gutes Verhältnis von Daten zu Monte Carlo zu haben. Dieses Ziel soll erreicht werden, indem optimale Variablen und Parameter für die BDTs gefunden werden.

Contents

1 Introduction

Over the last few decades the Standard Model (SM) of particle physics has proven itself successfully through experimental validation. It describes the fundamental particles using three generations of leptons and quarks. The interaction between these particles is described by the electromagnetic, the weak and the strong force mediated by four types of gauge bosons. The last not yet fully experimentally validated particle is the Higgs boson predicted in 1964 by Higgs, Brout and Englert [1] [2].
It manifests itself as an excitation of the Higgs field, which gives mass to massive particles by breaking the electroweak symmetry.
In July 2012, a Higgs-like particle was discovered at the Large Hadron Collider (LHC) with a mass of $126 \pm 0.4\,(\text{stat.}) \pm 0.4\,(\text{sys.})\,\text{GeV}$. The signal showed an excess from the background-only hypothesis expressed as the number of standard deviations of $5.0\,\sigma$ for CMS [3] and $6.0\,\sigma$ ATLAS [4].
In this thesis the Higgs boson in the decay channel $H \to WW \to l\bar{\nu}\bar{l}\nu$ will be studied, in the vector boson fusion production of the Higgs boson exclusively. This will be done by using a dataset with an integrated luminosity of $\mathcal{L} = 12.9\,\text{fb}^{-1}$, obtained at a center of mass energy of $\sqrt{s} = 8\,\text{TeV}$ at the ATLAS experiment. The two emitted neutrinos in this decay cannot be detected, which results in missing transverse energy. That leads to a limited mass resolution, but the channel is still interesting. The reason is that the coupling of a Higgs boson to vector bosons can be shown. Until now this is done by using a classical cut-based analysis to separate the signal from background processes. In this thesis, this will be done by using BDTs, a multivariate classifier where optimal cut values and variables are determined by an algorithm. The main challenge is to find the best parameters and determine which variables give the best signal versus background separation.
This thesis is divided into four parts. First, the theoretical background where a short introduction to the SM, the Higgs boson, and BDTs will be given. The second part will give an overview of the LHC, the ATLAS detector, the reconstruction and selection of objects, and event processing. The third and main part deals with the search for the Higgs boson using BDTs. It contains a description of the signal and background processes and the event selection. This outline will be followed by the description of systematic uncertainties. Afterwards different BDTs will be evaluated and compared with respect to parameters, variables, and channels. In the last part, an overview of the results will be given.

2 Theoretical Background

In this section, the theoretical background of the thesis will be illustrated. Starting with the SM and the Higgs mechanism, the description will be followed by the production and decay processes of the Higgs boson. In a further step the multivariate techniques used will be explained with an emphasis on BDTs. Finally, a short introduction to event shape variables will be given.

2.1 The Standard Model

The SM describes three generations of fermionic particles as listed in table 1. In every generation there are two quarks, each with three colors, one charged lepton (-1), and one uncharged lepton-neutrino. The up and down type quarks carry a charge of $+2/3$ and $-1/3$ respectively. Their masses increase with generation. Every particle has its antiparticle with reversed signs for charge-like quantum numbers.

	1st generation	2nd generation	3rd generation
Quarks	u	c	t
	d	s	b
Leptons	e	μ	τ
	ν_e	ν_μ	ν_τ

Table 1: The three generations of leptons and quarks in the SM [5].

The forces between these particles are mediated by gauge bosons as shown in table 2. Where the photon mediates the electromagnetic force, the $W^\pm$ and Z^0 transmit the weak force, and eight bi-colored gluons are the mediators for the strong force.

Force	Mediator	Charge [e]	Mass [GeV]
Strong	8 gluons	0	0
Electromagnetic	photon	0	0
Weak	$W^\pm$	± 1	80.385 ± 0.015
	Z^0	0	91.1876 ± 0.0021

Table 2: Overview of the mediator spin 1 bosons [6] [5].

2.2 The Standard Model Lagrangian

In quantum field theory the dynamics of a system are described by the Euler-Lagrange equation

$$\frac{\partial}{\partial x_\mu}\left(\frac{\partial \mathcal{L}}{\partial(\partial\phi/\partial x_\mu)}\right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0 \quad , \tag{1}$$

with fields $\phi(x,\, t)$, four vectors x_μ, and Lagrange density $\mathcal{L}$ [7]. The standard model Lagrangian can be decomposed into three parts: the electroweak, the quantum chromo dynamics (QCD), and the Higgs part. In the next section, these parts will be described.

Electroweak Lagrangian:
The electroweak interaction is described by the $SU(2) \times U(1)$ group. There are four fields: An isotriplet of vector fields $\mathbf{W}_\mu$ and a vector field B_μ. After spontaneous symmetry breaking (see section 2.3), a superposition of these fields form the four mediators which are mass eigenstates of the Lagrangian:

$$A_\mu = B_\mu \cos\theta_W + W^3_\mu \sin\theta_W \,, \tag{2a}$$

$$Z_\mu = -B_\mu \sin\theta_W + W^3_\mu \cos\theta_W \,, \tag{2b}$$

$$W^\pm_\mu = \frac{1}{\sqrt{2}}\left(W^1_\mu \mp i W^2_\mu\right) \,. \tag{2c}$$

Here θ_W is the weak mixing angle. The interaction is described by the covariant derivative

$$D_\mu = \partial_\mu - ig\frac{\boldsymbol{\tau}}{2}\mathbf{W}_\mu - ig'\frac{Y}{2}B_\mu \,, \tag{3}$$

where g and g' are the coupling strengths and $\boldsymbol{\tau}$ is the isospin. The relation between weak hypercharge Y, electromagnetic charge Q, and the third component of the isospin I_3 is

$$Y = 2(Q - I_3) \,. \tag{4}$$

The electroweak field strengths $B_{\mu\nu}$ and $\mathbf{W}_{\mu\nu}$ are defined as

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \tag{5}$$

$$\mathbf{W}_{\mu\nu} = \partial_\mu \mathbf{W}_\nu - \partial_\nu \mathbf{W}_\mu - g\mathbf{W}_\mu \times \mathbf{W}_\nu \,, \tag{6}$$

which leaves us with the Electroweak Lagrangian:

$$\mathcal{L}_{EW} = \sum_k i\bar{\psi}_k \gamma_\mu D^\mu \psi_k - \frac{1}{4}\mathbf{W}_{\mu\nu}\mathbf{W}^{\nu\mu} - \frac{1}{4}B_{\mu\nu}B^{\nu\mu} \,. \tag{7}$$

where the index k runs over the three generations of particles.

QCD Lagrangian:
The strong interaction is described by the $SU(3)$ group. The gauge fields are the gluons G^a_μ $(a = 1, ..., 8)$. The field strength is given via

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g f_{abc} G^b_\mu G^c_\nu \,, \tag{8}$$

where g_s is the strong coupling constant and f_{abc} are structure constants. With this the QCD Lagrangian can be written as follows:

$$\mathcal{L}_{QCD} = \sum_k i\bar{q}_k \gamma_\mu (D^\mu - i g_s G^\mu_a \frac{\lambda^a}{2}) q_k - \frac{1}{4} G^a_{\mu\nu} G^{\mu\nu}_a \,. \tag{9}$$

Again the index k runs over all three generations of particles. The $\lambda_a/2$ are the generators of the group. The term with the contravariant derivative describes the kinematic properties of the quarks and gluons and the interaction of quarks with the vector bosons. The last term is responsible for the self-interactions of gluons.

2.3 The Higgs Mechanism

Introducing a static mass term in the standard model Lagrangian would break gauge invariance of the $SU(3) \times SU(2) \times U(1)$ group. One possible solution is the spontaneous breaking of the symmetry by the ground state. To generate masses dynamically, one possible choice is to add a complex scalar doublet, the Higgs field [8] [9]

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} , \tag{10}$$

with a potential of

$$V\left(\phi^\dagger \phi\right) = \mu^2 \phi^\dagger \phi - \lambda \left(\phi^\dagger \phi\right)^2 . \tag{11}$$

For $\mu^2 < 0$ and $\lambda > 0$, the minimum of the potential has a finite value at $\phi\,\phi^\dagger = -\mu^2/2\,\lambda$. The choice of

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} , \tag{12}$$

where $v = \sqrt{-\mu^2/\lambda}$, is equivalent to spontaneously breaking the symmetry. After expanding around this minimum one is left with four real fields. Under a gauge

transformation three of them devolve to a third spin component of the vector bosons ($W^{\pm}$, Z^0). *"This process, in which a vector field eats a scalar one in the process of becoming massive, is known as the Higgs mechanism"* [9]. The component left is a scalar field. An excitation of this field would constitute a Higgs boson. It couples to each particle with a strength proportional to its mass [9].

Higgs Lagrangian:
Last but not least, the Higgs part of the Lagrangian can be written down.

$$\mathcal{L}_{Higgs} = |D_\mu \phi|^2 - V(\phi) - \sum_j c_j \bar{q}_j \phi_j q_j - \sum_k f_k \bar{\psi}_k \phi \psi_k \tag{13}$$

The indices j and k run over the three generations of particles. $V(\phi)$ as stated above is the potential of the Higgs field. The kinematic properties of the Higgs field and the coupling to the electroweak vector bosons are described by the covariant derivative. The last two terms show the coupling of the quarks and leptons.

2.4 Higgs Boson Production and Decay

Figure 1 shows the cross section of the five possible Higgs production mechanisms. These are the gluon-gluon fusion, the vector boson fusion, the associated W/Z boson production and the associated t quark production. Their respective Feynman graphs are depicted in figure 2. The production studied in this thesis is the vector boson fusion. Here, a Higgs boson is produced via two vector bosons. In addition, there are two jets which show a high separation in pseudorapidity (see figure 4). These tag jets are the distinguishing feature of this production.
Figure 3 shows a Higgs boson decaying into a W^+W^- pair. Each of the vector bosons then decays into a pair of charged lepton and neutrino. Since the Higgs boson has spin zero, the W bosons will have opposite spins of plus or minus one. Hence, both particles of the charged lepton neutrino pairs must have spin $+1/2$ or $-1/2$. The weak force only couples to left handed particles and right handed antiparticles. From this follows that in a projection onto the transverse plane the opening angle between the charged leptons should be small, the same applies for the neutrinos (see figure 4). As neutrinos can only be measured indirectly via the missing transverse energy, it can be deducted that the polar angle between the dilepton system and the missing transverse energy should be rather large. For a Higgs boson of a mass $m_{\mathrm{H}} = 125\,\mathrm{GeV}$, the cross section at a center of mass energy of $8\,\mathrm{TeV}$ is $\sigma = 1.578 \pm 0.047\,\mathrm{pb}$ if produced via the vector boson fusion [11]. This is about a factor of ten smaller than the cross section for the gluon-gluon fusion.

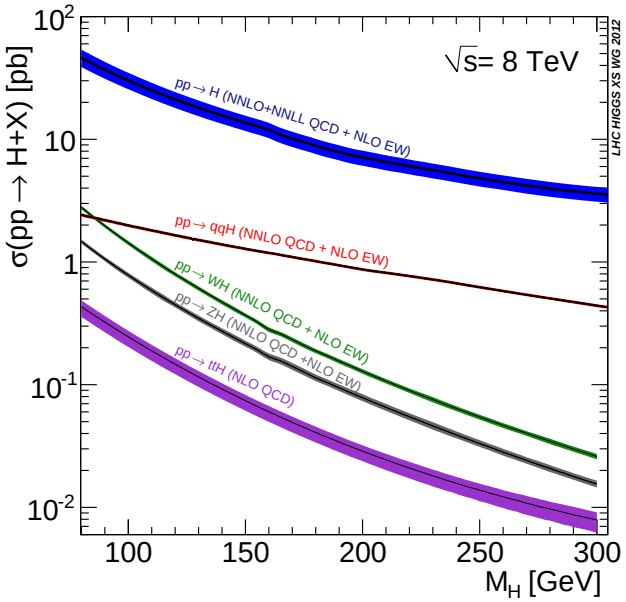


Figure 1: Higgs boson production cross section at a center of mass energy $E_{cm} = 8\,\text{TeV}$ for gluon-gluon fusion (blue), vector boson fusion (red), associated W (green), associated Z (gray), and associate t (purple) [10].

The branching ratio for the $H \to W^+W^-$ channel for the same Higgs mass as above is BR $= (216 \pm 9) \cdot 10^{-3}$ [12]. The branching ratio for $W^+W^- \to \bar{l}\nu l\bar{\nu}$ is BR $= (45.5 \pm 0.6) \cdot 10^{-3}$ [5]. This results in an overall branching ratio of BR $= (9.8 \pm 0.4) \cdot 10^{-3}$ for $H \to W^+W^- \to \bar{l}\nu l\bar{\nu}$. The l either stands for an electron or a muon, and the ν for its corresponding neutrino.
The analysis is performed in six channels: three opposite flavour channels and three same flavour channels. The opposite flavour channel is split into the $e\mu$ (electron muon) and the μe (muon electron) channel. Here, the first lepton is the one with the higher p_T. Without splitting, it is called the $e\mu\mu e$ channel. The same flavour channel is split with respect to flavour into the ee (electron electron) and the $\mu\mu$ (muon muon) channel. Again, these two can be combined to the $ee\mu\mu$ channel.
The main backgrounds for the vector boson fusion are $t\bar{t}$ in the opposite and Z + jets in the same flavour channel. A detailed description of the background processes will follow in section 4.1.

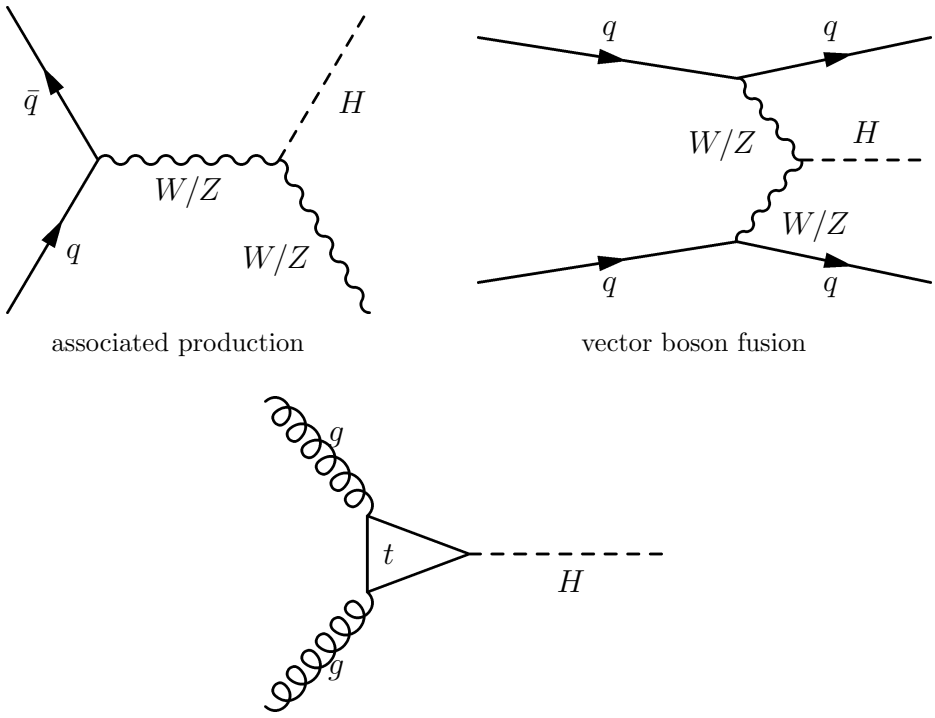


Figure 2: Feynman diagrams for the three most important Higgs boson production processes.

2.5 Boosted Decision Trees

To identify a new particle, a separation between background and signal is needed. Classically, this is done by cuts on variables (cut-based analysis) to obtain a signal-dominated region. In a BDT, the best cut value and variables are determined by an algorithm.

2.5.1 Mode of Operation

A **decision tree** as seen in figure 5 is a binary tree structured classifier [13]. To apply a BDT, the Monte-Carlo samples are divided into equally large training and testing samples. This can be done randomly or by even/odd separation. The tree is grown by repeated true/false decisions onto a single variable until a stop criterion is fulfilled. Possible stop criteria are: `nEventsMin`, the number of events in the final leave node, `NNodesMax`, the number of nodes, and `MaxDepth`, the maximum depth of

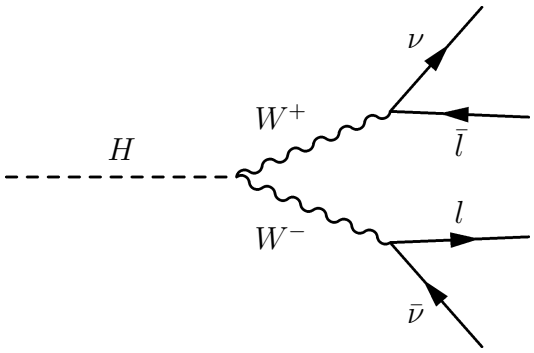


Figure 3: The Higgs boson decaying into a W^+W^- pair and further into two charged lepton neutrino pairs.

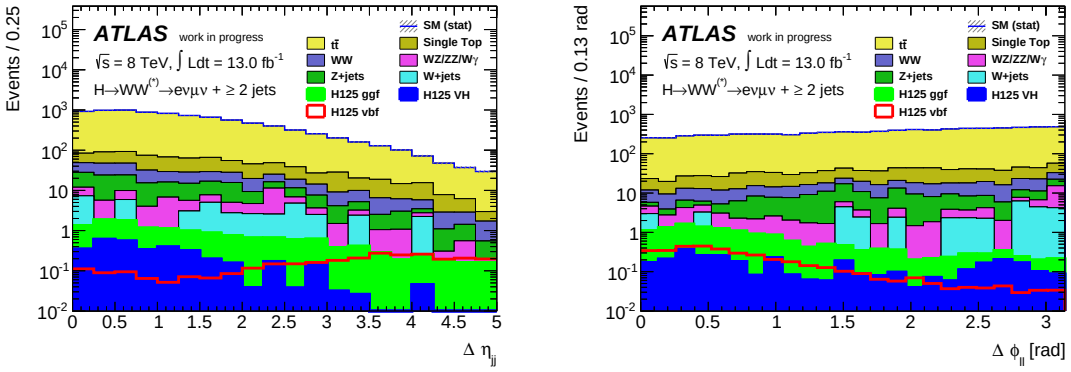


Figure 4: MC simulated distribution of $\Delta\eta_{jj}$ (left) and $\Delta\phi_{\ell\ell}$ (right) in logarithmic scale.

the tree. The training sample is split into many leaves where the split for each node is determined by finding the cut values and variable that gives the best signal over background separation. The number of cuts onto one variable and the separation type is configurable over the options `nCuts` and `SeparationType`. Depending on the majority of events in the final node, it is classified as signal or background like. After the training, the BDT is applied to the testing sample to obtain the BDT response a value between -1 (background-like) and +1 (signal-like).

Boosted Decision Trees: Many trees are consecutively grown from the same training sample but for every new tree the misclassified events are reweighted. The final classifier is obtained by averaging over all trees. The number of trees is adjusted via `NTrees`. The advantage of a BDT is that it stabilizes with respect to fluctuations.

The **boost method** used is gradient boost. A decision tree is described by a function $f(x, \alpha_m)$. The model response $F(x, \beta_m, \alpha_m)$ is a weighted sum (weights β_m) over the base functions $f(x, \alpha_m)$. The goal is to minimize the loss function $L(F, y) = (F(x, \beta_m, \alpha_m) - y)^2$, the difference between model response and true value y, with respect to the parameters α_m and β_m. The name derives from the minimization, which is done by calculating the gradient of the loss function. The learning rate of the boost algorithm can be adjusted by the `Shrinkage` parameter.

Overtraining of a BDT occurs if too many parameters of an algorithm are adjusted to too few data points, for example if `nEventsMin` is set to one: Every data point in the training sample will be classified perfectly as signal or background but the statistically independent points in the testing sample will not fit into the same parameters.
A method called pruning is used to remove statistically insignificant branches. It will not be further described since it is not implemented for gradient boost. Here, pruning has to be done manually by restricting the number of nodes and/or the depth of the tree.

Bagging uses resampling to stabilize the response of a classifier. It is not a boosting method. Here the classifier is repeatedly trained using resampled events. The resampling is done by randomly picking events from a parent sample, where one event can be picked several times. Bagging can be combined with the gradient boost via the flag `UseBaggedGrad` and the fraction of events used in each resampling is set by `GradBaggingFraction`.

Negative event weights (`NegWeightTreatment`) can occur in next-to-leading order Monte Carlo generators. Since they are unphysical, they are chosen to be ignored.
An important quantity used in the description of BDTs is purity. It is defined as:

$$p = \frac{S}{S+B}\,. \tag{14}$$

A detailed list of the used parameters can be found in appendix A. Unexplained options are trivial and are specific of Toolkit for Multivariate Data Analysis (TMVA). They do not affect the BDT.

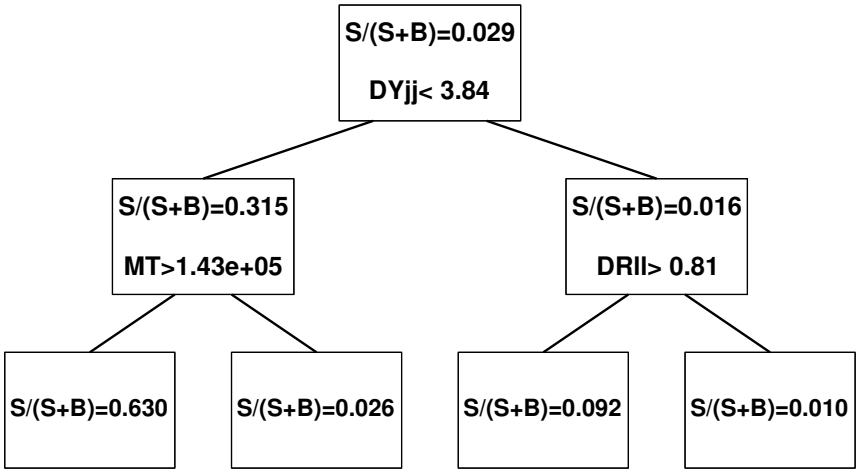


Figure 5: A BDT as obtained from the TMVA Gui [13]. The tree is grown from a sequence of splits onto variables. To obtain the highest purity $p = 0.630$, for example, the first split is onto $\Delta Y_{jj} < 3.84$ and then onto $m_T < 1.3 \cdot 10^5$. The stop criterion here is the depth of the tree, which is not to exceed two.

2.5.2 Evaluation of a BDT

In the following, it will be illustrated how to evaluate BDTs and display the results graphically [13].

Correlations between Input Variables
TMVA offers two possibilities to display correlations between variables. The first is a scatter plot and the second is done in form of a correlation matrix that contains the linear correlation coefficients (see figure 6). These plots are separated into the background and the signal samples.

Input Variable Distribution
In figure 6, one can see the normalized distribution of the input variables for signal and background. This is helpful to recognize how well a variable can be used to distinguish between background and signal.

Receiver Operating Characteristic (ROC) Curve
The ROC curve (figure 6) displays background rejection (probability of predicting signal if the true state is signal) versus signal efficiency (probability of predicting

background if the true state is background) [14]. It describes the efficiency of a BDT. An ideal ROC curve would be in the upper right corner. To get a numerical statement, the integral over this curve is taken. The closer it is to one, the more efficient the BDT.

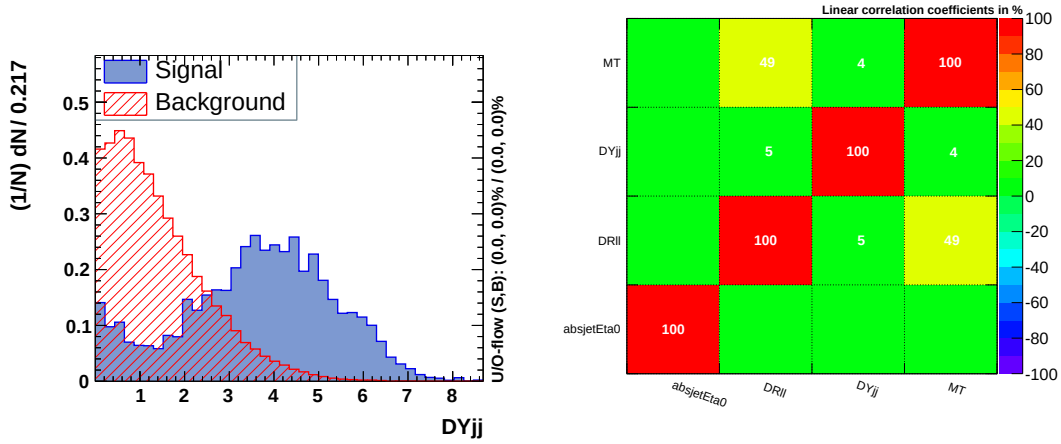
Overtraining Check
In order to determine how well a BDT can model the data, one has to check for overtraining. This is done by overlapping the classifier output distribution for testing and training sample (figure 6). A perfect overlap means no overtraining. The Kolmogorov-Smirnov test is used to classify the agreement between these two distributions. It returns the overtraining parameter, a number between zero and one. The closer it is to one, the better the agreement. The classifier output distribution is also called BDT response or BDT output.

Ranking of variables
TMVA offers a ranking of input variables. The occurrence of a variable used to split the sample is counted. This value is weighted by the number of events in the node and the gain in separation squared. However it breaks down in the case of correlated variables: If two variables are used that are about equally good in separation the algorithm will always pick the slightly better one and neglect the other one, which leads to a good ranking for one variable and a bad ranking for the other.

BDT response against variable
To evaluate how well a classifier works depending on the value of a variable, the BDT response is plotted against the variable in a two-dimensional histogram. The third axis displays the normalized number of events. This is done separately for all signal samples and all background samples. Thus, correlations can be visualized and regions in which a classifier works badly can be identified.

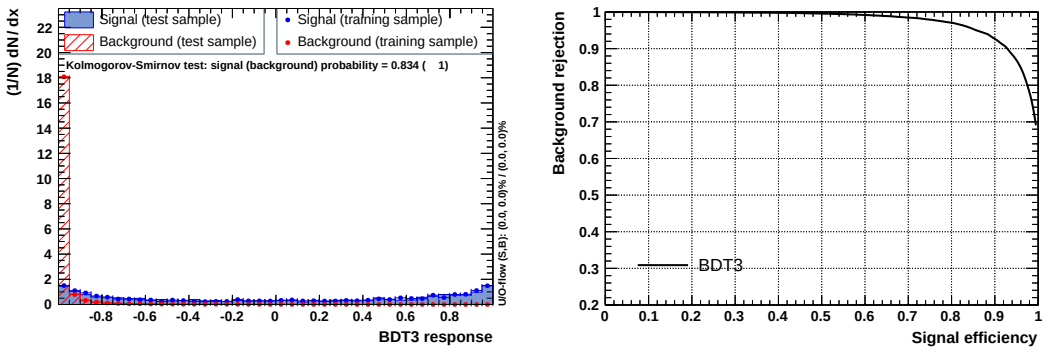
2.6 Event Shape Variables

Event shape observables describe the geometric properties of a particle collision. They quantify whether the energy flow is alongside one axis or distributed over the solid angle. In proton proton collisions, the transverse momenta are used to describe the shape of the event, since they are Lorentz-invariant under boosts along the beam axis. Usually, the event shape variables are calculated over the momenta of all outgoing particles but due to the finite detector acceptance at small angles and the inability to detect neutrinos only the sum over the available particles is used [15].



Input variable distribution

Correlation matrix



Overtraining check

ROC curve

Figure 6: BDT evaluation plots obtained via the TMVA Gui [13].

Transverse thrust: One event shape variable is the transverse thrust. It is defined as

$$T_T = \max_{|\hat{n}|=1} \frac{\sum_i |\vec{p}_{T,i} \cdot \hat{n}|}{\sum_i |\vec{p}_{T,i}|} \quad , \tag{15}$$

where the index i runs over all outgoing particles of the event and $\vec{p}_T$ is the transverse momentum of the particle with index i. The transverse thrust axis $\hat{n}_T$ is the unit vector $\hat{n}$ that maximizes equation 15. The transverse thrust ranges from $T_T = 1$ for a pencil-like event to $T_T = 2/\pi$ for an isotropic event.

Transverse thrust minor: The event plane is defined by the thrust axis $\hat{n}_T$ and the beam axis $\hat{z}$. The axis perpendicular to those is defined as the minor axis $\hat{n}_m = \hat{n}_T \times \hat{z}$ (see figure 7). It is calculated via

$$M_T = \frac{\sum_i |\vec{p}_{T,i} \cdot \hat{n}_m|}{\sum_i |\vec{p}_{T,i}|} \quad . \tag{16}$$

The transverse thrust minor has values from $M_T = 0$ for a pencil-like event to $M_T = 2/\pi$ for an isotropic event.

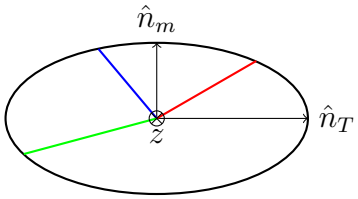


Figure 7: Illustration of the transverse thrust and minor. The colored lines represent outgoing particles.

Transverse sphericity: Another event shape variable is the sphericity. It is defined by using the sphericity tensor:

$$S^{\alpha\beta} = \frac{\sum_i p_i^\alpha \, p_i^\beta}{\sum_i |\vec{p}_i|^2} \, ; \quad \alpha, \beta = x, \, y, \, z \quad , \tag{17}$$

where p^α is the α'th component of the momentum. In case of the transverse sphericity, the z direction is omitted and the momentum $\vec{p}$ transfers to $\vec{p}_T$. The transverse sphericity ranges from $S_T = 0$ (isotropic event) to $S_T = 1$ (pencil-like event). It is obtained by

$$S_T = \frac{2\lambda_2}{\lambda_1 + \lambda_2} \, , \tag{18}$$

where $\lambda_{1/2}$ are the eigenvalues of the sphericity tensor.

3 Experimental Setup

This chapter describes the experimental setup. First, a brief summary of the LHC is given. As a second step the composition of the A Toroidal LHC Aparatus (ATLAS) detector is shortly outlined. Furthermore, a short section about triggers is included and both the object reconstruction and selection are outlined. In the end, the process and simulation of events will also be shortly described.

3.1 The Large Hadron Collider

The LHC near Geneva Switzerland is a two ring superconducting accelerator with a total circumference of 27 km. It reuses the former large electron positron (LEP) tunnel at a depth of 100 m. After a chain of pre-accelerators, the protons are injected into the LHC at an energy of 450 GeV, and can be further accelerated up to 7 TeV. For the 2012/13, run the energy was 4 TeV. The protons are accelerated in counterpropagating bunches with a bunch spacing of 50 ns, which is about twice the designed spacing. The two proton beams are kept on track by 1232 cryostatic dipole magnets cooled to a temperature of 1.9 K.
The LHC has four collision points. One of them is located at the ATLAS detector, the three others at Compact Muon Solenoid (CMS), A Large Ion Collider Experiment (ALICE), and Large Hadron Collider beauty (LHCb) [16] [17].

3.2 The ATLAS Detector

The ATLAS detector is composed of four major parts as seen in figure 8. These are beginning from the center around the interaction point: the inner detector, the electromagnetic calorimeter, the hadronic calorimeter, and the muon spectrometer.

3.2.1 The Inner Detector

The inner detector provides tracking measurements by combining precision trackers at small radii with broader resolution trackers at larger radii. Its goal is to achieve momentum and vertex resolution.
Each detector consists of concentric cylinders around the beam axis in the barrel region and disks perpendicular to the beam axis in the end-cap region. The inner most part is the pixel detector. It is surrounded by the semiconductor tracker (SCT) followed by the transition radiation tracker (TRT).
The whole inner detector is immersed in a two Tesla magnetic field generated by

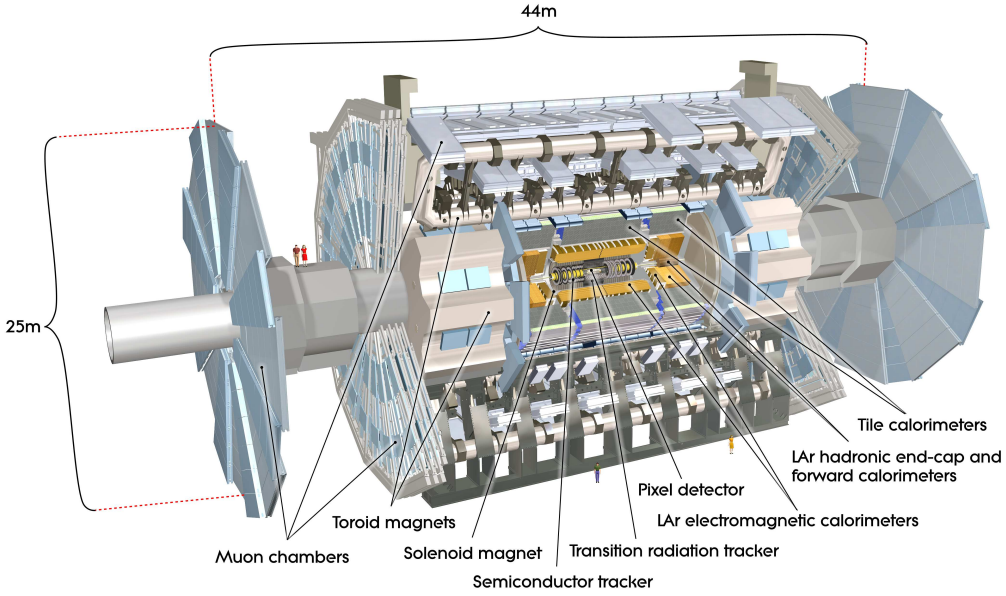


Figure 8: Overview of the ATLAS detector [18].

the central solenoid magnets. Its purpose is to deflect charged particles in order to measure their momentum and charge.

3.2.2 Electromagnetic Calorimeter

The electromagnetic calorimeter is used for precision measurements of electromagnetically interacting particles like electrons and photons. It is divided comparable to the inner detector into barrel and end-cap region. The electromagnetic calorimeter is a liquid Argon (LAr) detector with accordion-shaped kapton electrodes. Lead absorber plates enclose it.

3.2.3 Hadronic Calorimeter

The purpose of the hadronic calorimeter is the measurement of missing transverse energy E_T^{miss} and jets. The tile calorimeter, which is located in the barrel region, uses scintillating tiles read out by wavelength shifting fibers that couple into photomultipliers. The end-cap region is covered by the LAr hadronic calorimeter (HEC). The LAr forward calorimeter (FCal) can also be found there consisting of three parts. The first part made of copper measures electromagnetic interactions. The other two parts made of tungsten measure the hadronic ones.

3.2.4 Muon Spectrometer

The muons are deflected by using toroid magnets in the barrel as well as in the end-cap region. There are four types of muon chambers used in the detector: monitored drift tubes (MDT) for most of the pseudorapidity range and cathode strip chambers (CSC) for large pseudorapidity range. The thin gap chambers (TGC) in the end-cap region and the resistive-plate chambers (RPC) in the barrel region are used as trigger chambers.

3.2.5 The ATLAS Coordinate System

Here a brief summary of the coordinate system (see figure 9) used to describe the ATLAS detector will be given. The origin of the coordinate system is the nominal interaction point. The z-axis is defined along the beam direction, the x-y plane transverse to it. The positive y-axis is defined as pointing upwards, the positive x-axis is defined as pointing to the center of the LHC ring. The azimuthal angle ϕ is measured around and the angle θ from the beam axis [18]. Instead of θ, the Lorentz invariant pseudorapidity

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \tag{19}$$

is used. Another important quantity is the pseudorapidity-azimuthal angle space or radius

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \quad . \tag{20}$$

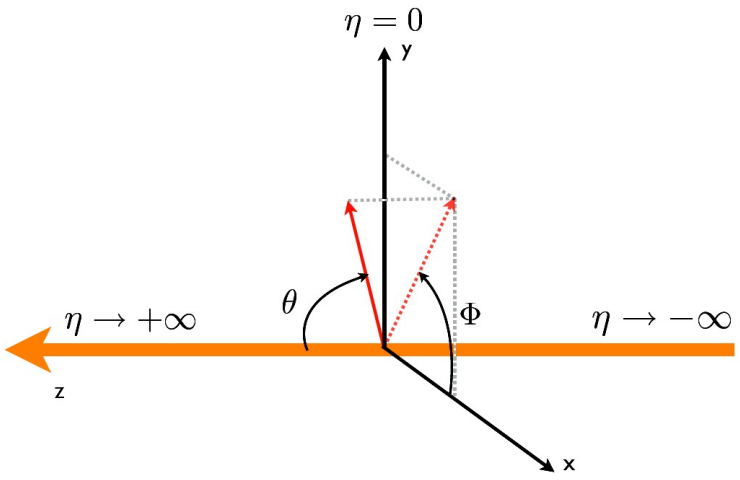


Figure 9: The ATLAS coordinate system [19].

3.3 Trigger

If the LHC runs at the designed specifications, it has a bunch crossing rate of 40 MHz with about 23 interactions per crossing. It is impossible to record this amount of data so a trigger system is used to filter the physically interesting events. The ATLAS trigger system consists of three triggers. The Level-1 trigger (L1), the Level-2 trigger (L2), and the event filter. The latter two form the High-Level trigger (HLT) [18].

The L1 is hardware-based and reduces the event rate from 40 MHz to about 100 kHz. With the information from the muon detector and the calorimeters, an initial event selection is performed. The L1 calorimeter trigger is able to identify $E_{\mathrm{T}}^{\mathrm{miss}}$ and high-E_{T} objects that could be electrons, photons, jets, or τ-leptons. The L1 muon trigger searches for high-p_{T} muons. Signatures passing the L1 trigger thresholds are defined as Regions-of-Interest (RoIs).

The software-based HLTs run on the server farm near the detector. The RoIs gathered by L1 are reevaluated by the L2, which has access to more detector information. The event rate is further reduced to about 4 kHz by this trigger. The event filter already reconstructs objects for its decisions. Events accepted by the event filter are stored, the final rate is 400 Hz.

In total, the trigger system reduces the amount of data by a factor of 10 000, which corresponds to roughly 1 GB/s.

3.4 Object Reconstruction and Selection

In this section it will be illustrated how the quantities measured in the ATLAS detector are used to describe physical objects [20]. The most important objects are electrons, muons, jets, here especially b-jets, and the missing transverse energy.

3.4.1 Electrons

Electrons are reconstructed by using clusters in the electromagnetic calorimeter representing a certain $\Delta\eta \times \Delta\phi$ area. The cluster has to contain an energy of at least 2.5 GeV and an associated track in the inner detector [21]. This works in the region of the track detectors $|\eta| < 2.5$. Electrons with larger pseudorapidity are reconstructed only with clusters that satisfy $E_T > 5$ GeV.

There are three electron selection categories: loose, medium, and tight. This selection is done by a multivariate cut-optimization and cuts onto certain variables. The tighter the criteria, the more variables are added to the multivariate method and the stricter are the cut values. The electrons used in this analysis are tight++. Electrons have to fulfill certain isolation requirements to distinguish them from particles misidentified as electrons or vice versa. Electrons are often misidentified as W+jets. To achieve this, the following isolation and vertex impact parameter cuts are applied:

- topoEtConeCor30/$p_T < 0.16$
- PtCone30/$p_T = \begin{cases} 0.12 & \text{if } 15\,\text{GeV} < p_T < 25\,\text{GeV} \\ 0.16 & \text{if } p_T > 25\,\text{GeV} \end{cases}$
- d_0 significance < 3.0
- $z_0 \sin\theta < 0.4\,\text{mm}$

Here d_0 significance is the transverse impact parameter divided by its error and $z_0 \sin\theta$ the projection of the longitudinal impact parameter onto the transverse plane. The variable used for track isolation is pTcone30. It is the sum of all transverse momenta within a cone of 0.3 around the electron. The same is done for the calorimeter isolation topoEtConeCor30, it is the sum of all cell energy deposits projected onto the transverse plane.

3.4.2 Muons

There are three different ways to reconstruct muons: standalone, combined, and tagging [22]. Since the muons used for this analysis are Staco combined, the method will be described shortly. Muons are reconstructed by extrapolating a track from the muon spectrometer to a matching track from the inner detector. This is done with the Staco (statistical combination) algorithm. For muons isolation requirements and cuts have to be fulfilled as well. A large source of misidentification for muons are jets faking them. The isolation and vertex impact parameter cuts are listed below:

- EtConeCor30/$p_T < 0.014\,p_T - 0.15$ and EtConeCor30/$p_T < 0.2$
- PtCone30/$p_T < 0.01\,p_T - 0.105$ and PtCone30/$p_T < 0.15$
- d_0 significance < 3.0
- $z_0 \sin\theta < 1.0\,\text{mm}$

The definitions for the variables and the electrons described above are the same.

3.4.3 Jets

Jets are reconstructed by using the anti-k_t algorithm [23]. Two definitions of distances are introduced. The distance between two objects d_{ij} and the distance between an object and beam d_{iB}.

$$d_{ij} = \min\left(k_{ti}^{2p}, k_{tj}^{2p}\right) \frac{(y_i - y_j)^2 - (\phi_i - \phi_j)^2}{R^2}, \tag{21a}$$

$$d_{iB} = k_{ti}^{2p}, \tag{21b}$$

where k_t is the transverse momentum, y the rapidity, and ϕ the azimuthal angle. The algorithm is called k_t if $p = 1$ and anti-k_t if $p = -1$. The algorithm tries to find the smallest distance between objects. In the case of this being d_{ij}, they are combined, in the case of this being d_{iB}, they are identified as a jet and are removed from the list. This procedure is repeated till no objects are left. In this analysis the cut parameter is 0.4.
To pass as a jet, certain criteria on the Jet Vertex Fraction (JVF) [1], jet p_T, and pseudorapidity have to be met. The applied cuts and criteria are as follows:

- Jet $p_T > \begin{cases} 25\,\text{GeV} & \text{if} \quad |\eta| <= 2.5 \\ 30\,\text{GeV} & \text{if} \quad |\eta| > 2.5 \end{cases}$
- $|\text{JVF}| > 0.5$

B-jets: The MV1 b-tagger algorithm uses a neural network to identify b-jets. Different working points and jet transverse momenta can be chosen. This is important to identify the b quarks produced in single top or $t\bar{t}$ processes.

3.4.4 Missing Transverse Energy

The only non detectable particle is the neutrino. Therefore the energy deposited in the calorimeters does not add up to zero. Missing transverse energy ($E_{\text{T}}^{\text{miss}}$) is calculated from the negative sum of calibrated cell energies for the x and y components respectively.
The corrected quantity used is $E_{\text{T,rel}}^{\text{miss}}$. It is defined as $E_{\text{T,rel}}^{\text{miss}} = E_{\text{T}}^{\text{miss}} \sin\Delta\phi_{\text{min}}$ with $\Delta\phi_{\text{min}} = \min(\Delta\phi, \pi/2)$, where $\Delta\phi$ is the azimuthal difference between the $E_{\text{T}}^{\text{miss}}$ vector and the nearest lepton or jet with $p_{\text{T}} > 25$GeV.

[1] The JVF is a measure that describes whether the jet is from a primary vertex.

3.5 Event Processing and Simulation

To compare the obtained results with theoretical predictions, events are simulate by using MC event generators. Data and MC samples must have the same format. To achieve this, the full ATLAS detector is modelled with Geant4 a simulation toolkit [24].
The used generators are listed below with a short description of their properties. Table 3 shows a list of signal and background processes, and the generators used to simulate them.

- **MC@NLO**
  MC@NLO is a next-to-leading order (NLO) generator calculating QCD processes using a parton shower simulation. The top background is modelled with this generator [25].

- **Alpgen**
  Alpgen calculates matrix elements in leading order for QCD and electroweak processes. Leptonically decaying vector bosons that were produced with N partons are calculated with this generator [26].

- **AcerMC**
  The single top production in the t-channel is calculated with this generator [27].

- **gg2WW** and **gg2ZZ**
  Dileptonically decaying vector bosons that are produced via gluons are an important background for the Higgs search. They are simulated with these generators [28].

- **Pythia8**
  Hadronisation and parton showers are simulated with this generator. The associated Higgs production uses this generator for modelling [29].

- **Powheg**
  Powheg combines NLO calculations with parton shower generators. It produces only positive event weights and is independent of the MC generator used for showering. The dibosonic fully leptonically decaying samples where generated by using Powheg [30].

process	generator	$\sigma \cdot \mathrm{Br}\,[\mathrm{pb}]$
ggF	PowHeg+Pythia8	0.44
VBF	PowHeg+Pythia8	0.036
WH/ZH	Pythia8	0.13
WW	Powheg	5.68
$gg \to WW$	gg2WW	0.16
ZZ	Powheg	1.2
$gg \to ZZ$	gg2ZZ	0.003
$W+$ jets	Alpgen	30824.6
$Z+$ jets	Alpgen	13766.6
$W\gamma$	Alpgen	320.5
$W(Z/\gamma^*)$	Powheg	19.8
$t\bar{t}$	MC@NLO	238.1
single top	MC@NLO AcerMC	52.6

Table 3: Used MC generators for signal and background processes with corresponding cross sections times branching ratio.

4 Search for the Higgs Boson

The aim of this section is to optimize parameters of the BDT, compare this analysis with the cut based analysis, and find (new) variables for the distinction between background and signal. Furthermore the different analysis channels will be compared.

4.1 Processes

In this section, a closer look at the background processes will be taken. Their key features and signatures are described and it will become clear how to distinguish them from the signal.

4.1.1 Signal Processes

Since this analysis focuses on the coupling of a Higgs boson to vector bosons, only the vector boson fusion is used as signal process. All other formerly discussed Higgs production processes (see section 2.4) will be treated as background.

4.1.2 Background Processes

The most important background processes which show a similar structure to the vector boson fusion are illustrated in figure 10.

- The $\mathbf{t\bar{t}}$ production is the largest and most difficult background to deal with. The structure is very similar to the vector boson fusion process because of the two emitted b-jets and the W bosons, where each boson decays into a pair of charged lepton and neutrino.

- Two jets along with a charged lepton and a neutrino are produced by the **single top** background. This process is distinguished from the signal because it only has one charged lepton.

- The **Drell-Yan** process plays an important role in the same flavour channel since the Z boson only decays into same flavour leptons. The lack of neutrinos is used to suppress this background. Moreover the two leptons are emitted

back to back in contrast to the signal.

- The **Z + jets** background produces jets and two charged leptons in the final state. The distinction criterion between background and signal will - as in the Drell-Yan background - be the missing neutrinos.

- The **W + jets** background as stated in the name contains at least two jets but only one pair of charged lepton and neutrino as opposed to the signal.

- Another background producing two charged leptons and two neutrinos is the **WW** production.

- The **ZZ** production has two charged leptons and two neutrinos in the final state but they are not produced in pairs of charged lepton and neutrino. So, the process differs from the signal by two antiparallel emitted leptons.

- In the **WZ** background, there are a Z boson decaying into two charged leptons and a W boson decaying into a pair of charged lepton and neutrino. The difference to the signal are the three charged leptons.

- The **associated production** with a vector boson as already stated produces either a W or a Z boson in association with a Higgs boson.

- In the **gluon gluon fusion**, only a Higgs boson is produced. Thus, the distinction criteria here are the two tag jets.

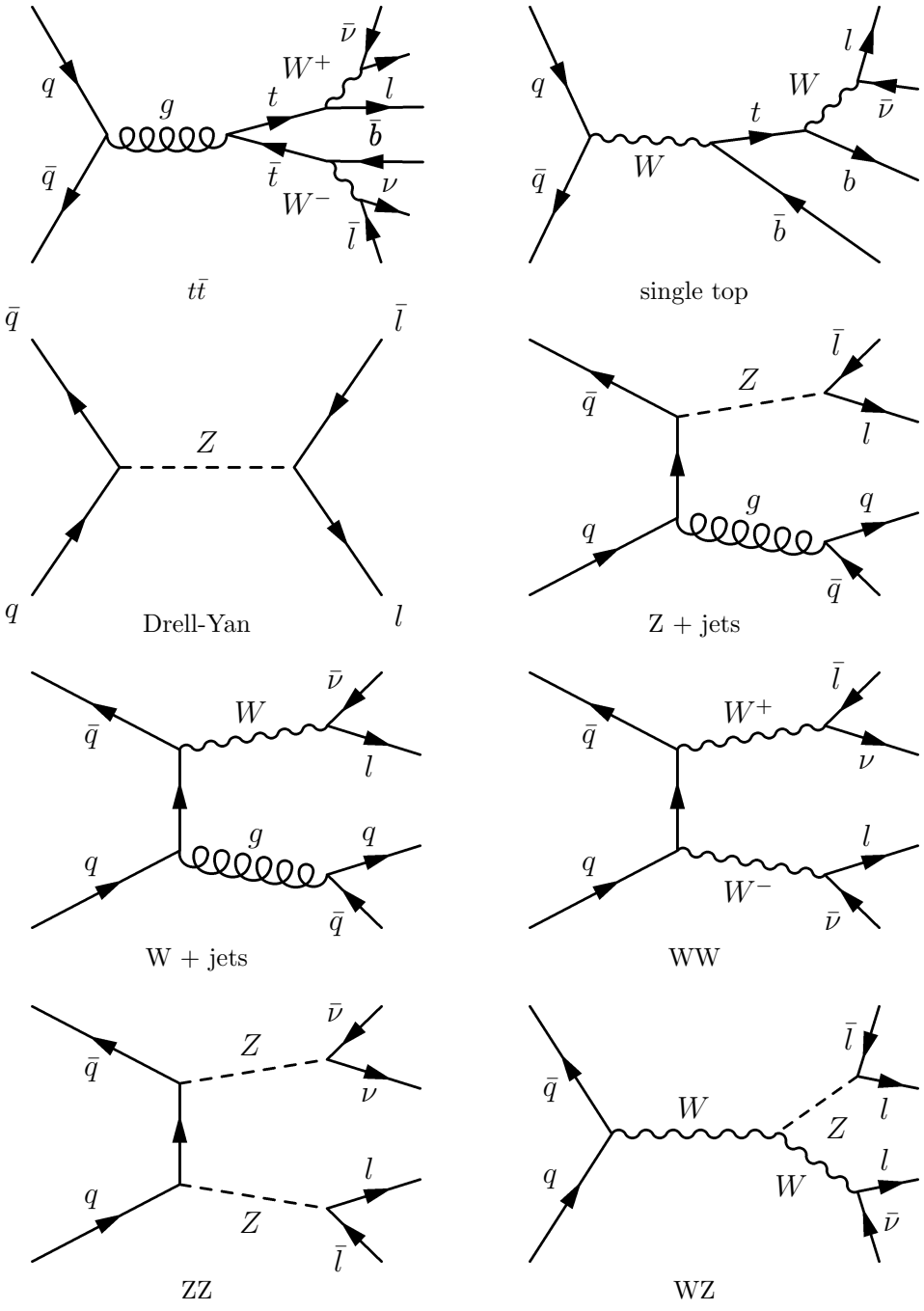


Figure 10: Feynman diagrams of some of the most important background processes.

4.2 Event Selection

In order to filter the previously discussed background from the signal processes, selection criteria on the physical variables (so called cuts) will be used to reduce certain backgrounds [31]. These cuts are illustrated in this section.

- **At least two jets cut**
  In the vector-boson-fusion the Higgs boson is produced with two jets so at least two jets are required.

- **Opposite sign lepton cut**
  Since there are a charged lepton and a charged antilepton in the final state, two oppositely charged leptons are required.

- **Lepton $\mathbf{p_T}$ cut**
  The transverse momentum p_T of the leading charged lepton has to be larger than 25 GeV. The p_T of the subleading charged lepton should be at least 15 GeV. This is done to exclude misidentified leptons.

- **$\mathbf{m_{ll}}$ cut**
  To suppress Υ resonances, the dilepton invariant mass is required to be more than 12 GeV in the same flavour channel and more than 10 GeV in the opposite flavour channel. This cut is necessary since there are model dependent uncertainties on the MC which describes these resonances.

- **Z veto**
  Drell-Yan production of $Z/\gamma^\star$ is reduced in the same flavour channel by cutting out the Z-peak. It is required that the dilepton invariant mass differs from the Z mass by at least 15 GeV.

- **$\mathbf{E^{miss}_{T,rel}}$ cut**
  Since the two charged leptons are accompanied by two neutrinos, a minimal $E^{\text{miss}}_{\text{T,rel}}$ is required. It has to be more than 25 GeV in the opposite and more than 45 GeV in the same flavour channel.

4.3 Cutflow and Event Yield

This section gives a short overview of the event yields. The dataset used has an integrated luminosity of $\mathcal{L} = 12.9\,\mathrm{fb}^{-1}$ and was obtained at a center of mass energy of $\sqrt{s} = 8\,\mathrm{TeV}$. A more detailed cutflow for all channels can be found in appendix B. As seen in tables 4, 5, 6, and 7, the dominating background is $t\bar{t}$ with about 90% of the total background in the opposite flavour channel and about 80% in the same flavour channel. After the last preselection cut, most events are expected for the $e\mu$ channel with 8835 ± 41, followed by the μe channel with 7695 ± 38. For the $\mu\mu$ channel, 5458 ± 40 events are expected, and the least are expected for the ee channel with 3504 ± 31. There is overall reasonable data over MC ratio agreement with values varying ± 0.04 around one. The significance ranges from 0.03 for the same flavour channel to 0.04 for the opposite flavour channel.

	VBF 125	$t\bar{t}$	Total Bkg.	Observed
at least 2 jets	9.2 ± 0.1	14954.1 ± 50.5	17441.0 ± 61.0	18181
lepton $p_{\mathrm{T}} > 25, 15\,\mathrm{GeV}$	8.0 ± 0.1	14132.9 ± 49.1	16325.9 ± 58.7	17067
OS leptons	8.0 ± 0.1	14094.7 ± 49.1	16092.6 ± 56.4	16701
$m_{\ell\ell} > 12, 10$ GeV	7.9 ± 0.1	14085.4 ± 49.1	16080.1 ± 56.4	16688
$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45, 25$ GeV	3.9 ± 0.1	7956.8 ± 36.9	8834.6 ± 40.9	9150

Table 4: Cutflow for the $e\mu$ channel. Only statistical uncertainties are given.

	VBF 125	$t\bar{t}$	Total Bkg.	Observed
at least 2 jets	7.7 ± 0.1	13089.8 ± 47.2	15307.2 ± 66.4	15611
lepton $p_{\mathrm{T}} > 25, 15\,\mathrm{GeV}$	6.8 ± 0.1	12442.4 ± 46.0	14377.9 ± 64.3	14681
OS leptons	6.8 ± 0.1	12421.5 ± 45.9	14159.4 ± 52.5	14385
$m_{\ell\ell} > 12, 10$ GeV	6.7 ± 0.1	12410.9 ± 45.9	14143.7 ± 52.4	14373
$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45, 25$ GeV	3.3 ± 0.1	6968.3 ± 34.4	7695.4 ± 38.2	7814

Table 5: Cutflow for the μe channel. Only statistical uncertainties are given.

	VBF 125	$t\bar{t}$	Total Bkg.	Observed
at least 2 jets	6.9 ± 0.1	11338.3 ± 43.5	209430.9 ± 753.0	215016
lepton $p_{\mathrm{T}} > 25, 15\,\mathrm{GeV}$	6.0 ± 0.1	10707.3 ± 42.3	198944.4 ± 740.3	203216
OS leptons	6.0 ± 0.1	10662.1 ± 42.2	197718.5 ± 738.2	202033
$m_{\ell\ell} > 12, 10$ GeV	5.9 ± 0.1	10631.8 ± 42.1	196942.1 ± 738.2	199795
Z veto (for $ee, \mu\mu$)	5.8 ± 0.1	8340.1 ± 37.4	31993.9 ± 200.8	33343
$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45, 25$ GeV	1.6 ± 0.0	2943.7 ± 22.2	3504.0 ± 30.9	3470

Table 6: Cutflow for the ee channel. Only statistical uncertainties are given.

	VBF 125	$t\bar{t}$	Total Bkg.	Observed
at least 2 jets	10.5 ± 0.1	16805.9 ± 54.0	342250.7 ± 977.4	342328
lepton $p_{\mathrm{T}} > 25, 15\,\mathrm{GeV}$	9.3 ± 0.1	15997.7 ± 52.7	327847.3 ± 963.2	327552
OS leptons	9.3 ± 0.1	15990.5 ± 52.7	327774.6 ± 963.2	327321
$m_{\ell\ell} > 12, 10$ GeV	9.1 ± 0.1	15945.6 ± 52.6	326355.4 ± 963.1	322871
Z veto (for $ee, \mu\mu$)	9.0 ± 0.1	12441.5 ± 46.5	53712.8 ± 266.8	54224
$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45, 25$ GeV	2.4 ± 0.1	4559.7 ± 28.2	5457.9 ± 40.1	5342

Table 7: Cutflow for the $\mu\mu$ channel. Only statistical uncertainties are given.

4.4 Control Regions

Since the MC generators cannot model reality perfectly, control regions are used to account for small discrepancies between data and MC-simulated data. Therefore, cuts are used to obtain a region dominated by the desired background process (i.e. $t\bar{t}$). Then the background but the desired background process is subtracted from the data and divided by the desired background process. The resulting number is called normalization factor. It is listed for the various channels in table 8. The error is calculated from the statistical error of all quantities by using gaussian error propagation. Figure 11 and 12 show the dilepton invariant mass before and after applying the normalization factors.

	$e\mu$	μe	$e\mu\mu e$
Z+jets	-	-	-
$t\bar{t}$	1.08 ± 0.01	1.06 ± 0.02	1.07 ± 0.01

	ee	$\mu\mu$	$ee\mu\mu$
Z+jets	1.17 ± 0.03	1.17 ± 0.02	1.17 ± 0.02
$t\bar{t}$	1.04 ± 0.02	1.04 ± 0.02	1.04 ± 0.02

Table 8: Normalization factors obtained from the Z+jets and $t\bar{t}$ control regions in the different channels with statistical error.

4.4.1 $t\bar{t}$

The $t\bar{t}$ control region is obtained by adding the condition to have at least one b-tagged jet (`nJets_Pt25_MV1_85`) to the at least two jets cut. Figure 11 shows the distribution of the dilepton invariant mass before and after applying the normalization factors. It can be seen that the MC tends to underestimate the $t\bar{t}$ distribution.

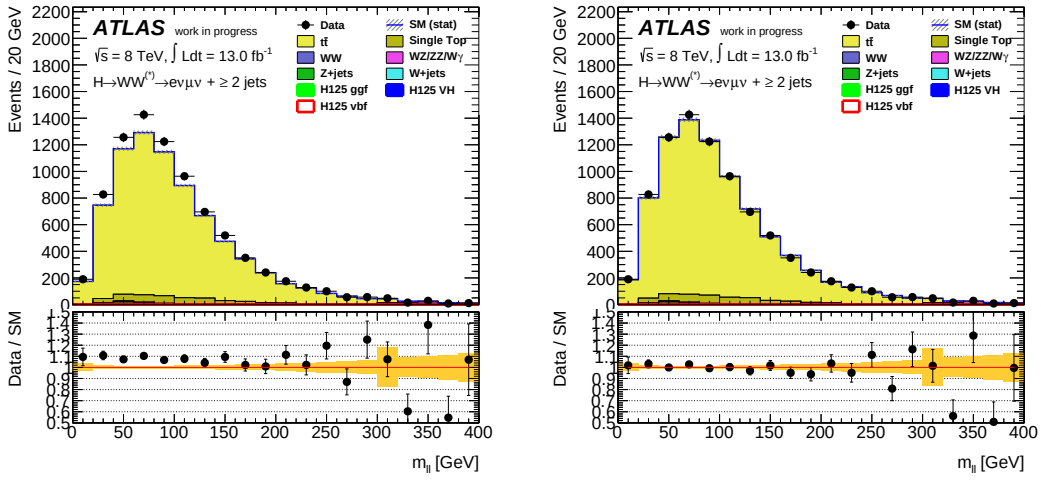


Figure 11: Dilepton invariant mass distribution in the $t\bar{t}$ control region before (left) and after (right) applying the normalization factors. Only the statistical errors are shown.

4.4.2 Z + jets

To obtain the Z + jets control region all cuts up to the dilepton invariant mass cut are used. The additional condition is that the absolute difference between the dilepton invariant mass and the Z mass is smaller than 15 GeV. Furthermore, the difference in rapidity between the two jets is required to be larger than 2.5 and the dijet invariant mass larger than 30 GeV. Of course, this control region is only used in the same flavour channels. Figure 12 shows the dilepton invariant mass distribution before and after applying the normalization factors. Like in the $t\bar{t}$ case, the MC tends to underestimate the Z + jets distribution.

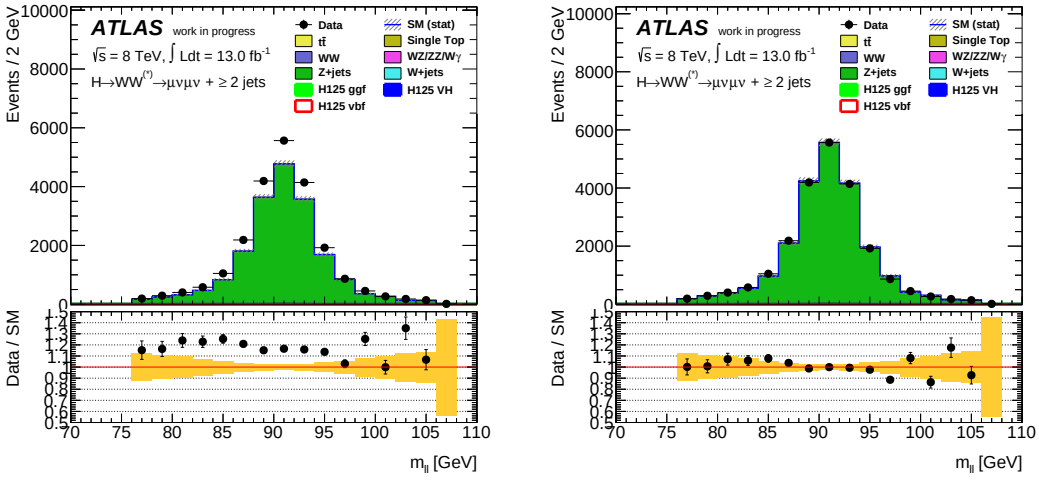


Figure 12: Dilepton invariant mass distribution in the Z+jets control region before (left) and after (right) applying the normalization factors. Only the statistical errors are shown.

4.5 Systematic Uncertainties

There are many sources of systematic uncertainties. They can be divided into experimental and theoretical categories. In this section, it will be explained which uncertainties have to be taken into account and how they are calculated.

The experimental uncertainties include scale and resolution of objects such as leptons, jets, and E_T^{miss}. Reconstruction efficiencies and the luminosity (3.6%) are also afflicted with uncertainties. The largest contribution here is the jet energy scale $(1-5\%)$, the jet energy resolution $(5-20\%)$, and the b-tagging efficiency $(5-12\%)$ [32].
Theoretical uncertainties are reflected in the cross section. They are influenced by the QCD renormalization the Parton Distribution Function (PDF) and the parton shower model.

channel	WW/top	Z+jets	W+jets	$WZ/ZZ/W\gamma$	signal
$e\mu$	0.13	0.14	0.41	0.10	0.091
μe	0.13	0.13	0.42	0.12	0.099
ee	0.13	0.24	0.4	0.13	0.14
$\mu\mu$	0.13	0.21	0.41	0.12	0.12

Table 9: Relative systematic uncertainties on the event yields separated by background and channel.

The calculation of systematic uncertainties is done as follows. The uncertainty considered is varied up and down by one standard deviation. These changes are propagated through the calculation. Later, the MC samples are produced including these changes. The uncertainties on the event yield are obtained by the difference in the yield between the nominal and the varied samples. The event yields considered are separated by process category (WW/top, Z+jets, W+jets, $WZ/ZZ/W\gamma$ and signal) and channel. To obtain the relative systematic uncertainty, the uncertainty on the event yield is divided by the nominal result.
For this analysis, the results of the $H \to WW \to \ell\nu\ell\nu$ working group calculated with the method described above were used. The relative systematic uncertainties on the event yields are displayed in table 9, separated by background and channel.

4.6 Baseline BDT

This section describes a baseline BDT using default parameters and basic standard variables. First, the variables used are explained and control plots are shown to ensure the correct modelling and data over MC ratio. Then, the results of the baseline BDT are explained and discussed. The only stop criterion for the BDT is the minimum number of events to be 200. The procedure will be shown explicitly for the combined opposite flavour channel $e\mu\mu e$. For the other channels, only the results will be shown in section 4.7 with the same procedure applied.

4.6.1 Variables Used

The 12 basic variables used in the training and application of the BDT are as follows:

- The transverse mass $m_T = \sqrt{\left(E_T^{\ell\ell} + E_T^{miss}\right)^2 - |\mathbf{p}_T^{\ell\ell} + \mathbf{p}_T^{miss}|^2}$
- The missing transverse energy $E_{T,rel}^{miss}$
- The opening angle between the two charged leptons $\Delta\phi_{\ell\ell}$
- The difference in pseudorapidity between the two charged leptons $\Delta\eta_{\ell\ell}$
- Difference in radius between the two charged leptons $\Delta R = \sqrt{\Delta\eta_{\ell\ell}^2 + \Delta\phi_{\ell\ell}^2}$
- Dilepton invariant mass $m_{\ell\ell}$
- Dilepton transverse momentum $p_{T,\ell\ell}$
- Opening angle between leading and subleading jet $\Delta\phi_{jj}$
- Difference in pseudorapidity between leading and subleading jet $\Delta\eta_{jj}$
- Difference in radius between leading and subleading jet $\Delta R_{jj} = \sqrt{\Delta\eta_{jj}^2 + \Delta\phi_{jj}^2}$
- Dijet invariant mass m_{jj}
- Dijet transverse momentum $p_{T,jj}$

4.6.2 Modelling of the Input Variables

In order for the BDT to classify the data correctly, it is important that the input variables are correctly modelled and the agreement between data and MC is good. In the following, some selected input variables are discussed. All of the input variables are shown in appendix C.
All input variables have good data over MC agreement. One of the many variables that show good data over MC agreement in all channels is for example the transverse mass m_T as shown in figure 14. The only mismodelling can be seen in the same flavour channel for the event shape variables (see figure 13). It looks as if the data is shifted relative to the MC. One reason for this could be, that the required missing transverse energy in the same flavour channel is about twice as much as in the opposite flavour channel. This is because of worse missing transverse energy resolution in the same flavour channel due to pile-up effects. As previously mentioned missing transverse energy is not included in the calculation of the event shape variables.

4.6.3 Results of the Baseline BDT

In this section, the results of the baseline BDT are shown and discussed. In figure 15, the BDT response plots are shown both in linear and logarithmic scale. In the lower part, the data over MC ratio or the significance $S/\sqrt{B}$ is displayed. One sees a reasonable agreement of data over MC. The significance increases from the first to the last bin of the BDT response.
In table 10, values for the three control variables (signal and background overtraining parameter and ROC integral) and the significance are shown: The background overtraining parameter is excellent at 1.0, the one for signal is acceptable at 0.4. The ROC integral is at 0.9628 and the significance has a value of 0.4.

To evaluate the performance of a classifier depending on the value of a variable, the BDT response is plotted against a variable for the signal or the background sample into a two dimensional histogram. Figure 16 shows these plots for $\Delta\eta_{jj}$ and $\Delta\eta_{ll}$. For $\Delta\eta_{jj}$, it can be seen that in the background sample almost all events are classified as background-like and as expected the number of events is higher for small values of $\Delta\eta_{jj}$. The signal sample shows events misclassified as background like for values of $\Delta\eta_{jj} < 3$, then an almost linear dependence of the classifier for values of $\Delta\eta_{jj}$ between 3 and 4, and then correct classification of events with $\Delta\eta_{jj} > 4$.
In the same figure, these plots are shown for $\Delta\eta_{ll}$. Again, the background is classified correctly. In the signal sample plot, of course, the same peak of misclassification occurs but no dependence on the value of the variable is visible. This demonstrates

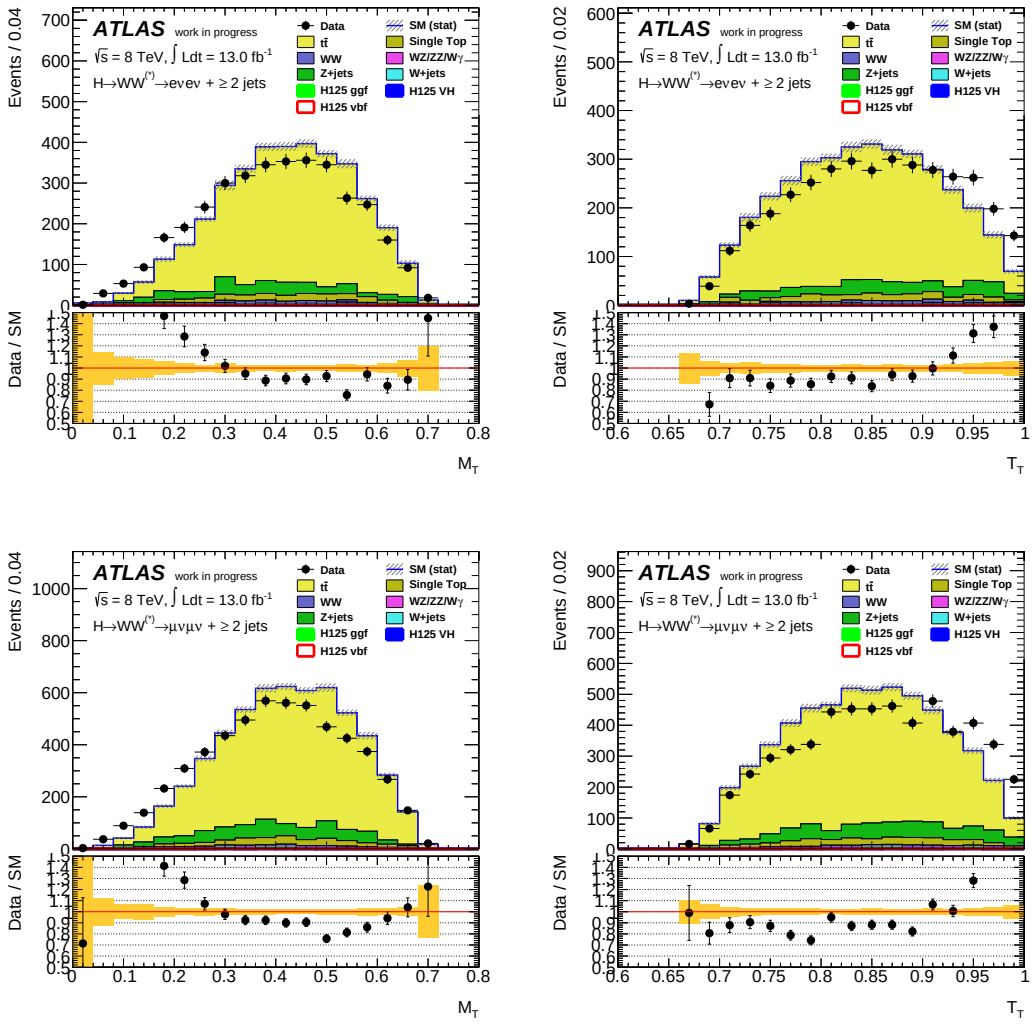


Figure 13: Transverse thrust and minor in the same flavour channel. Only statistical uncertainties are shown.

that variables are needed, which are able to classify signal events for small $\Delta\eta_{jj}$ regions.

The main focus will now be to remove or suppress the misclassification of the signal. While at the same time improving the performance in significance, the ROC integral, the signal overtraining parameter, and the data over MC ratio.

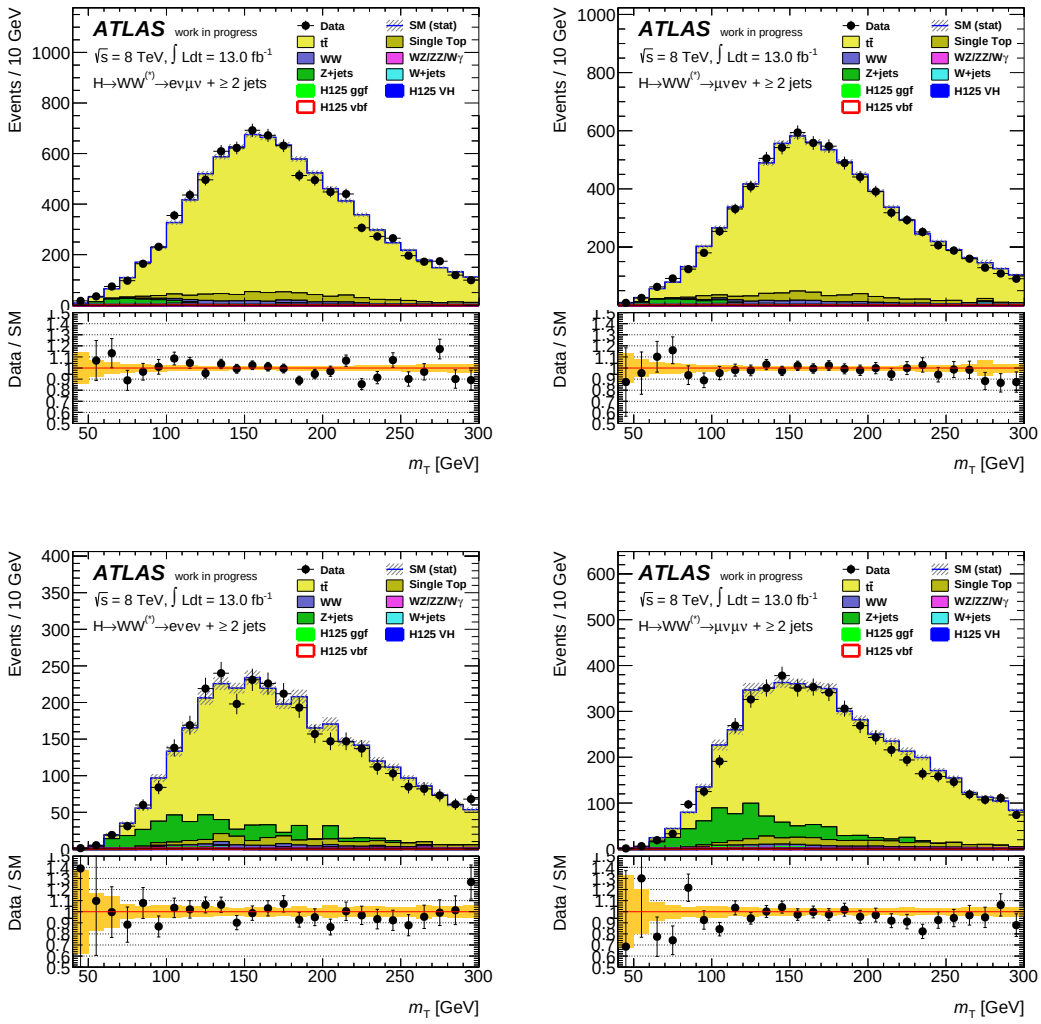


Figure 14: Transverse mass m_T for all four channels. Only statistical uncertainties are shown.

4.6.4 Determining Variables

In this section, improvements of the BDT performance are described. In each run variables will be successively added and the plots for the BDT response against $\Delta\eta_{jj}$ will be used to evaluate the classifier, with the values of the ROC integral, significance, and overtraining parameters in mind. The results are shown in figure 21 and

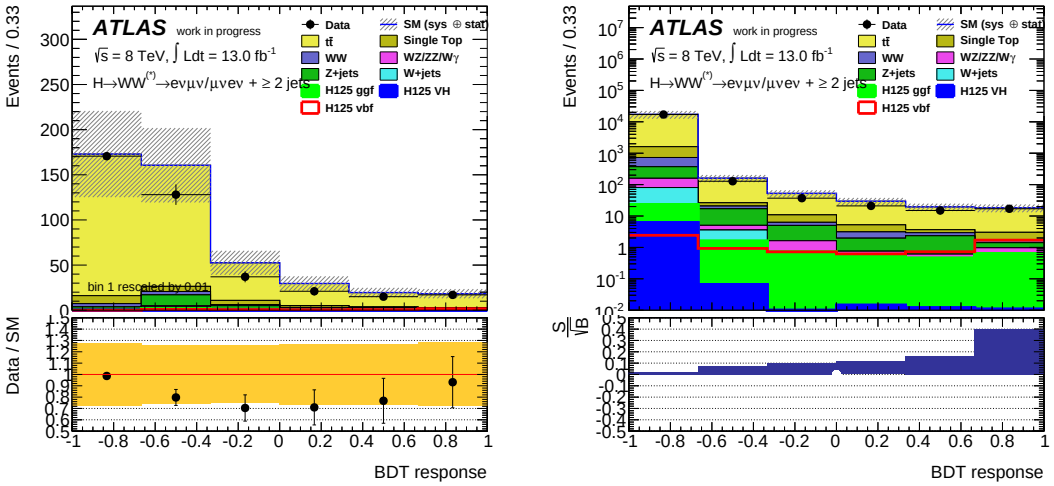


Figure 15: BDT response in linear and logarithmic scale with data over MC ratio (left) and significance $S/\sqrt{B}$ (right) for the baseline BDT. The error band shows statistical and systematic uncertainties.

the values for the control variables and the significance are listed in table 10. The runs are as follows:

run	base	1	2	3	4	5	6
overtraining sig	0.4	0.43	0.58	0.57	0.64	0.52	0.25
overtraining bkg	1.0	1.0	1.0	1.0	1.0	1.0	1.0
ROC integral	0.964	0.964	0.964	0.964	0.964	0.979	0.980
significance	0.4	0.42	0.32	0.34	0.34	0.48	0.54

Table 10: Values of the three control variables and the significance over the seven described runs.

Run 1: First, the event shape variables are added. Only a slight improvement in the significance and the signal overtraining parameter is visible because the distribution for signal and background of the transverse event shape variables is rather similar (see figure 17).

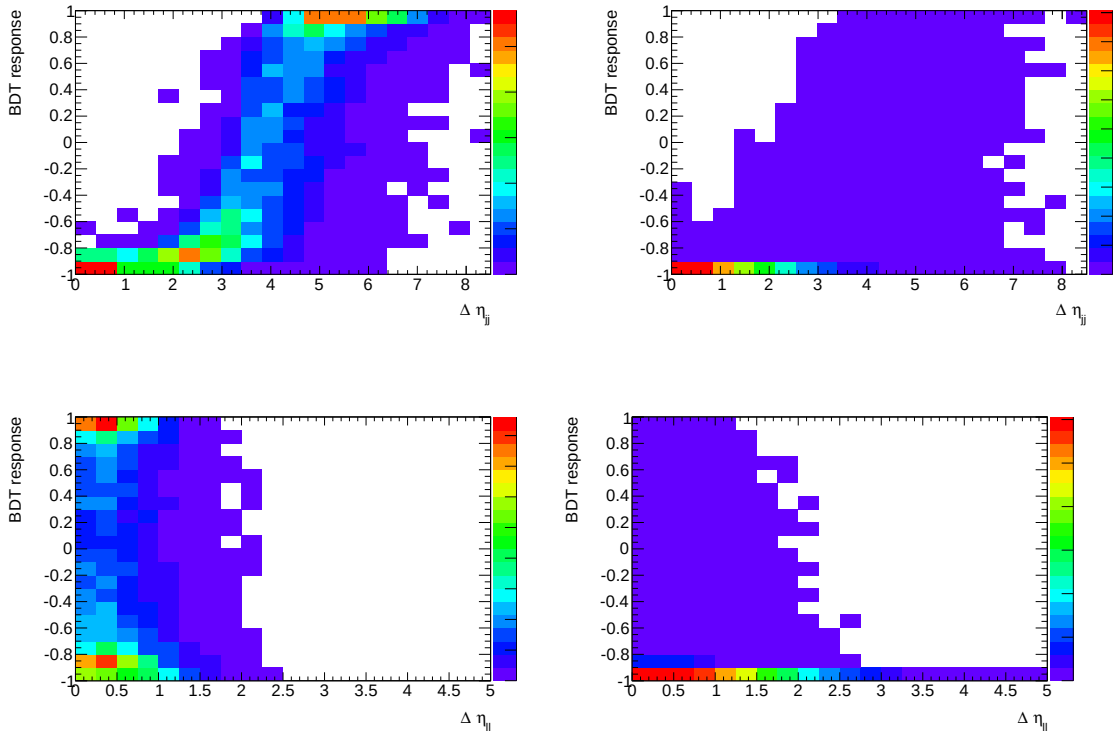


Figure 16: BDT response of the baseline BDT plotted against $\Delta\eta_{jj}$ for the signal sample (top left) and background sample (top right). The same is shown for $\Delta\eta_{ll}$ (bottom left, right). The third axis displays the normalized number of events.

Run 2: It adds the b-tag variables, the output of the b-tagging algorithm for the leading and subleading jet $\mathrm{MV1}_{jet1/2}$, and the number of b-tagged jets for different working points. This improves the signal overtraining parameter but strongly decreases the significance. All other variables stay at the same value. In figure 18, it can be seen that these variables show no separation between background and signal.

Run 3: The next variable added is the difference in the azimuthal angle between the dilepton system and the missing transverse energy $\Delta\phi_{ll,MET}$. After adding this variable, a slight increase in significance occurs. Figure 19 shows a little separation between background and signal.

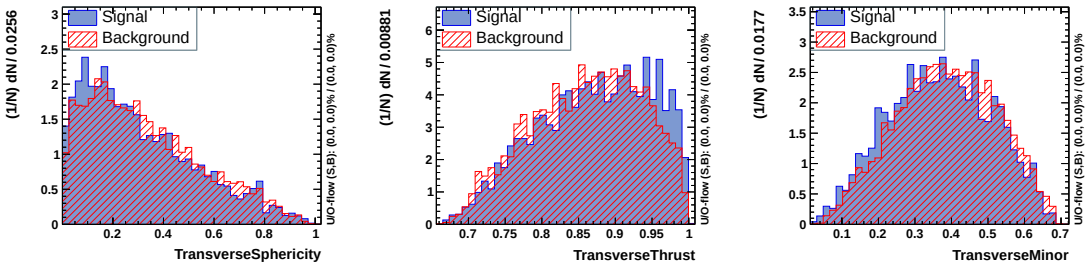


Figure 17: Input variable distribution for the transverse event shapes: Transverse sphericity, transverse thrust and transverse minor.

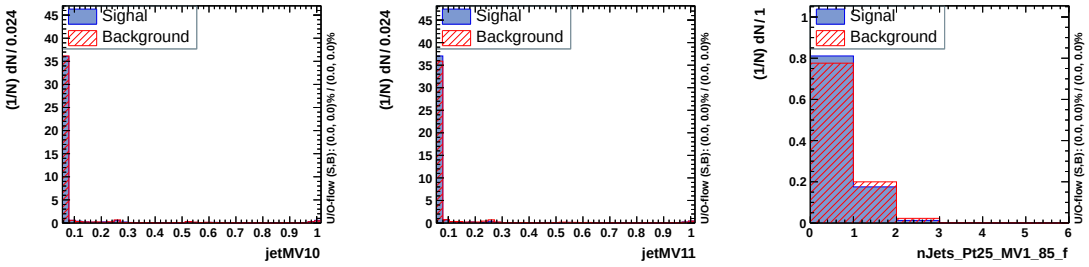


Figure 18: Input variable distribution for $\mathrm{MV1}_{jet1/2}$ and the number of b-tagged jets.

Run 4: A further variable to add is the difference in rapidity between the leading and subleading jet ΔY_{jj}. The control variable values stay the same and no improvement can be observed. A strong separation between background and signal can be seen in figure 19.

Run 5: The next variable studied is the maximum difference in pseudorapidity between two jets $\max(\Delta\eta_{jj})$. This variable shows an increase in the ROC integral and the significance. Furthermore, it strongly reduces the $\Delta\eta_{jj}$ dependence of the classifier. As before, a strong background versus signal separation (see figure 19) is visible.

Run 6: It adds two further variables which are the absolute of the pseudorapidity for leading and subleading jet $|\eta_{jet1/0}|$. Figure 20 shows a separation between signal and background.

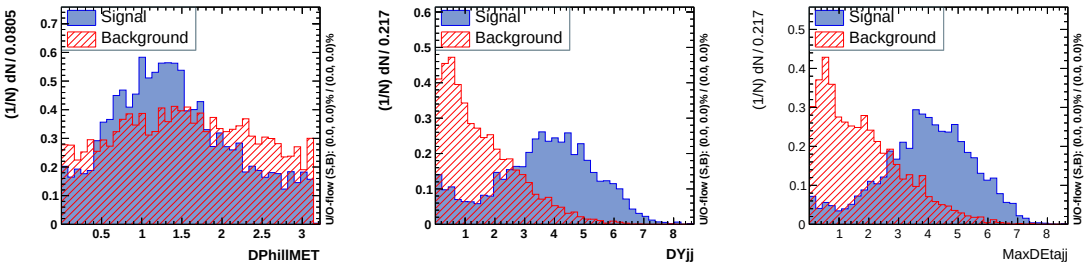


Figure 19: Input variable distribution for $\Delta\phi_{ll,MET}$, ΔY_{jj} and $max(\Delta\eta_{jj})$.

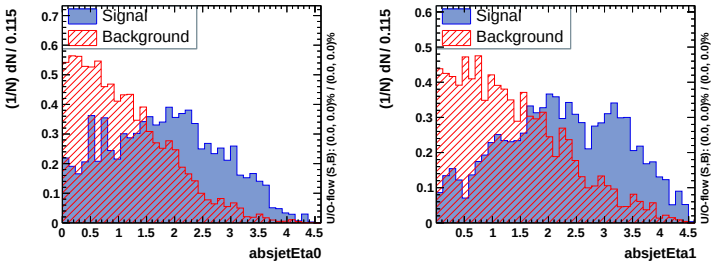


Figure 20: Input variable distribution for $|\eta_{jet1/0}|$.

In conclusion, the dependence on ΔY_{jj} was strongly reduced (see figure 21) especially by the addition of the variable max($\Delta\eta_{jj}$). All other variables did show slight improvement.
From this can be concluded that a lot of the basic variables are correlated. The linear correlation coefficients for some variables are shown in figure 22. Correlated variables are about equally good to separate signal from background. Thus, the classifier will always pick one over the other. So adding correlated variables will not significantly improve the classifier. The final choice of input variables will be to use all the ones described above and let the BDT do the work of picking the best ones. The only reason to leave out some of the correlated variables would be to improve the speed of the BDT. But then one would have to check by hand, which variables to use to obtain similar results.

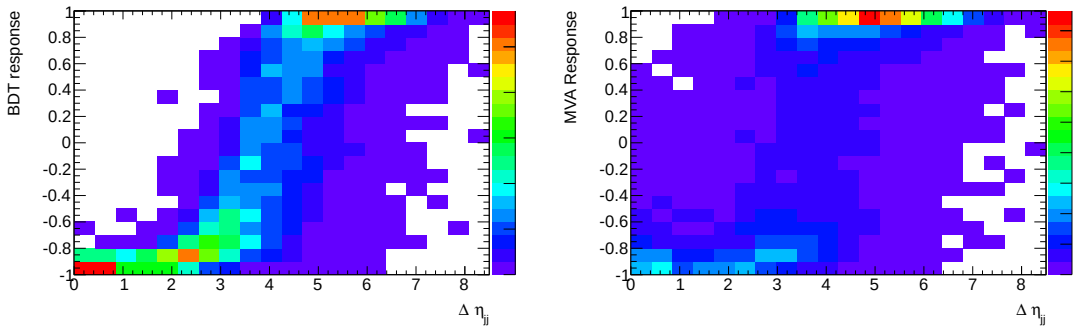


Figure 21: Improvement of the classifier over the above described runs. Baseline on the left and run 6 on the right.

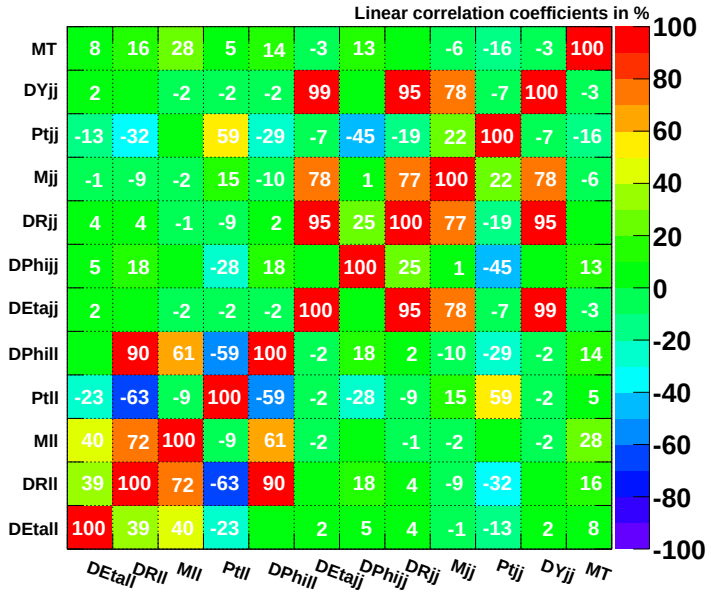


Figure 22: Linear correlation coefficients for certain input variables.

4.6.5 Parameter Optimization

The BDT parameters were optimized by varying the values in the six different channels. The used parameters and their values are listed in table 11. They result in a total of 18800 combinations.

parameter	min	max	step size
`NTrees`	500	1000	100
`GradBaggingFraction`	0.1	1.0	0.1
`Shrinkage`	0.1	0.5	0.1
`nEventsMin`	100	300	20
`MaxDepth`	2	8	1

Table 11: List of BDT parameters and the range within they were varied.

The optimization is done with respect to the ROC integral because as explained earlier it is sensitive to the efficiency. At the same time the signal and background overtraining parameter was required to be larger than 0.5. From the ten best runs fulfilling these conditions, the most suitable one is chosen. In table 12, the optimized parameters are shown for all six channels. The corresponding values of the control variables are listed in table 13.

4.7 Comparison between channels

So far, only the electron muon channel without separation in electron and muon momenta was described ($e\mu\mu e$). Now the other channels will be examined, using the same variables, but optimizing the parameters separately for every channel.

channel	`Fraction`	`Shrinkage`	`NTrees`	`nEventsMin`	`MaxDepth`
$e\mu$	0.7	0.2	1000	220	2
μe	0.4	0.1	800	260	2
$\mu\mu$	0.7	0.1	900	180	2
ee	0.3	0.1	500	240	2
$e\mu\mu e$	0.3	0.4	900	140	2
$ee\mu\mu$	0.6	0.3	800	120	2

Table 12: Optimized BDT parameters for the six channels: Gradient Bagging Fraction, shrinkage, number of trees, number of minimum events, and maximum depth.

channel	significance	overtrain sig	overtrain bkg	ROC integral
$e\mu$	0.49	0.61	1.0	0.97
μe	0.50	0.55	1.0	0.98
$\mu\mu$	0.29	0.52	1.0	0.98
ee	0.28	0.54	1.0	0.97
$e\mu\mu e$	0.48	0.76	1.0	0.98
$ee\mu\mu$	0.36	0.52	1.0	0.98

Table 13: ROC integral, overtraining parameters, and significance in the last bin of the BDT response for the six channels after optimization.

The values of the optimized parameters for the different channels are shown in table 12. The maximum depth of the tree is always 2. No dependencies can be seen in the number of trees, the minimum number of events, the gradient bagging fraction, and the shrinkage. These values seem almost arbitrary.
The obtained values for the ROC integral, the overtraining parameters, and the

significance in the last bin of the BDT response are depicted in table 13. The significance in the last bin, which is similar to a cut onto the BDT response with a value larger than 0.67, is chosen, in order to have still one observed event in the channel with the least events (ee). The ROC integral is around 0.97 for all channels. The background overtraining parameter also stays constant at 1.0. The signal overtraining parameter is best for the $e\mu$ channel with 0.6 and around 0.5 for the other channels. The significance is around 0.5 for the opposite flavour and around 0.3 for the same flavour channels.

channel	VBF 125	Total Bkg.	Obs.	$S/\sqrt{B}$
$e\mu$	$1.2 \pm 0.0 \pm 0.1$	$5.5 \pm 1.0 \pm 0.6$	6	$0.5 \pm 0.1 \pm 0.1$
μe	$0.8 \pm 0.0 \pm 0.1$	$2.5 \pm 0.7 \pm 0.3$	4	$0.5 \pm 0.1 \pm 0.2$
ee	$0.3 \pm 0.0 \pm 0.0$	$0.8 \pm 0.4 \pm 0.1$	1	$0.3 \pm 0.1 \pm 0.2$
$\mu\mu$	$0.6 \pm 0.0 \pm 0.1$	$4.7 \pm 0.9 \pm 0.6$	4	$0.3 \pm 0.1 \pm 0.2$
$e\mu\mu e$	$3.0 \pm 0.1 \pm 0.4$	$36.8 \pm 2.6 \pm 5.4$	28	$0.5 \pm 0.0 \pm 0.1$
$ee\mu\mu$	$1.6 \pm 0.0 \pm 0.3$	$18.9 \pm 1.8 \pm 2.9$	17	$0.4 \pm 0.0 \pm 0.2$

Table 14: Expected signal and background events compared to observed events in the last bin (BDT response > 0.67) for all channels. Statistical (first row) and systematic (second row) uncertainties are also shown. Also given is the significance based on simulated data.

It can be seen from figure 23 to 28 that the $t\bar{t}$ or Z + jets background is clearly classified as background and thus accumulated on the left. Only a few background events are located in the signal bin. This is also visible in the increase of the significance from the first to the last bin of the BDT response.
In Table 14, the expected signal and background events compared to the observed events are shown. There is good agreement within the statistical error except for the $e\mu\mu e$ channel, where a slight deviation of 1.5 standard deviations (statistical plus systematic error) is observed.

4.7.1 Opposite flavour channel

eμ channel

After optimizing the parameters and using the same input variables as described above, the BDT response as seen in figure 23 is obtained. A significance of $0.5 \pm$

0.1 (stat.) ± 0.1 (sys.) could be obtained in the last bin of the BDT response, where the signal and background overtraining parameter is 0.61 and 1.0 respectively. The ROC integral is 0.97 (see table 13). It can be seen that the data over MC ratio is very good within the statistical error. In the cut onto the last bin of the BDT response there are 6 measured events where $6.7 \pm 1.0\,(\text{stat.}) \pm 0.6\,(\text{sys.})$ events were expected (see table 14).

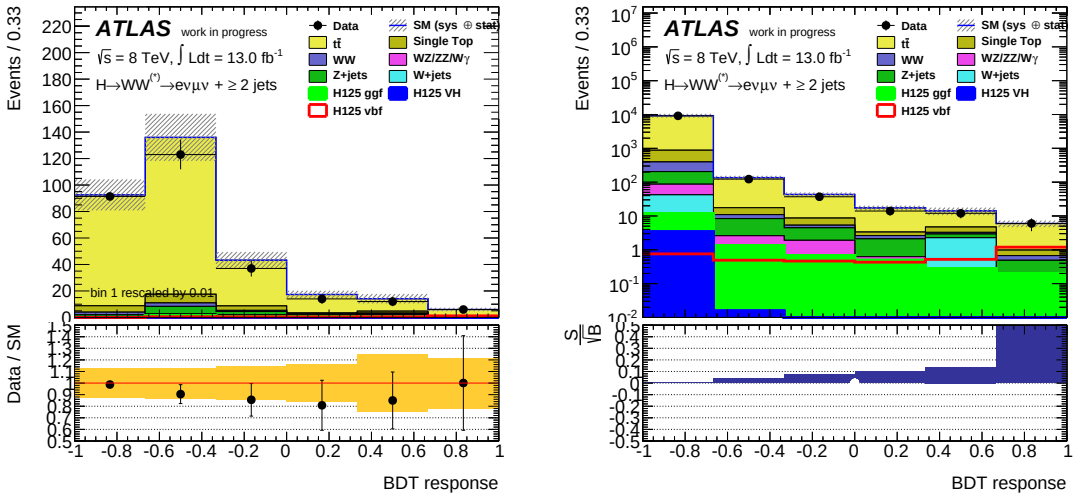


Figure 23: BDT response for the $e\mu$ channel. Left: Linear scale with data over MC ratio. Right: Logarithmic scale and significance. The error bands include systematic and statistical uncertainties.

μe channel

In figure 24, where the results for this channel can be seen, the significance ($0.5 \pm 0.1\,(\text{stat.}) \pm 0.2\,(\text{sys.})$) and the background overtraining parameter (1.0) stay almost the same. The signal overtraining parameter (0.55) is about 10% worse, therefore the ROC integral is slightly better: 0.98 (see table 13). The overall data over MC agreement is good and except for the second bin within the statistical error. The number of expected events in the last bin of the BDT response is $3.3 \pm 0.7\,(\text{stat.}) \pm 0.3\,(\text{sys.})$ compared to 4 measured events (see table 14).

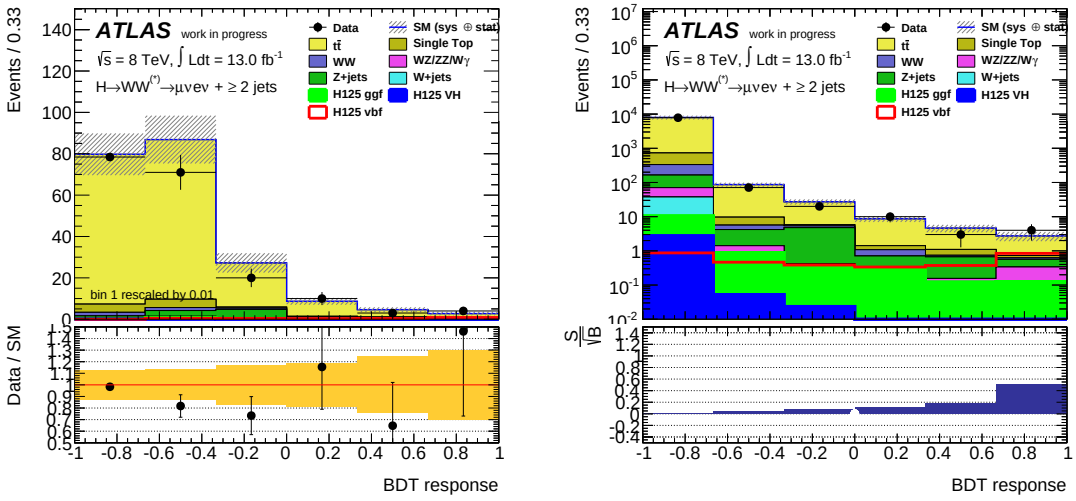


Figure 24: BDT response for the μe channel. Left: Linear scale with data over MC ratio. Right: Logarithmic scale and significance. The error bands include systematic and statistical uncertainties.

$e\mu\mu e$ channel

Figure 25 shows the BDT response for the $e\mu\mu e$ channel. Here, the two opposite flavour channels were trained together. The data over MC agreement is good except for the last bin BDT response, where $39.8 \pm 2.7\,(\text{stat.}) \pm 5.4\,(\text{sys.})$ events were expected but only 28 were measured (see table 14). The background overtraining parameter and the efficiency are as good as for the channels described above. The signal overtraining parameter with 0.76 is very good, whereas the significance of $0.5 \pm 0.0\,(\text{stat.}) \pm 0.1\,(\text{sys.})$ is the same as for the $e\mu$ and μe channel (see table 13). Since the significance for the separately trained channels together is $0.7 \pm 0.1\,(\text{stat.}) \pm 0.1\,(\text{sys.})$, it is better to train separately and combine the results afterwards. The data over MC agreement is better for the separately trained channels, too.

4.7.2 Same flavour channel

ee channel

Figure 26 shows the distribution of the BDT response for the ee channel. Good agreement in the data over MC ratio within the statistical error can be seen. While

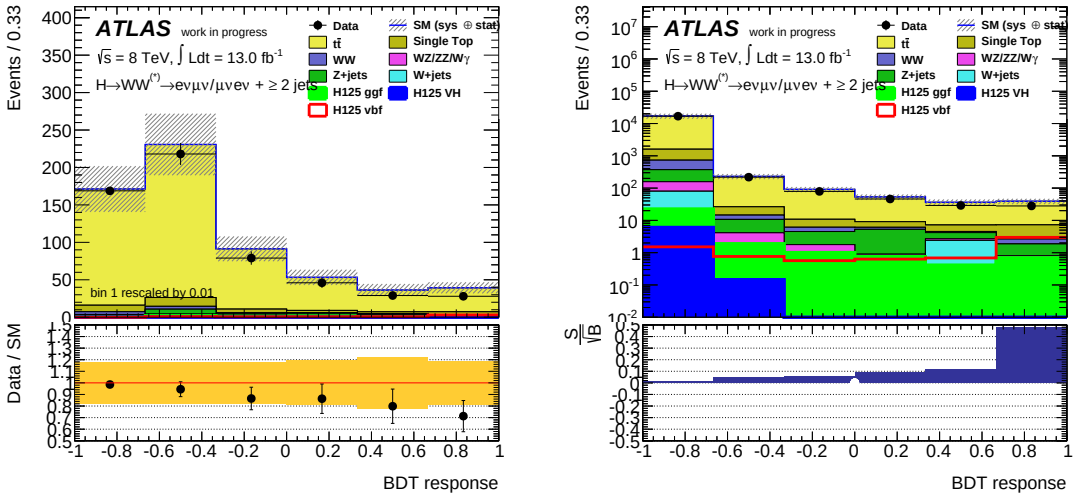


Figure 25: BDT response for the $e\mu\mu e$ channel. Left: Linear scale with data over MC ratio. Right: Logarithmic scale and significance. The error bands include systematic and statistical uncertainties.

the background (1.0) and signal overtraining parameters (0.54) stay as well as for the previously described channels, a significance of only $0.3 \pm 0.1\,(\text{stat.}) \pm 0.2\,(\text{sys.})$ could be obtained. The ROC integral with a value of 0.97 is comparable to the previously calculated (see table 13). Looking at the cutflow for the last bin BDT response, it can be seen that $1.1 \pm 0.4\,(\text{stat.}) \pm 0.1\,(\text{sys.})$ events were expected and 1 event was measured (see table 14).

$\mu\mu$ channel

Figure 27 shows the BDT response for the $\mu\mu$ channel. Except for the fourth bin a good agreement between data and MC is visible. The signal (0.52) and background (1.0) overtraining parameters are as good as before. The significance is $0.3 \pm 0.1\,(\text{stat.}) \pm 0.2\,(\text{sys.})$ and the ROC integral 0.98 (see table 13). $5.3 \pm 0.9\,(\text{stat.}) \pm 0.6\,(\text{sys.})$ events were expected and 4 were measured (see table 14).

ee$\mu\mu$ channel

The BDT response plots for the $ee\mu\mu$ channel are shown in figure 28. Like in the opposite flavour channel, the two same flavour channels were trained together. Good

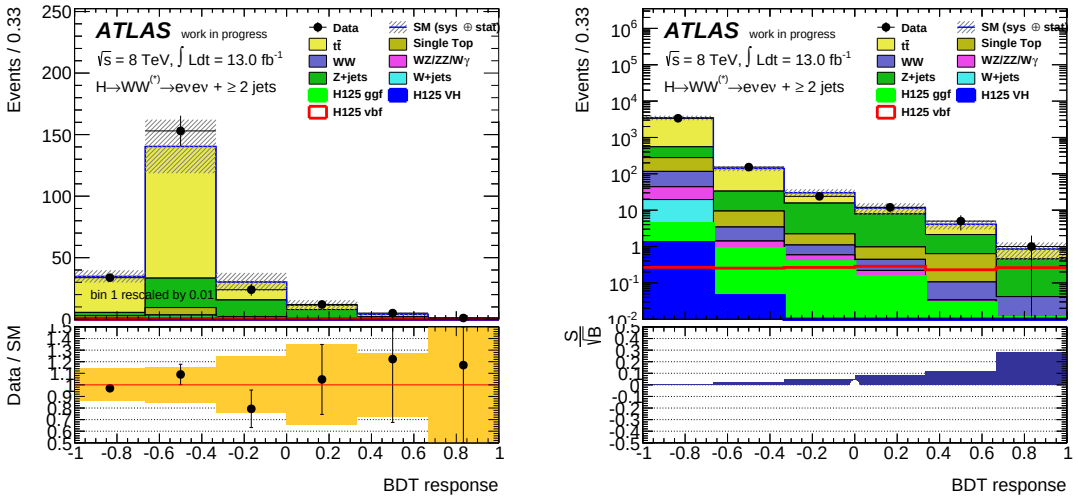


Figure 26: BDT response for the *ee* channel. Left: Linear scale with data over MC ratio. Right: Logarithmic scale and significance. The error bands include systematic and statistical uncertainties.

agreement in the data over MC ratio is observed. The ROC integral with 0.98 is the highest of all channels. The background and signal overtraining parameters are as good as in the above described channels with 0.52 and 1.0 respectively (see table 13). The number of expected events in the last bin BDT response is $20.5 \pm 1.8\,(\text{stat.}) \pm 2.9\,(\text{sys.})$ but only 17 were observed (see table 14).
Unlike in the combined opposite flavour channel, the significance with $0.4 \pm 0.0\,(\text{stat.}) \pm 0.2\,(\text{sys.})$ is as high as the significance for the separately trained channels taken together, which is $0.4 \pm 0.1\,(\text{stat.}) \pm 0.1\,(\text{sys.})$. So concerning the significance, the number of expected events and the control variables, there is no significant difference in separate or combined training. But the overall data over MC agreement is better in the combined same flavour channel.

4.8 Comparison Cut Based Analysis

Only a limited comparison with the cut-based analysis is possible, since associated production and vector boson fusion are used there as signals whereas in this analysis only vector boson fusion was used. Also, the cut-based analysis does not look into the same flavour channels. To overcome this problem, a BDT with vector boson

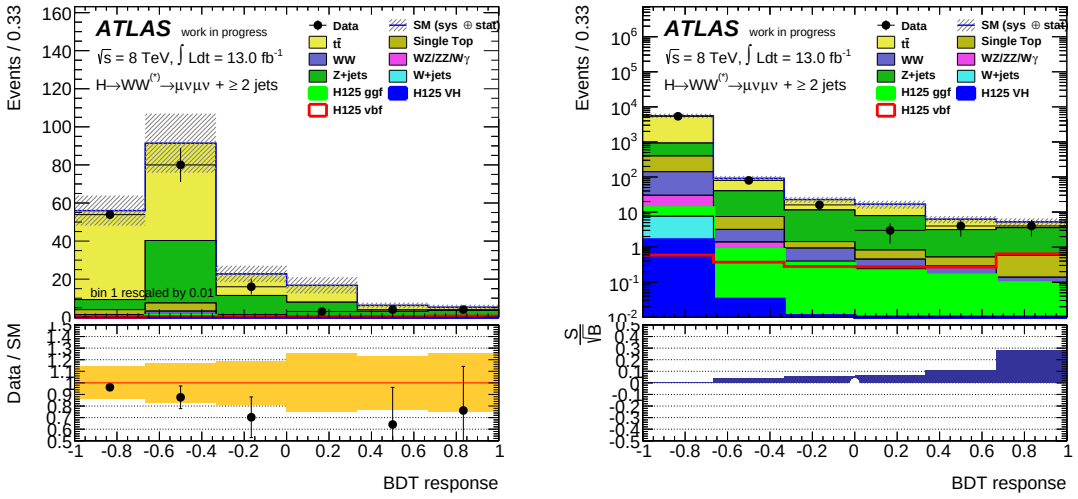


Figure 27: BDT response for the $\mu\mu$ channel. Left: Linear scale with data over MC ratio. Right: Logarithmic scale and significance. The error bands include systematic and statistical uncertainties.

fusion and associated production as signal was trained with the same input variables and optimization procedure. The results are shown in figure 29. Overtraining, ROC integral, and data over MC ratio are as good as the previously discussed results.
Table 15 shows the significance for the cut-based analysis compared to the BDT analysis. The significance in the BDT analysis is around 0.5 for the opposite flavour channel. There is no big difference whether associated production is signal or background. The cut based analysis is about 60% better in the $e\mu$ channel and 40% in the μe channel. In the opposite flavour channel, the significance is around 0.3. As previously mentioned a comparison is not possible.
Even though the BDT analysis uses additional information, the significance is lower than that for the cut based analysis. An explanation for this is, that the BDT is not optimized for recognizing the signal of the associated production. In order for this signal to have a structure similar to the vector boson fusion, there has to be initial state radiation, where the jet is emitted in forward direction. The Higgs boson then has to decay into a $b\bar{b}$ pair, where one jet is misidentified as lepton and the other jet is emitted in random direction. The probability for this is extremely low.

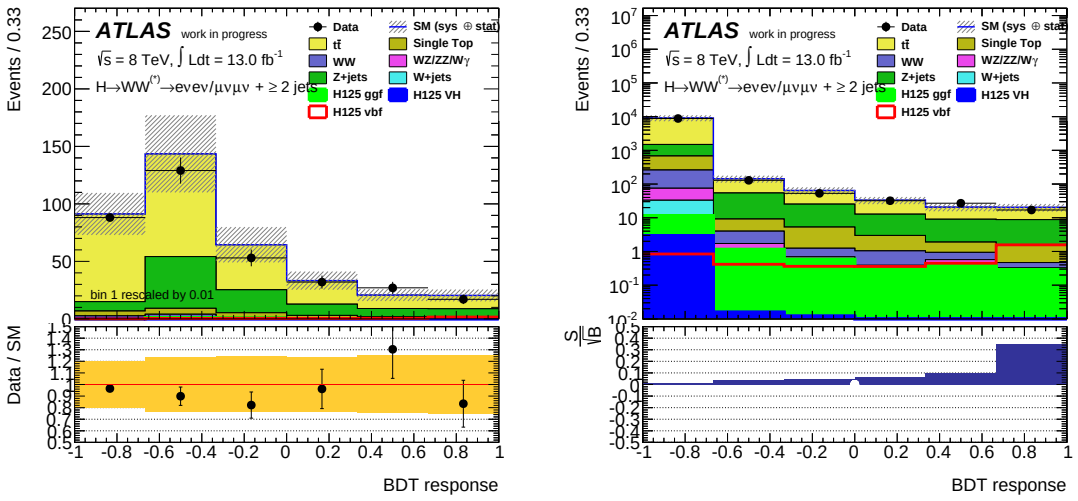


Figure 28: BDT response for the $ee\mu\mu$ channel. Left: Linear scale with data over MC ratio. Right: Logarithmic scale and significance. The error bands include systematic and statistical uncertainties.

channel	cut based	BDT(VBF + VH)	BDT (VBF)
$e\mu$	0.8	0.5	0.5
μe	0.7	0.5	0.5
ee	-	-	0.3
$\mu\mu$	-	-	0.3

Table 15: Significance of the cut based analysis compared to the BDT analysis with cuts onto the BDT response at 0.67 (last bin). The BDT significance is shown for vector boson fusion (VBF) as signal and vector boson fusion plus associated production as signal (VBF + VH).

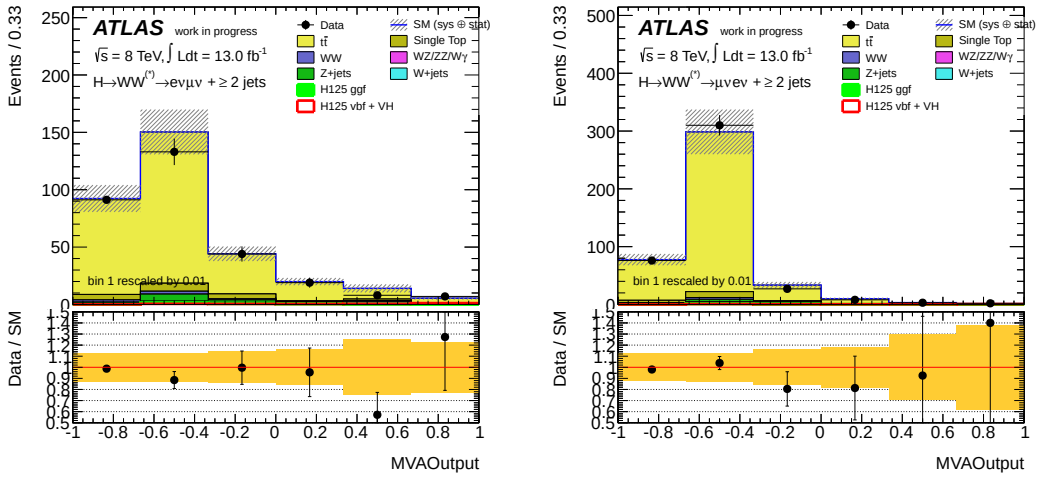


Figure 29: BDT response in the $e\mu$ (left) and μe channel (right) with associated production as additional signal. Both error bands include statistical and systematic uncertainties.

5 Summary and Outlook

This analysis searched for a Higgs boson in H $\rightarrow WW$ decays in the vector boson fusion channel. With the help of BDTs, $13\,\text{fb}^{-1}$ of ATLAS data at $\sqrt{s} = 8\,\text{TeV}$ were analyzed. No significant excess of events over statistical and systematic uncertainties was observed in this dataset. However, the enormous $t\bar{t}$ background could be reduced with the BDTs, yielding a significance of 0.4 in the same flavour and 0.7 in the opposite flavour channel. The significance of the cut-based analysis could not be reached or exceeded so far. Furthermore, a promising new variable, the maximum difference in pseudorapidity $max(\Delta\eta_{jj})$, was found , giving good signal versus background separation and reducing the dependence of the classifier on $\Delta\eta_{jj}$. Unfortunately, this cannot be said for the event shape variables. Nevertheless they could be used in further analyses to understand mismodelling in the same flavour channel. BDTs provide many free parameters and variables to adjust for optimization, since a change of variables or parameters strongly influences the BDT response. Here, the BDT parameters were optimized with respect to the signal and background overtraining and the ROC integral. This however, consumes a lot of computing power and time. In further analyses it should be recommended to use both, the cut based and the BDT analysis in order to compare and compete. In this thesis, only $\mathcal{L} = 13\,\text{fb}^{-1}$ of data were used. Expanding the analysis to the full $21\,\text{fb}^{-1}$ of 2012 data should increase the significance for H $\rightarrow WW$ decays in vector boson fusion production explicitly.

A BDT parameters

Short list of BDT parameters that have been the same for all channels and runs.

`VarTransform`	I
`NTrees`	channel specific
`BoostType`	Grad
`UseBaggedGrad`	True
`GradBaggingFraction`	channel specific
`Shrinkage`	channel specific
`SeparationType`	GiniIndex (default)
`nCuts`	20 (default)
`nEventsMin`	channel specific
`NNodesMax`	100000 (default)
`MaxDepth`	channel specific
`NegWeightTreatment`	IgnoreNegWeights

B Cutflow

	Signal	Signal other	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	Single Top	Z+jets	W+jets	Total Bkg.	Observed
at least 2 jets	9.2 ± 0.1	35.9 ± 0.7	358.3 ± 3.9	222.2 ± 11.7	14954.1 ± 50.5	845.8 ± 19.0	867.4 ± 16.0	157.4 ± 19.8	17441.0 ± 61.0	18649
lepton $p_{\mathrm{T}} > 25, 15$ GeV	8.0 ± 0.1	32.0 ± 0.7	339.6 ± 3.8	197.6 ± 10.9	14132.9 ± 49.1	802.3 ± 18.3	680.0 ± 14.3	141.4 ± 19.0	16325.9 ± 58.7	17464
OS leptons	8.0 ± 0.1	30.2 ± 0.6	338.7 ± 3.8	102.4 ± 8.2	14094.7 ± 49.1	776.1 ± 16.0	662.9 ± 13.3	87.7 ± 15.9	16092.6 ± 56.4	17098
$m_{\ell\ell} > 12, 10$ GeV	7.9 ± 0.1	30.0 ± 0.6	338.4 ± 3.8	101.8 ± 8.2	14085.4 ± 49.1	775.4 ± 16.0	661.5 ± 13.3	87.7 ± 15.9	16080.1 ± 56.4	17085
Z veto (for $ee, \mu\mu$)	7.9 ± 0.1	30.0 ± 0.6	338.4 ± 3.8	101.8 ± 8.2	14085.4 ± 49.1	775.4 ± 16.0	661.5 ± 13.3	87.7 ± 15.9	16080.1 ± 56.4	17085
$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45, 25$ GeV	3.9 ± 0.1	14.8 ± 0.4	195.5 ± 2.9	47.3 ± 7.8	7956.8 ± 36.9	459.4 ± 11.9	128.4 ± 5.9	32.4 ± 8.1	8834.6 ± 40.9	9330
NF factors					NF = 1.08	NF = 1.08				
NFs applied	3.9 ± 0.1	14.8 ± 0.4	195.5 ± 2.9	47.3 ± 7.8	8558.1 ± 39.6	494.1 ± 12.8	128.4 ± 5.9	32.4 ± 8.1	9470.6 ± 43.6	9330
BDTResponse > 0.67	1.2 ± 0.0	0.2 ± 0.0	0.2 ± 0.1	0.0 ± 0.0	4.5 ± 0.9	0.3 ± 0.3	0.3 ± 0.2	0.0 ± 0.0	5.5 ± 1.0	6
BDTResponse > 0.7	1.1 ± 0.0	0.2 ± 0.0	0.2 ± 0.1	0.0 ± 0.0	4.0 ± 0.9	0.3 ± 0.3	0.3 ± 0.2	0.0 ± 0.0	4.9 ± 0.9	4
BDTResponse > 0.8	0.8 ± 0.0	0.1 ± 0.0	0.2 ± 0.1	0.0 ± 0.0	2.7 ± 0.7	0.3 ± 0.3	0.1 ± 0.1	0.0 ± 0.0	3.4 ± 0.7	2
BDTResponse > 0.9	0.5 ± 0.0	0.1 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	1.0 ± 0.4	0.3 ± 0.3	0.0 ± 0.0	0.0 ± 0.0	1.4 ± 0.5	1

	Signal	Signal other	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	Single Top	Z+jets	W+jets	Total Bkg.	Observed
at least 2 jets	7.7 ± 0.1	30.4 ± 0.7	312.6 ± 3.6	210.6 ± 8.9	13089.8 ± 47.2	725.4 ± 15.5	771.3 ± 16.1	167.2 ± 40.0	15307.2 ± 66.4	16006
lepton $p_{\mathrm{T}} > 25, 15$ GeV	6.8 ± 0.1	26.7 ± 0.7	298.1 ± 3.5	186.3 ± 8.8	12442.4 ± 46.0	683.0 ± 14.7	599.3 ± 14.8	142.2 ± 38.7	14377.9 ± 64.3	15004
OS leptons	6.8 ± 0.1	25.0 ± 0.6	297.5 ± 3.5	87.8 ± 5.3	12421.5 ± 45.9	670.8 ± 14.4	585.2 ± 13.7	71.7 ± 14.3	14159.4 ± 52.5	14708
$m_{\ell\ell} > 12, 10$ GeV	6.7 ± 0.1	24.9 ± 0.6	297.4 ± 3.5	87.1 ± 5.3	12410.9 ± 45.9	670.7 ± 14.4	583.1 ± 13.7	69.8 ± 14.2	14143.7 ± 52.4	14696
Z veto (for $ee, \mu\mu$)	6.7 ± 0.1	24.9 ± 0.6	297.4 ± 3.5	87.1 ± 5.3	12410.9 ± 45.9	670.7 ± 14.4	583.1 ± 13.7	69.8 ± 14.2	14143.7 ± 52.4	14696
$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45, 25$ GeV	3.3 ± 0.1	12.3 ± 0.4	171.4 ± 2.7	33.7 ± 3.1	6968.3 ± 34.4	380.2 ± 10.8	102.2 ± 4.7	27.1 ± 10.9	7695.4 ± 38.2	7950
NF factors					NF = 1.06	NF = 1.06				
NFs applied	3.3 ± 0.1	12.3 ± 0.4	171.4 ± 2.7	33.7 ± 3.1	7355.2 ± 36.3	401.3 ± 11.4	102.2 ± 4.7	27.1 ± 10.9	8103.3 ± 40.1	7950
BDTResponse > 0.67	0.8 ± 0.0	0.1 ± 0.0	0.0 ± 0.0	0.1 ± 0.1	1.9 ± 0.7	0.1 ± 0.1	0.2 ± 0.2	0.0 ± 0.0	2.5 ± 0.7	4
BDTResponse > 0.7	0.8 ± 0.0	0.1 ± 0.0	0.0 ± 0.0	0.1 ± 0.1	1.5 ± 0.6	0.1 ± 0.1	0.2 ± 0.2	0.0 ± 0.0	2.1 ± 0.7	4
BDTResponse > 0.8	0.6 ± 0.0	0.1 ± 0.0	0.0 ± 0.0	0.1 ± 0.1	0.9 ± 0.4	0.1 ± 0.1	0.1 ± 0.1	0.0 ± 0.0	1.4 ± 0.5	3
BDTResponse > 0.9	0.4 ± 0.0	0.1 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.2 ± 0.3	0.1 ± 0.1	0.1 ± 0.1	0.0 ± 0.0	0.4 ± 0.3	1

Table 16: Full cutflow for the $e\mu$ (top) and the μe channel (bottom). Only statistical uncertainties are shown.

	Signal	Signal other	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	Single Top	Z+jets	W+jets	Total Bkg.	Observed
at least 2 jets	6.9 ± 0.1	29.0 ± 0.7	275.9 ± 3.4	503.7 ± 11.2	11338.3 ± 43.5	626.2 ± 14.4	196470.2 ± 751.2	187.7 ± 22.8	209430.9 ± 753.0	215608
lepton $p_{\mathrm{T}} > 25, 15$ GeV	6.0 ± 0.1	25.5 ± 0.7	260.8 ± 3.3	464.0 ± 10.9	10707.3 ± 42.3	591.3 ± 14.0	186720.1 ± 738.5	175.4 ± 22.2	198944.4 ± 740.3	203697
OS leptons	6.0 ± 0.1	23.5 ± 0.6	259.1 ± 3.3	345.2 ± 8.7	10662.1 ± 42.2	579.5 ± 13.4	185743.3 ± 736.7	105.7 ± 16.9	197718.5 ± 738.2	202514
$m_{\ell\ell} > 12, 10$ GeV	5.9 ± 0.1	23.2 ± 0.6	258.6 ± 3.3	342.7 ± 8.7	10631.8 ± 42.1	577.7 ± 13.4	185002.5 ± 736.6	105.7 ± 16.9	196942.1 ± 738.2	200276
Z veto (for $ee, \mu\mu$)	5.8 ± 0.1	20.4 ± 0.5	206.6 ± 2.9	138.1 ± 8.3	8340.1 ± 37.4	451.7 ± 11.6	22755.1 ± 196.3	81.9 ± 12.8	31993.9 ± 200.8	33824
$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45, 25$ GeV	1.6 ± 0.0	5.7 ± 0.3	75.2 ± 1.7	25.3 ± 3.8	2943.7 ± 22.2	164.4 ± 7.0	274.3 ± 19.2	15.3 ± 5.4	3504.0 ± 30.9	3571
NF factors					NF = 1.04	NF = 1.04	NF = 1.17			
NFs applied	1.6 ± 0.0	5.7 ± 0.3	75.2 ± 1.7	25.3 ± 3.8	3051.5 ± 23.0	170.4 ± 7.3	321.2 ± 22.5	15.3 ± 5.4	3664.7 ± 33.6	3571
BDTResponse > 0.67	0.3 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.4 ± 0.3	0.0 ± 0.0	0.4 ± 0.2	0.0 ± 0.0	0.8 ± 0.4	1
BDTResponse > 0.7	0.2 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.2 ± 0.3	0.0 ± 0.0	0.2 ± 0.2	0.0 ± 0.0	0.5 ± 0.3	1
BDTResponse > 0.8	0.1 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.2 ± 0.2	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.2 ± 0.2	1
BDTResponse > 0.9	0.1 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.2 ± 0.2	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.2 ± 0.2	0

	Signal	Signal other	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	Single Top	Z+jets	W+jets	Total Bkg.	Observed
at least 2 jets	10.5 ± 0.1	45.8 ± 0.9	410.4 ± 4.2	442.5 ± 4.6	16805.9 ± 54.0	907.7 ± 17.4	323594.2 ± 975.7	44.2 ± 9.7	342250.7 ± 977.4	343081
lepton $p_{\mathrm{T}} > 25, 15$ GeV	9.3 ± 0.1	41.8 ± 0.9	391.6 ± 4.1	419.0 ± 4.4	15997.7 ± 52.7	864.3 ± 17.0	310101.4 ± 961.6	31.5 ± 8.1	327847.3 ± 963.2	328210
OS leptons	9.3 ± 0.1	39.4 ± 0.8	391.6 ± 4.1	376.4 ± 3.8	15990.5 ± 52.7	855.4 ± 16.8	310098.1 ± 961.6	23.1 ± 7.0	327774.6 ± 963.2	327979
$m_{\ell\ell} > 12, 10$ GeV	9.1 ± 0.1	38.8 ± 0.8	390.9 ± 4.1	374.7 ± 3.8	15945.6 ± 52.6	853.8 ± 16.8	308730.6 ± 961.5	21.0 ± 6.7	326355.4 ± 963.1	323529
Z veto (for $ee, \mu\mu$)	9.0 ± 0.1	34.3 ± 0.7	307.8 ± 3.6	81.5 ± 2.6	12441.5 ± 46.5	669.8 ± 15.0	40163.9 ± 262.2	14.1 ± 5.3	53712.8 ± 266.8	54882
$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45, 25$ GeV	2.4 ± 0.1	8.7 ± 0.3	114.0 ± 2.2	17.5 ± 1.1	4559.7 ± 28.2	250.6 ± 9.3	501.3 ± 26.6	6.0 ± 3.5	5457.9 ± 40.1	5496
NF factors					NF = 1.04	NF = 1.04	NF = 1.17			
NFs applied	2.4 ± 0.1	8.7 ± 0.3	114.0 ± 2.2	17.5 ± 1.1	4751.7 ± 29.4	261.1 ± 9.7	585.0 ± 31.0	6.0 ± 3.5	5744.2 ± 44.0	5496
BDTResponse > 0.67	0.6 ± 0.0	0.1 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	1.6 ± 0.6	0.4 ± 0.4	2.6 ± 0.5	0.0 ± 0.0	4.7 ± 0.9	4
BDTResponse > 0.7	0.6 ± 0.0	0.1 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	1.5 ± 0.6	0.4 ± 0.4	2.2 ± 0.5	0.0 ± 0.0	4.2 ± 0.9	3
BDTResponse > 0.8	0.5 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	1.0 ± 0.5	0.4 ± 0.4	1.9 ± 0.5	0.0 ± 0.0	3.4 ± 0.8	3
BDTResponse > 0.9	0.3 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.5 ± 0.4	0.0 ± 0.0	0.9 ± 0.3	0.0 ± 0.0	1.5 ± 0.5	1

Table 17: Full cutflow for the ee (top) and the $\mu\mu$ channel (bottom). Only statistical uncertainties are shown.

	Signal	Signal other	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	Single Top	Z+jets	W+jets	Total Bkg.	Observed
at least 2 jets	16.9 ± 0.1	66.2 ± 1.0	670.9 ± 5.3	432.9 ± 14.7	28043.9 ± 69.1	1571.2 ± 24.6	1638.7 ± 22.7	324.6 ± 44.6	32748.3 ± 90.2	34655
lepton $p_{\mathrm{T}} > 25, 15$ GeV	14.7 ± 0.1	58.7 ± 1.0	637.7 ± 5.2	383.9 ± 14.0	26575.3 ± 67.3	1485.3 ± 23.4	1279.3 ± 20.5	283.6 ± 43.1	30703.8 ± 87.1	32468
OS leptons	14.7 ± 0.1	55.1 ± 0.9	636.1 ± 5.2	190.1 ± 9.8	26516.2 ± 67.2	1446.9 ± 21.5	1248.1 ± 19.1	159.4 ± 21.4	30252.0 ± 77.0	31806
$m_{\ell\ell} > 12, 10$ GeV	14.6 ± 0.1	54.9 ± 0.9	635.8 ± 5.2	188.8 ± 9.8	26496.2 ± 67.2	1446.1 ± 21.5	1244.5 ± 19.1	157.4 ± 21.3	30223.8 ± 77.0	31781
Z veto (for $ee, \mu\mu$)	14.6 ± 0.1	54.9 ± 0.9	635.8 ± 5.2	188.8 ± 9.8	26496.2 ± 67.2	1446.1 ± 21.5	1244.5 ± 19.1	157.4 ± 21.3	30223.8 ± 77.0	31781
$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45, 25$ GeV	7.1 ± 0.1	27.1 ± 0.6	366.9 ± 3.9	81.1 ± 8.4	14925.1 ± 50.4	839.6 ± 16.0	230.6 ± 7.6	59.5 ± 13.6	16530.0 ± 55.9	17280
NF factors					NF = 1.07	NF = 1.07				
NFs applied	7.1 ± 0.1	27.1 ± 0.6	366.9 ± 3.9	81.1 ± 8.4	15913.2 ± 53.8	895.2 ± 17.1	230.6 ± 7.6	59.5 ± 13.6	17573.6 ± 59.2	17280
BDTResponse > 0.67	3.0 ± 0.1	0.7 ± 0.1	0.8 ± 0.2	0.1 ± 0.1	29.9 ± 2.3	4.3 ± 1.1	1.1 ± 0.3	0.0 ± 0.0	36.8 ± 2.6	28
BDTResponse > 0.7	2.9 ± 0.1	0.6 ± 0.1	0.7 ± 0.2	0.1 ± 0.1	27.3 ± 2.2	3.9 ± 1.1	0.8 ± 0.3	0.0 ± 0.0	33.5 ± 2.5	25
BDTResponse > 0.8	2.5 ± 0.1	0.5 ± 0.1	0.6 ± 0.1	0.1 ± 0.1	18.8 ± 1.8	2.4 ± 0.9	0.3 ± 0.2	0.0 ± 0.0	22.6 ± 2.0	17
BDTResponse > 0.9	1.9 ± 0.0	0.3 ± 0.1	0.4 ± 0.1	0.1 ± 0.1	10.6 ± 1.4	1.8 ± 0.7	0.2 ± 0.1	0.0 ± 0.0	13.4 ± 1.6	9

	Signal	Signal other	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	Single Top	Z+jets	W+jets	Total Bkg.	Observed
at least 2 jets	17.4 ± 0.1	74.8 ± 1.2	686.4 ± 5.4	946.1 ± 12.1	28144.2 ± 69.4	1533.9 ± 22.6	520064.3 ± 1231.4	231.9 ± 24.7	551681.6 ± 1233.8	558689
lepton $p_{\mathrm{T}} > 25, 15$ GeV	15.2 ± 0.1	67.3 ± 1.1	652.4 ± 5.2	883.0 ± 11.7	26705.0 ± 67.6	1455.6 ± 22.0	496821.4 ± 1212.5	206.9 ± 23.7	526791.7 ± 1214.8	531907
OS leptons	15.2 ± 0.1	62.9 ± 1.0	650.8 ± 5.2	721.7 ± 9.5	26652.6 ± 67.5	1434.9 ± 21.5	495841.5 ± 1211.3	128.7 ± 18.3	525493.1 ± 1213.6	530493
$m_{\ell\ell} > 12, 10$ GeV	15.0 ± 0.1	62.0 ± 1.0	649.5 ± 5.2	717.4 ± 9.5	26577.4 ± 67.4	1431.4 ± 21.5	493733.0 ± 1211.2	126.7 ± 18.2	523297.5 ± 1213.4	523805
Z veto (for $ee, \mu\mu$)	14.8 ± 0.1	54.7 ± 0.9	514.4 ± 4.6	219.5 ± 8.7	20781.6 ± 59.7	1121.4 ± 19.0	62919.0 ± 327.6	96.0 ± 13.9	85706.6 ± 334.0	88706
$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45, 25$ GeV	4.0 ± 0.1	14.5 ± 0.4	189.2 ± 2.8	42.9 ± 4.0	7503.5 ± 35.9	415.0 ± 11.6	775.6 ± 32.8	21.3 ± 6.4	8961.9 ± 50.6	9067
NF factors					NF = 1.04	NF = 1.04	NF = 1.17			
NFs applied	4.0 ± 0.1	14.5 ± 0.4	189.2 ± 2.8	42.9 ± 4.0	7803.2 ± 37.3	431.6 ± 12.1	906.3 ± 38.3	21.3 ± 6.4	9408.9 ± 55.4	9067
BDTResponse > 0.67	1.6 ± 0.0	0.3 ± 0.1	0.1 ± 0.1	0.0 ± 0.0	11.1 ± 1.5	1.1 ± 0.7	6.2 ± 0.9	0.0 ± 0.0	18.9 ± 1.8	17
BDTResponse > 0.7	1.5 ± 0.0	0.3 ± 0.1	0.1 ± 0.0	0.0 ± 0.0	10.8 ± 1.4	0.8 ± 0.6	5.8 ± 0.8	0.0 ± 0.0	17.9 ± 1.8	15
BDTResponse > 0.8	1.3 ± 0.0	0.2 ± 0.0	0.1 ± 0.0	0.0 ± 0.0	6.2 ± 1.2	0.5 ± 0.5	4.5 ± 0.7	0.0 ± 0.0	11.5 ± 1.5	8
BDTResponse > 0.9	1.0 ± 0.0	0.1 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	2.9 ± 0.8	0.5 ± 0.5	2.7 ± 0.6	0.0 ± 0.0	6.1 ± 1.1	7

Table 18: Full cutflow for the $e\mu\mu e$ (top) and the $ee\mu\mu$ channel (bottom). Only statistical uncertainties are shown.

C Control plots

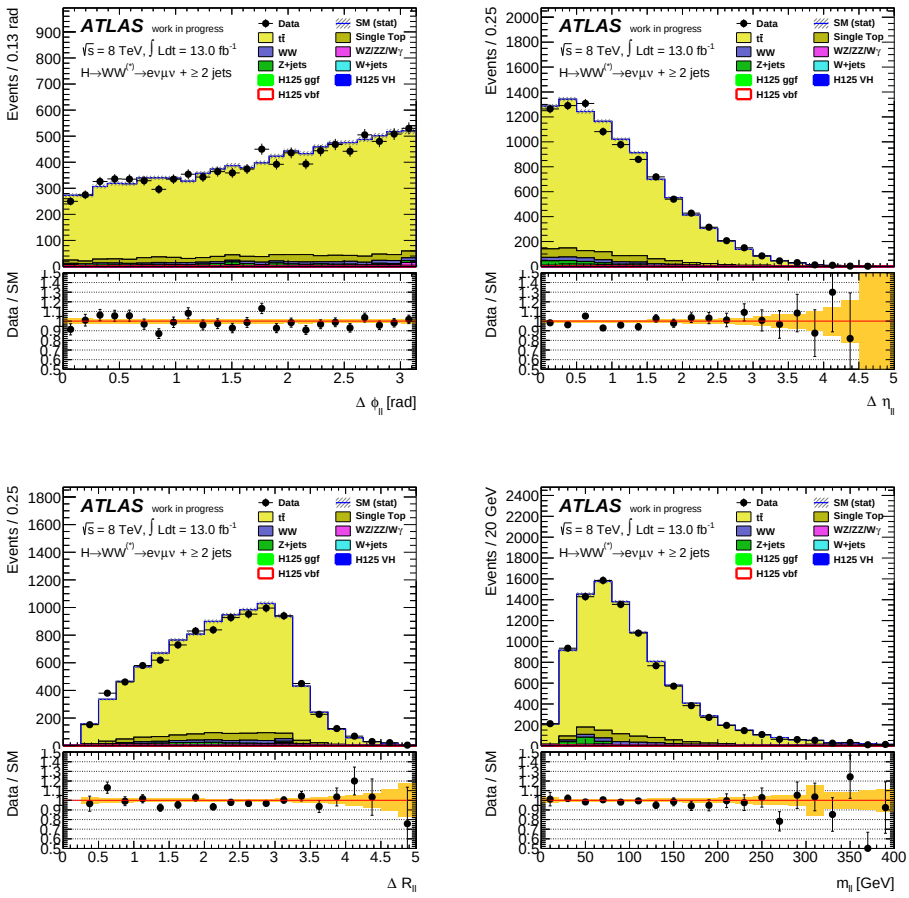


Figure 30: Input variables for the $e\mu$ channel with statistical error band.

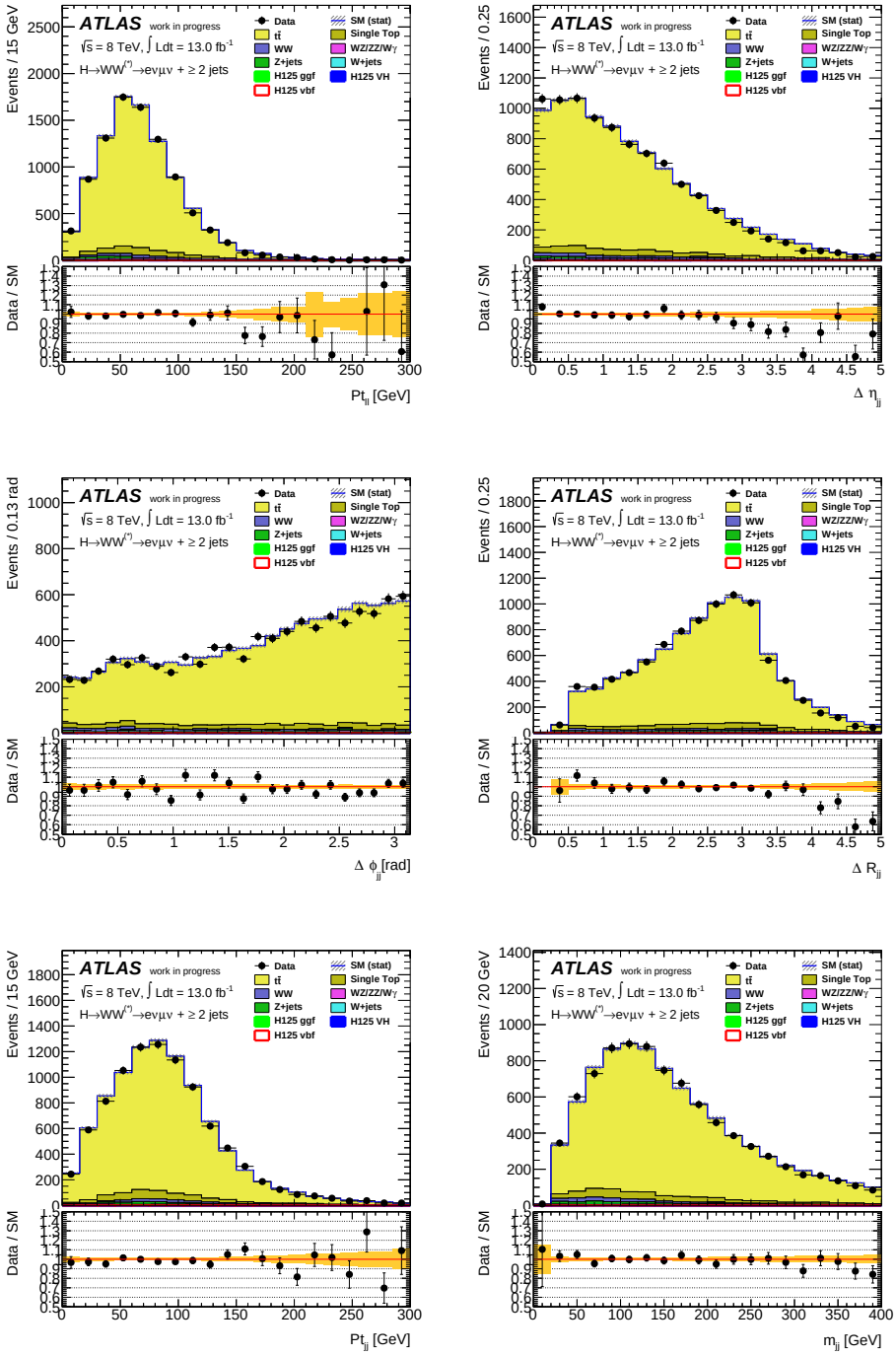


Figure 31: Input variables for the $e\mu$ channel with statistical error band.

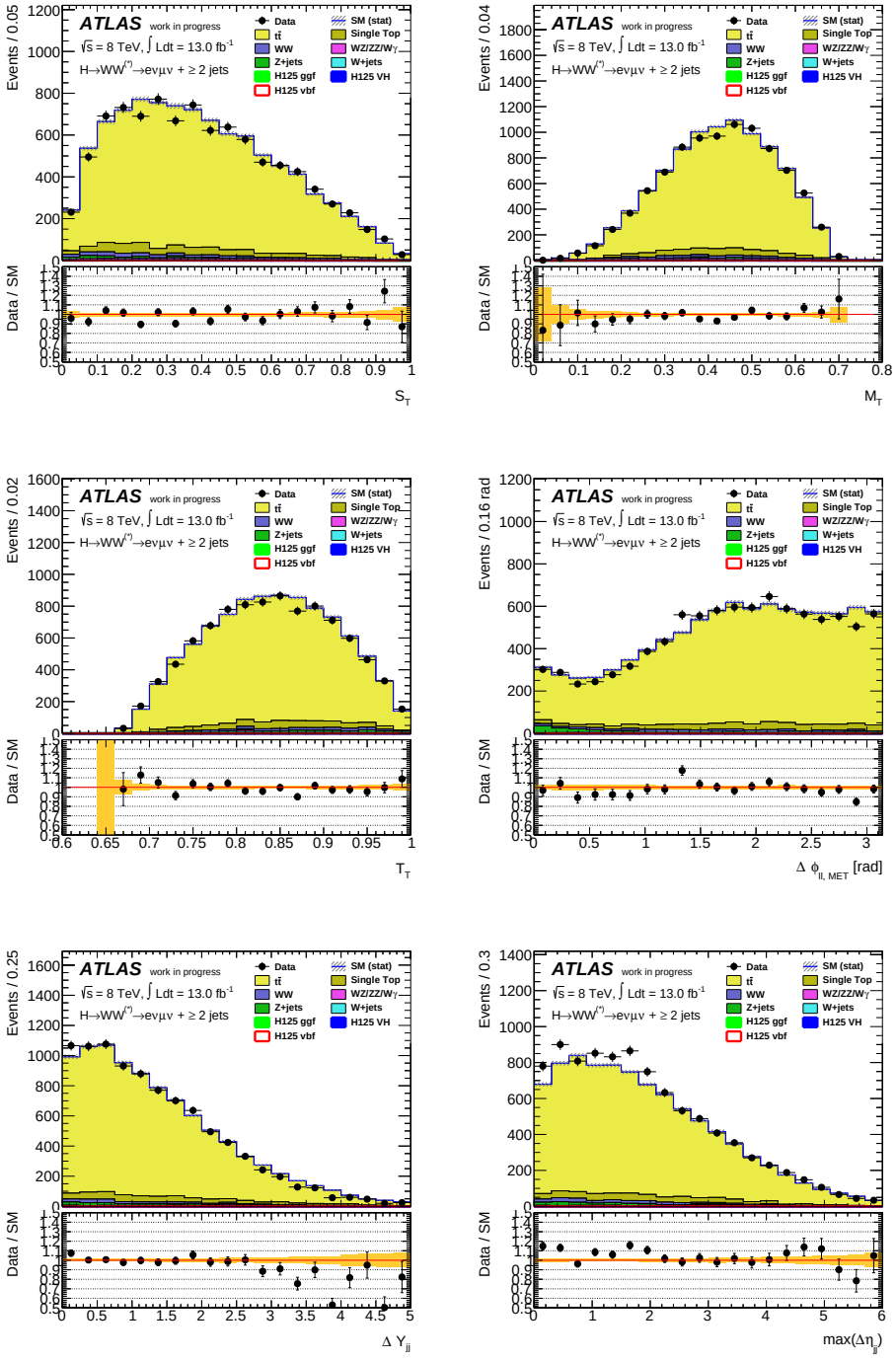


Figure 32: Input variables for the $e\mu$ channel with statistical error band.

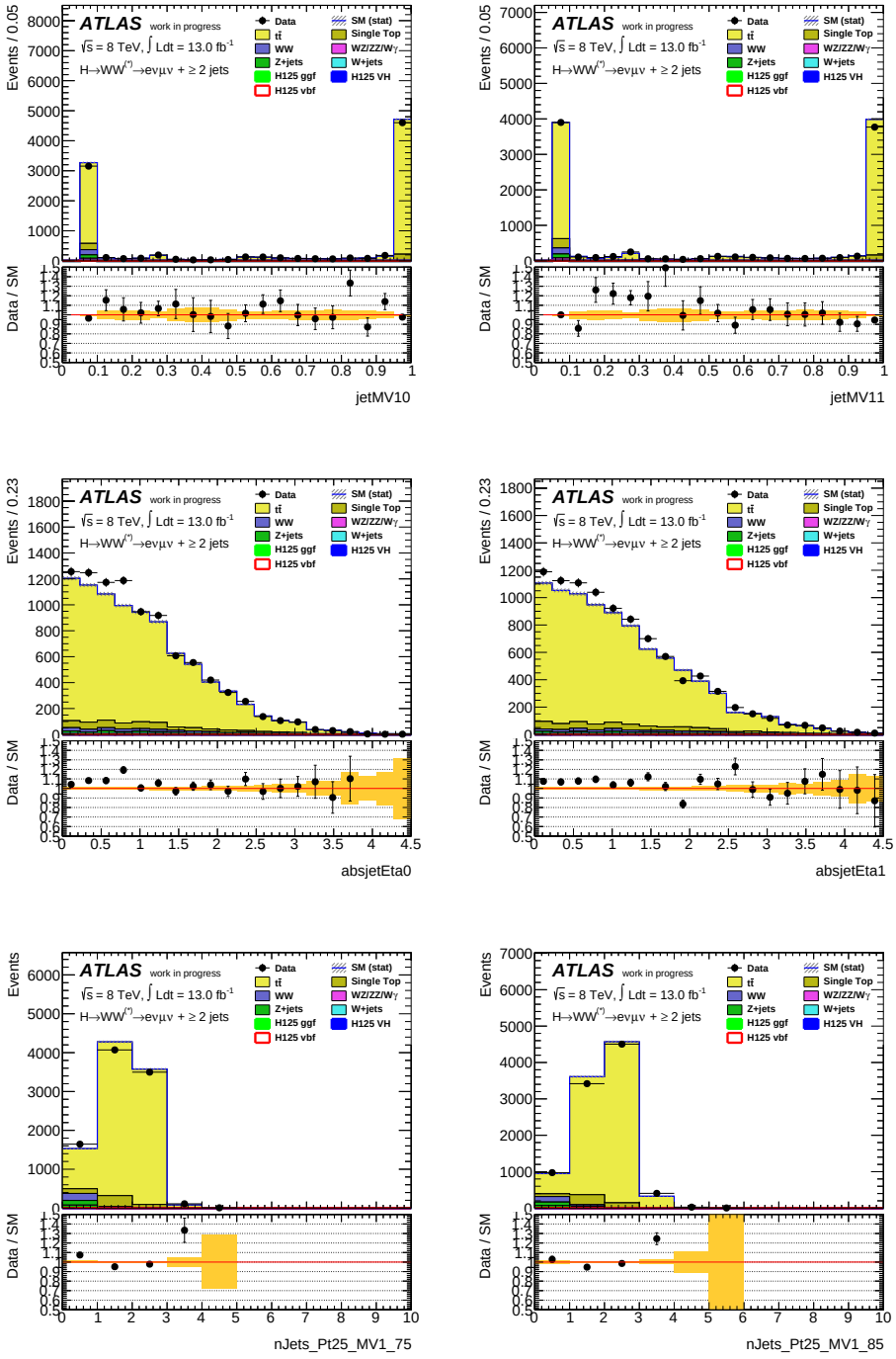


Figure 33: Input variables for the $e\mu$ channel with statistical error band.

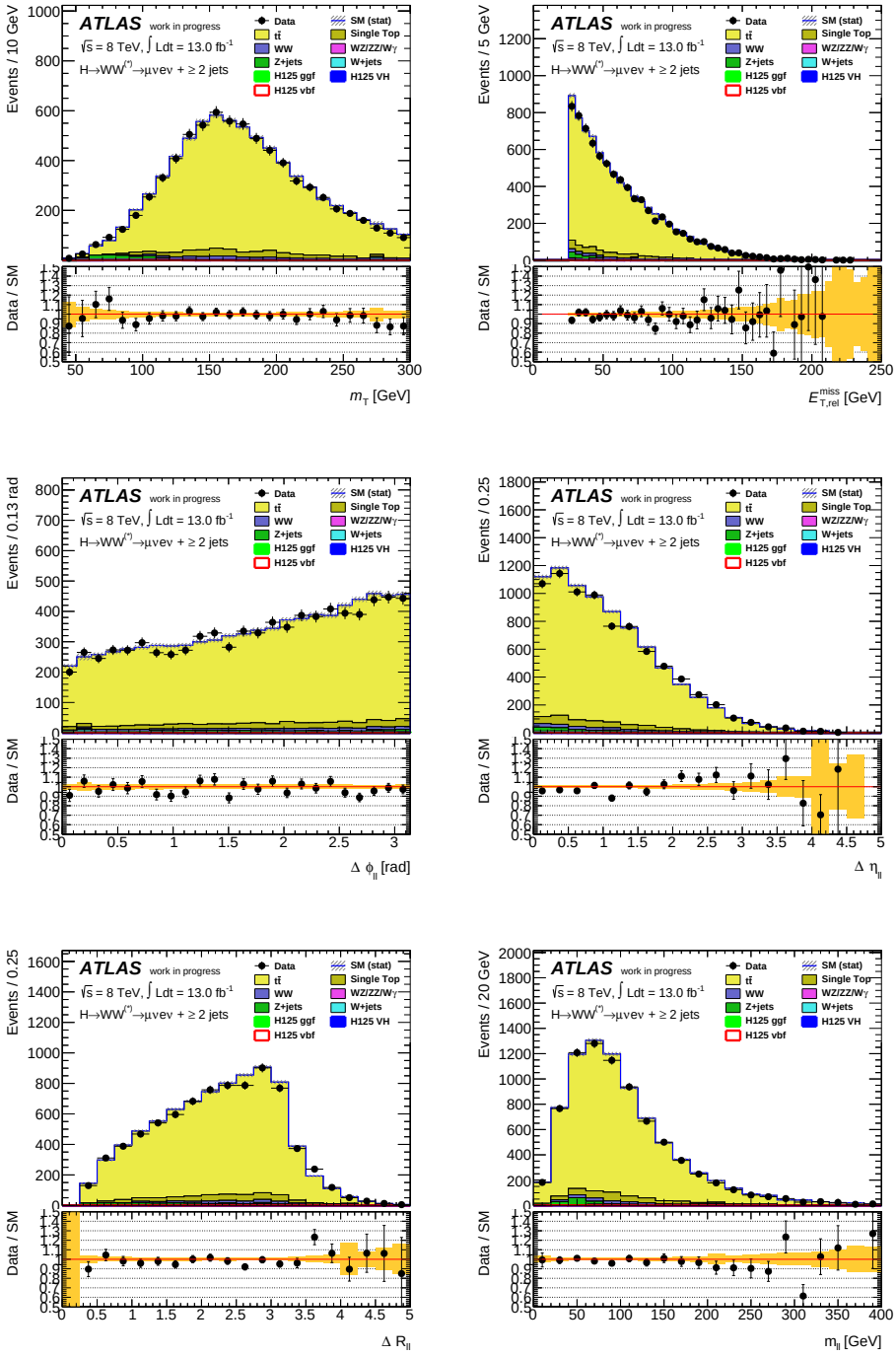


Figure 34: Input variables for the μe channel with statistical error band.

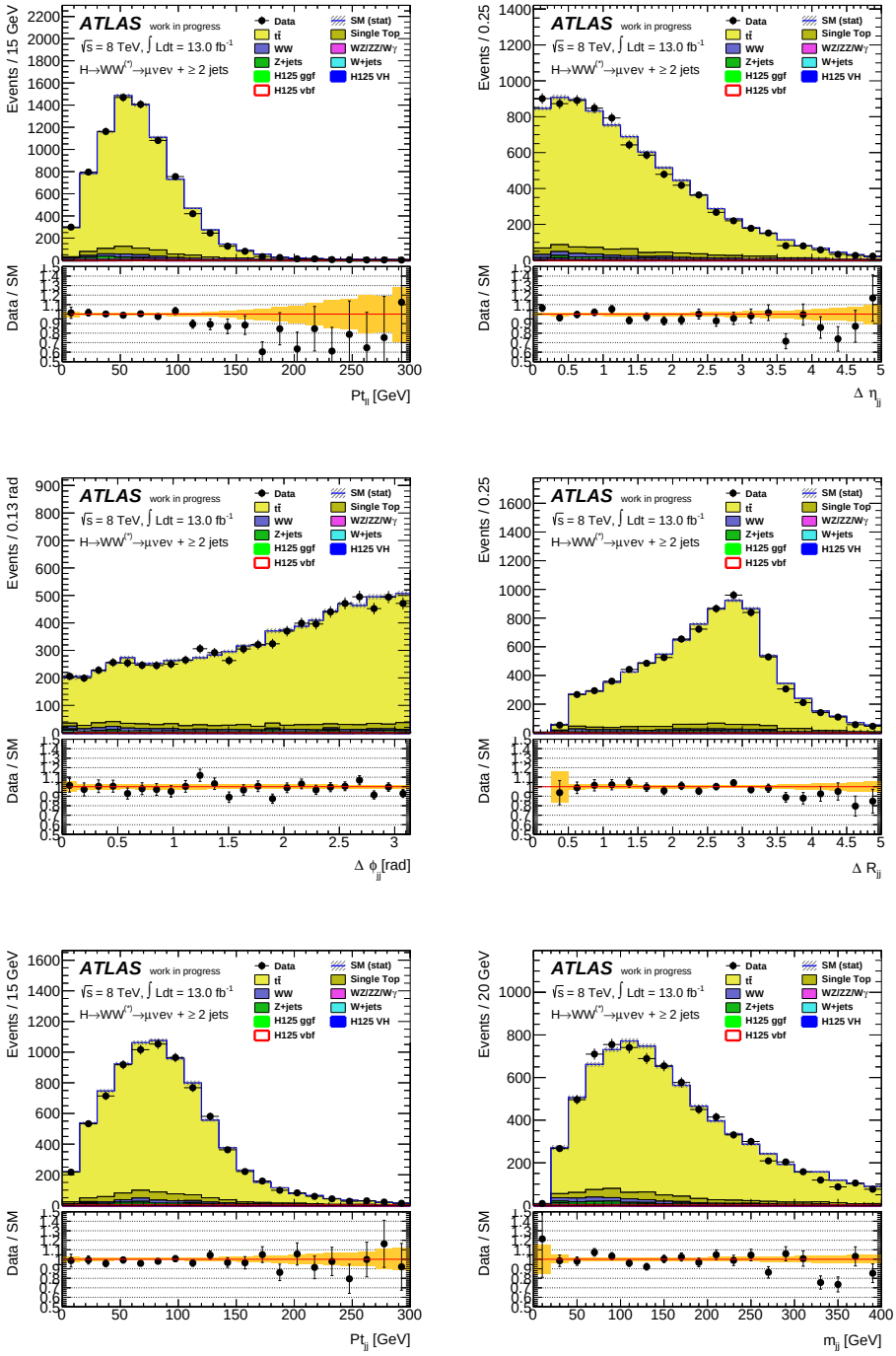


Figure 35: Input variables for the μe channel with statistical error band.

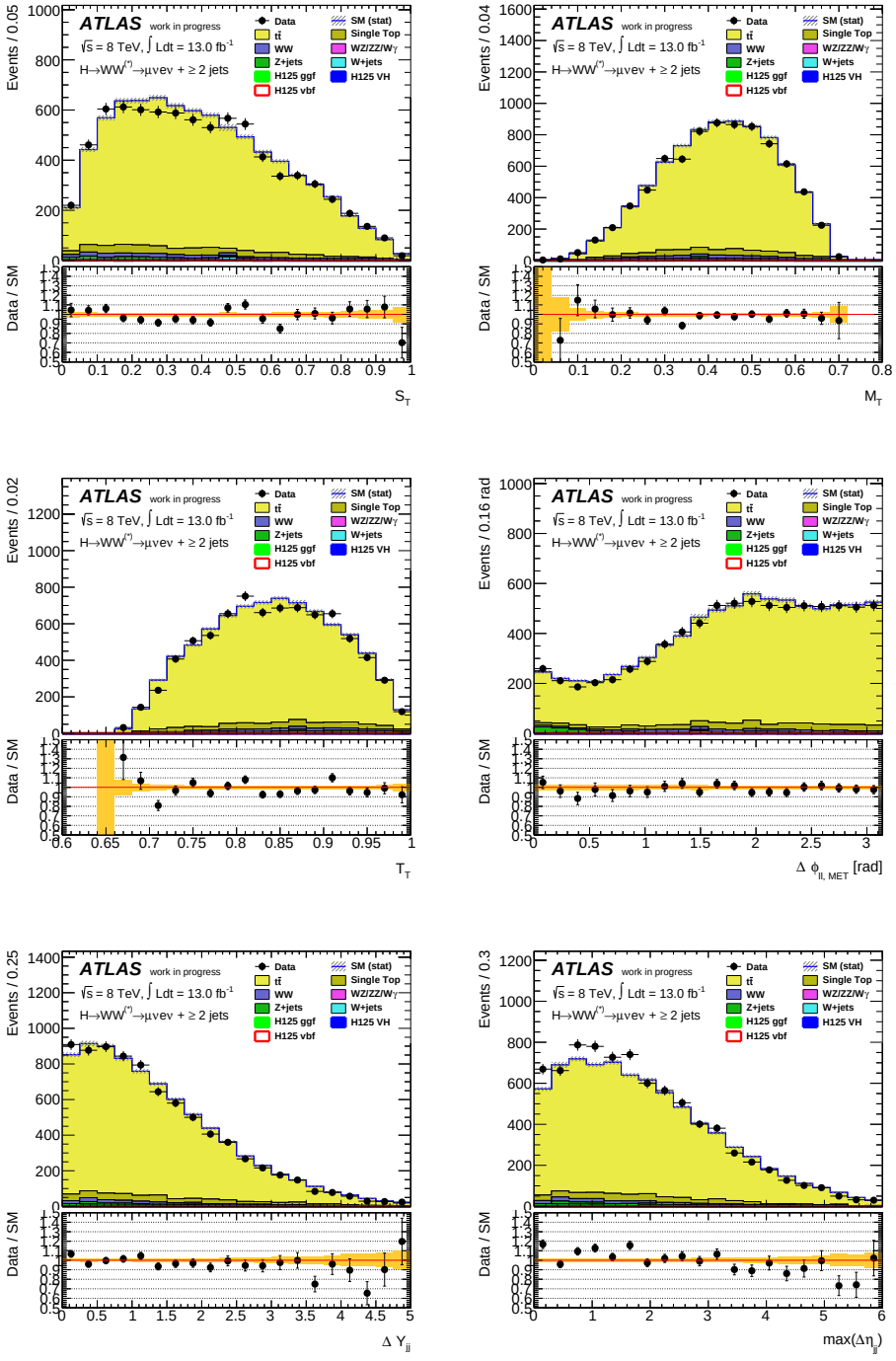


Figure 36: Input variables for the μe channel with statistical error band.

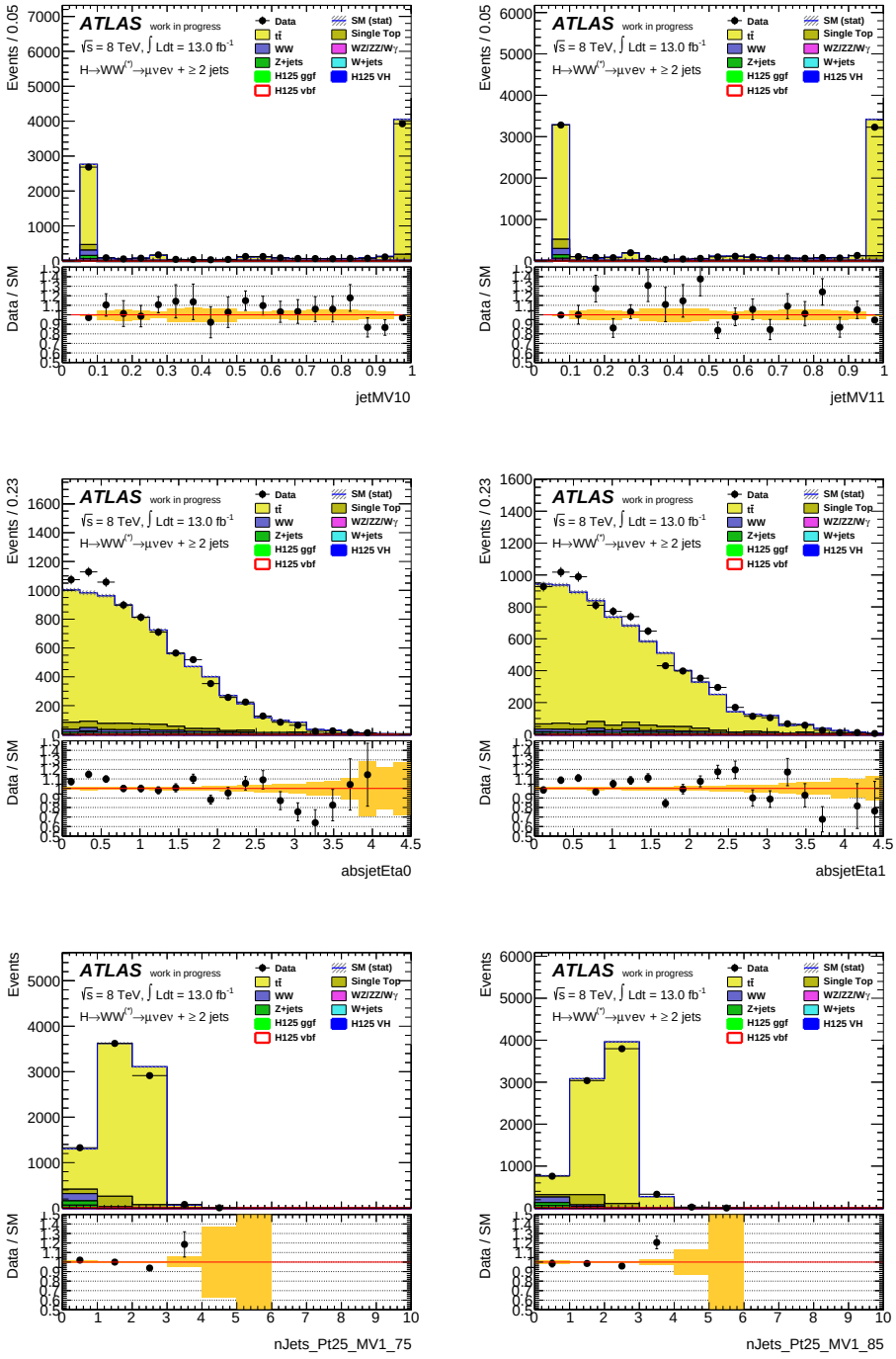


Figure 37: Input variables for the μe channel with statistical error band.

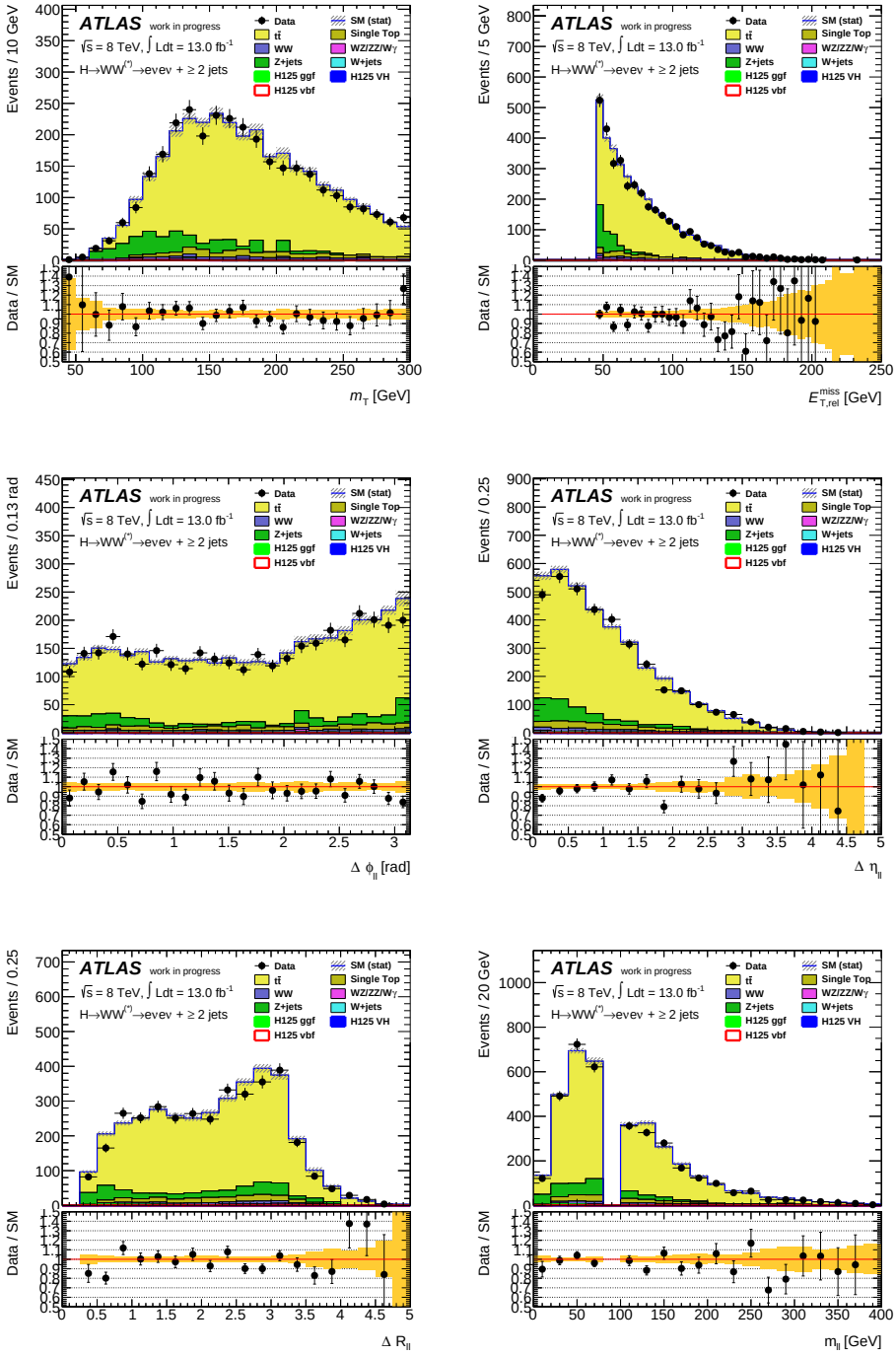


Figure 38: Input variables for the *ee* channel with statistical error band.

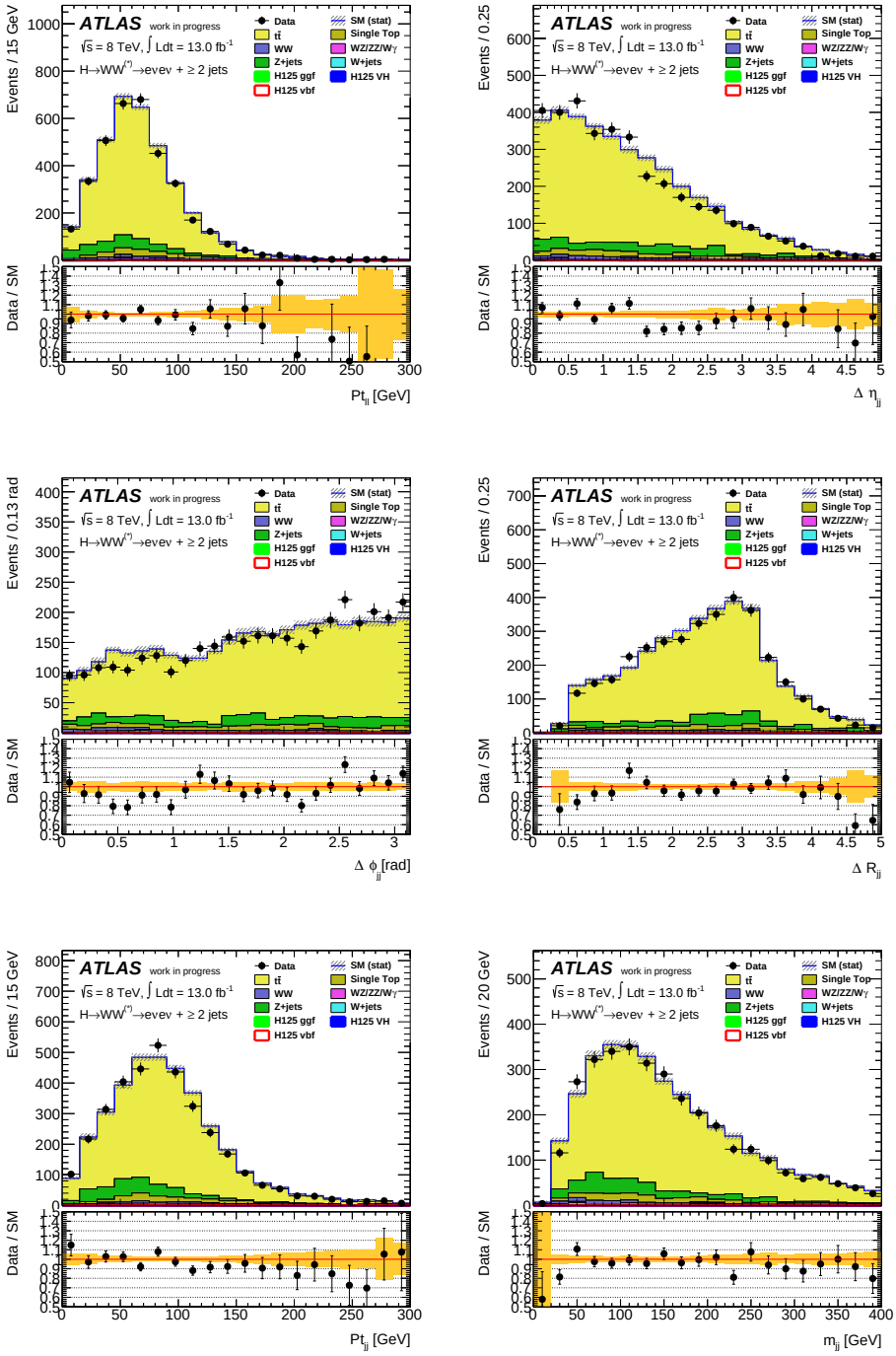


Figure 39: Input variables for the *ee* channel with statistical error band.

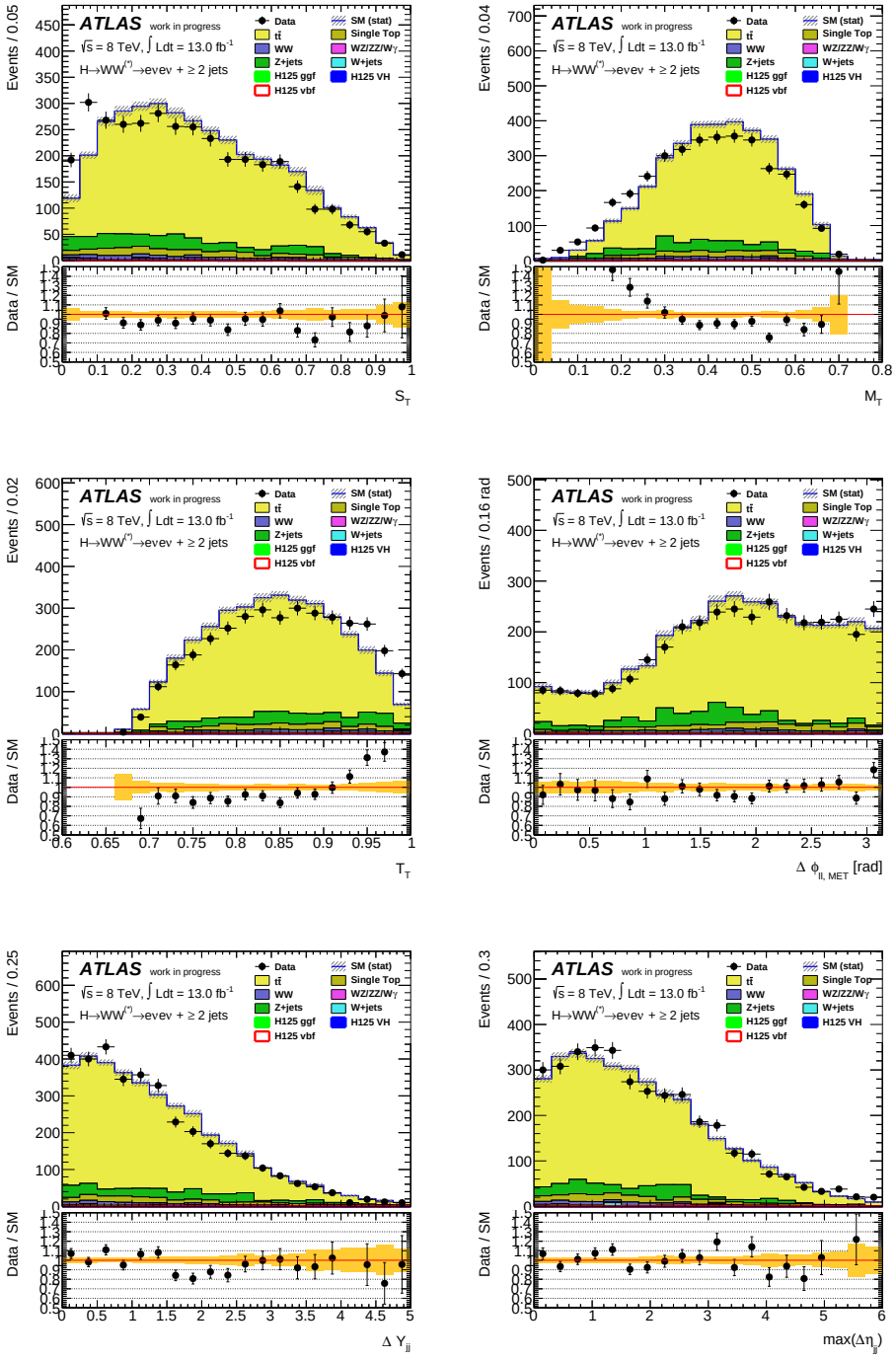


Figure 40: Input variables for the *ee* channel with statistical error band.

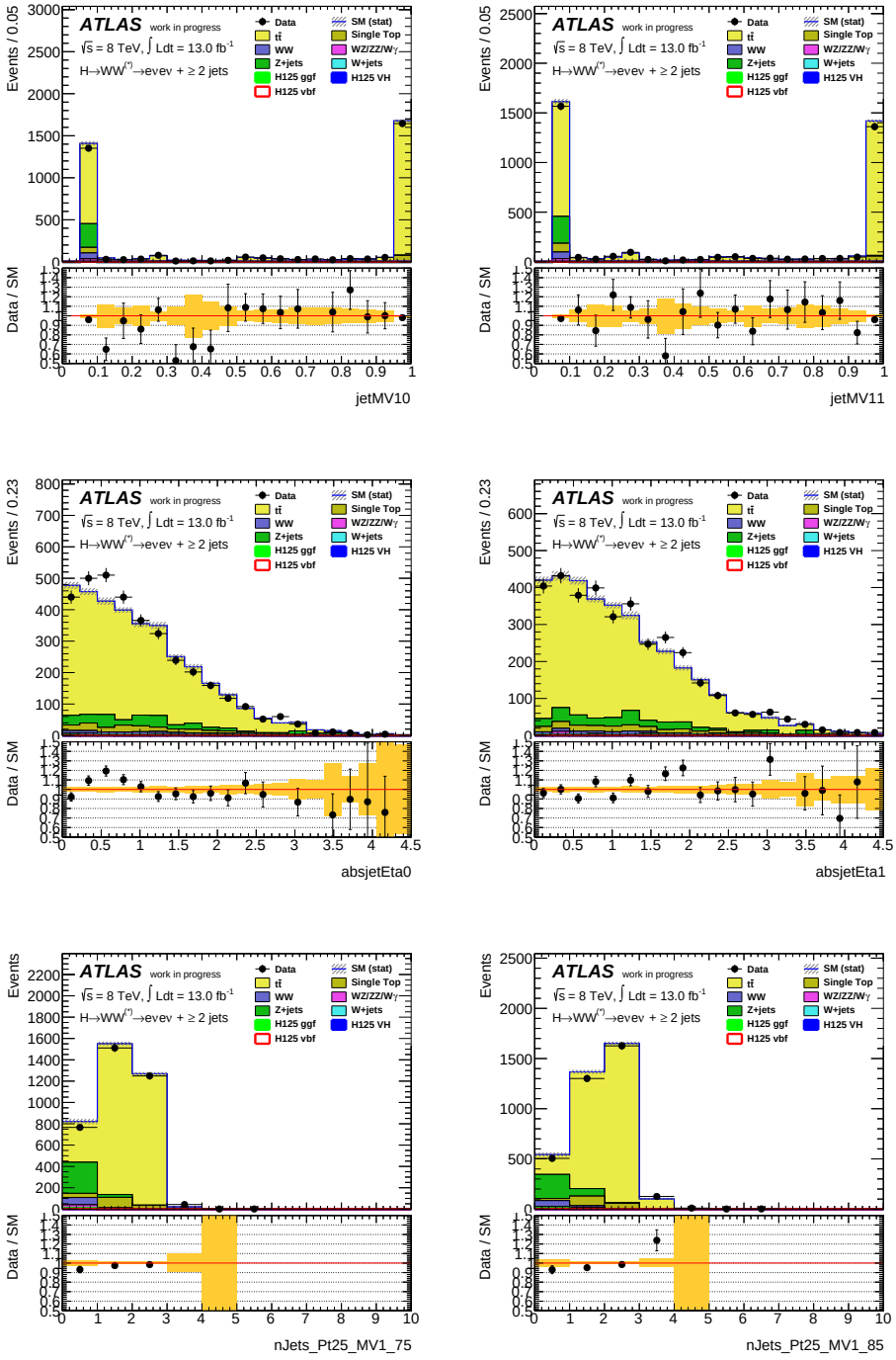


Figure 41: Input variables for the *ee* channel with statistical error band.

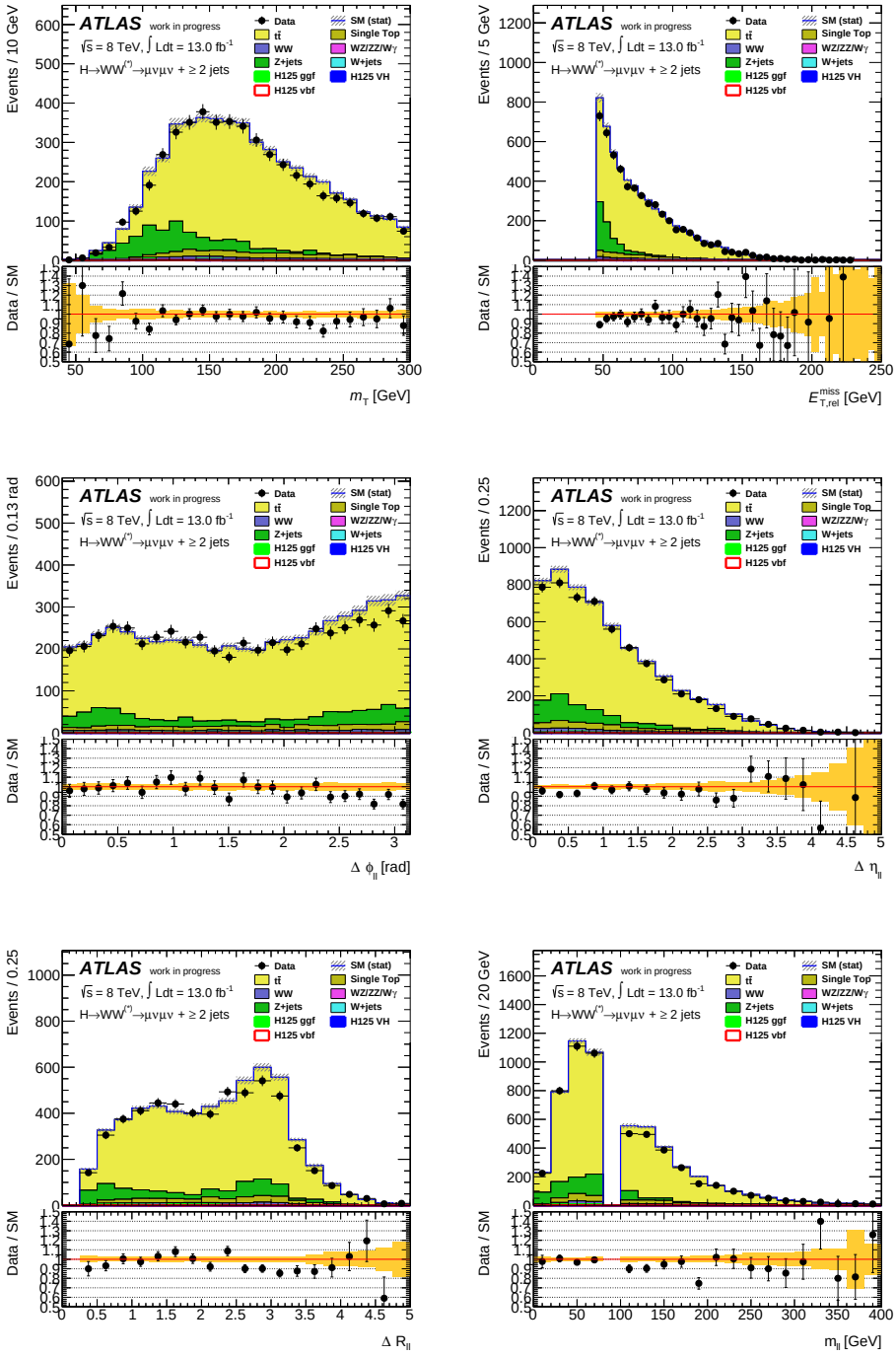


Figure 42: Input variables for the $\mu\mu$ channel with statistical error band.

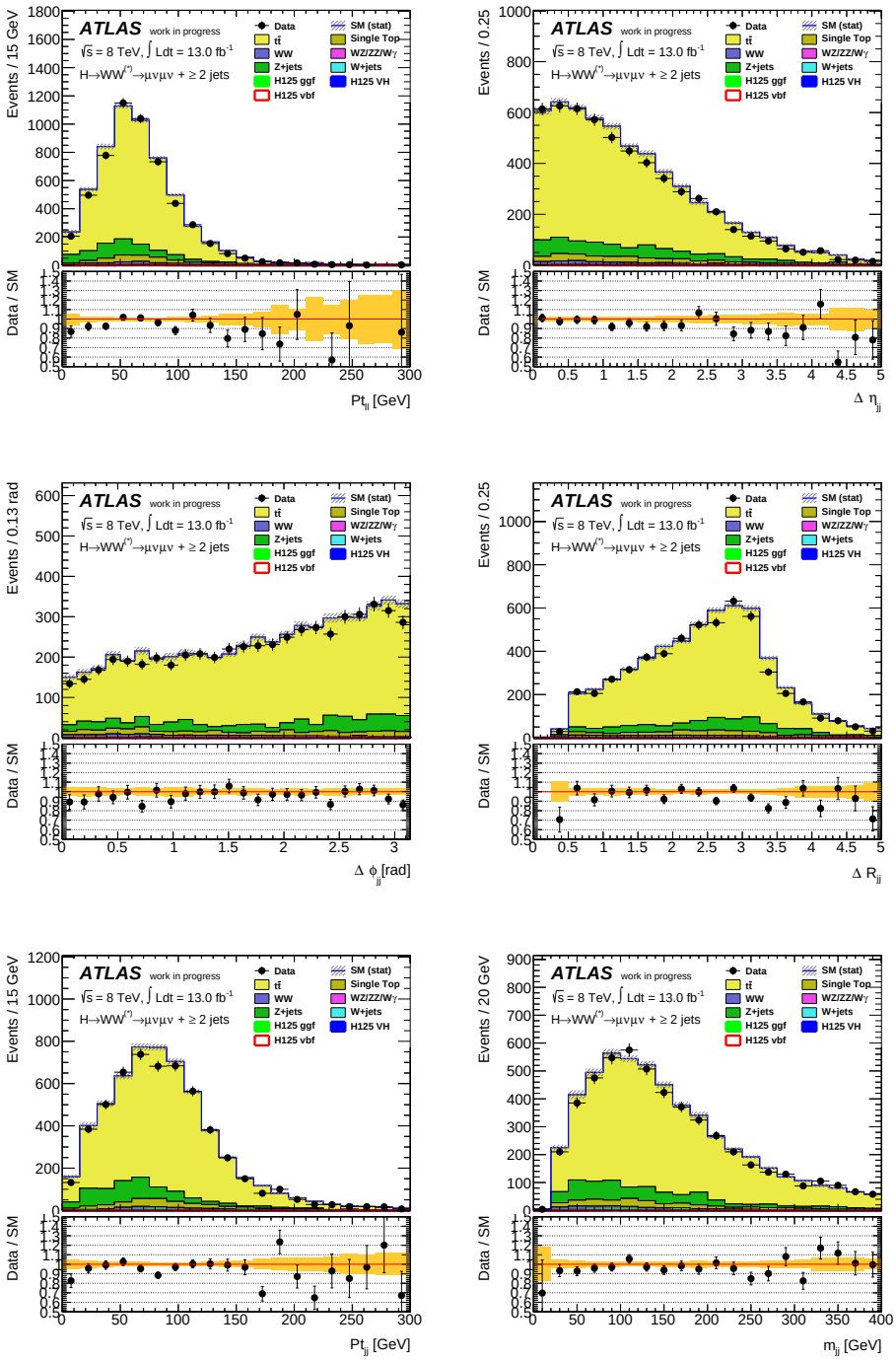


Figure 43: Input variables for the $\mu\mu$ channel with statistical error band.

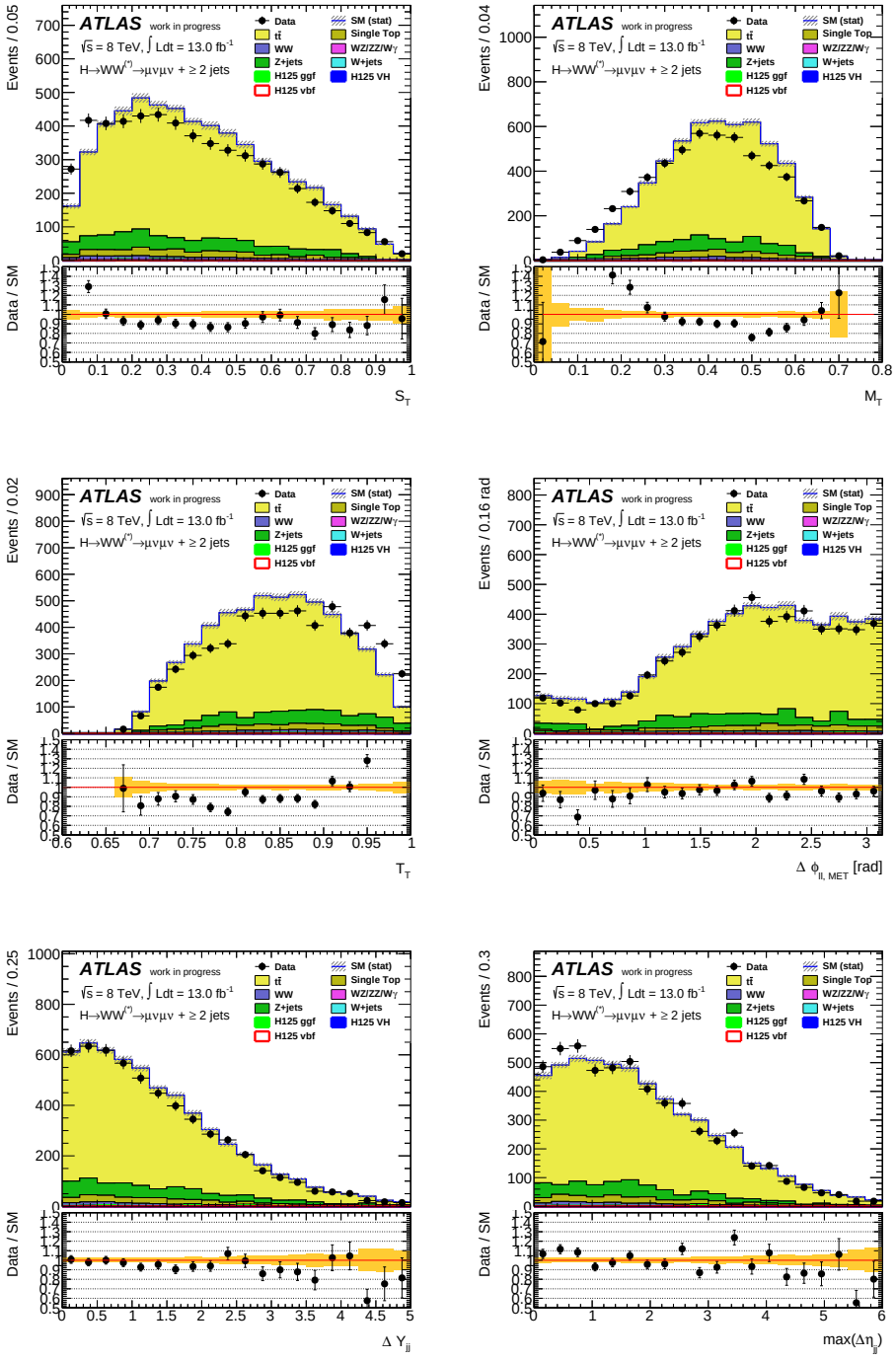


Figure 44: Input variables for the $\mu\mu$ channel with statistical error band.

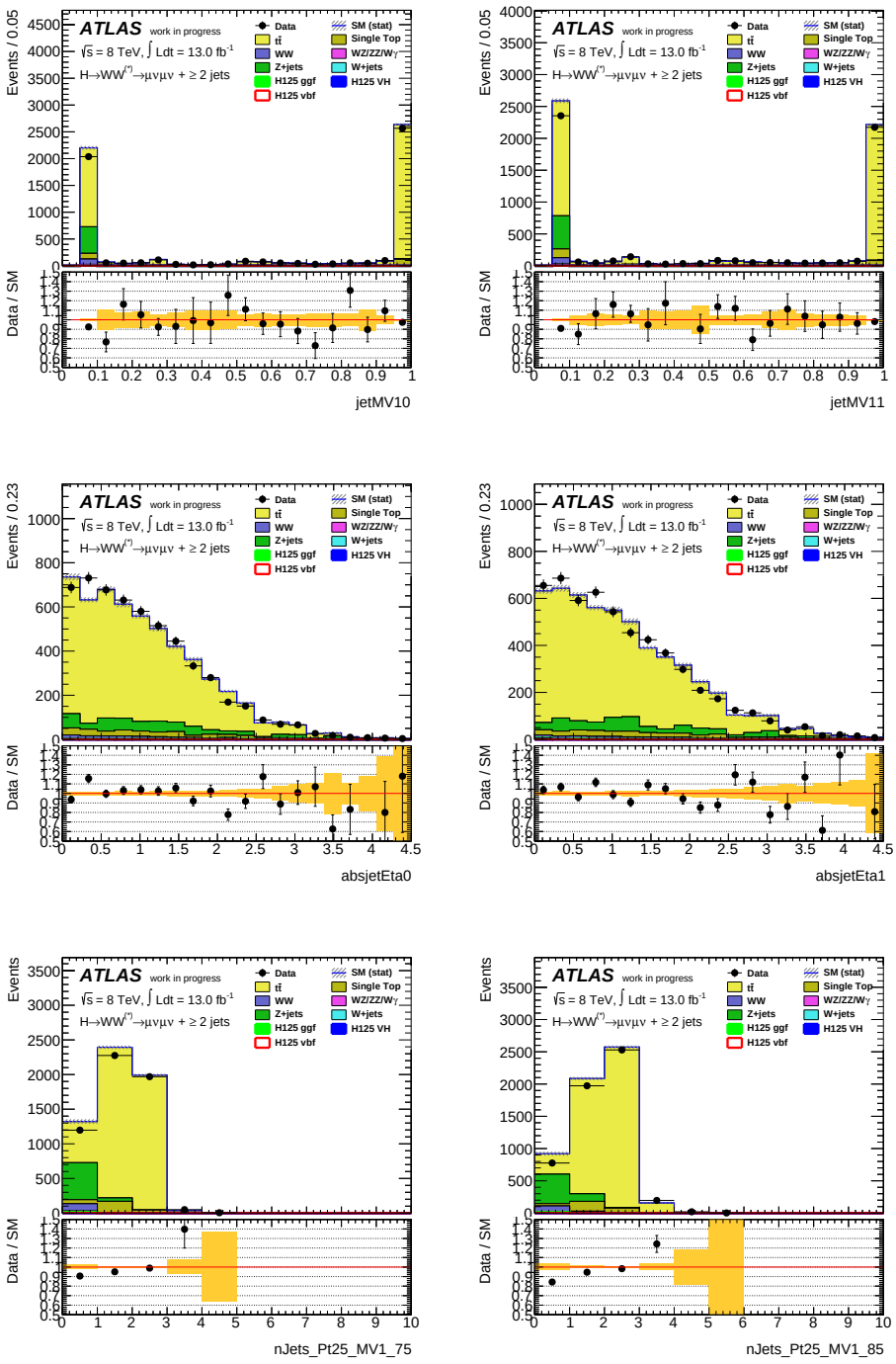


Figure 45: Input variables for the $\mu\mu$ channel with statistical error band.

References

[1] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," *Phys. Rev. Lett.*, vol. 13, pp. 508–509, Oct 1964.

[2] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons," *Phys. Rev. Lett.*, vol. 13, pp. 321–323, 1964.

[3] S. Chatrchyan *et al.*, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Phys.Lett.*, vol. B716, pp. 30–61, 2012.

[4] The ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," 2012.

[5] J. Beringer *et al.*, "Review of Particle Physics (RPP)," *Phys.Rev.*, vol. D86, p. 010001, 2012.

[6] D. Griffiths, *Introduction to Elementary Particles*. Physics Textbook, Wiley, 2008.

[7] F. Halzen and A. Martin, *Quarks and leptons: an introductory course in modern particle physics*. Wiley, 1984.

[8] J. Ebke, "The search for the Higgs Boson at the ATLAS Experiment using Multivariate Techniques," 2008.

[9] C. Burgess and G. Moore, *The Standard Model: A Primer*. Cambridge books online, Cambridge University Press, 2006.

[10] S. Dittmaier, S. Dittmaier, C. Mariotti, G. Passarino, R. Tanaka, *et al.*, "Handbook of LHC Higgs Cross Sections: 2. Differential Distributions," 2012.

[11] "SM Higgs production cross sections at $\sqrt{8}\,\mathrm{TeV}$." `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt8TeV`, 2012.

[12] "SM Higgs Branching Ratios and Partial-Decay Widths." `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageBR`, 2011.

[13] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, "TMVA: Toolkit for Multivariate Data Analysis," *PoS*, vol. ACAT, p. 040, 2007.

[14] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. Springer series in statistics, Springer, 2009.

[15] G. Aad *et al.*, "Measurement of charged-particle event shape variables in $\sqrt{s} = 7$ TeV proton-proton interactions with the ATLAS detector," 2012.

[16] Brüning, Oliver Sim and Collier, Paul and Lebrun, P and Myers, Stephen and Ostojic, Ranko and Poole, John and Proudlock, Paul, *LHC Design Report Vol. 2.* Geneva: CERN, 2004.

[17] Brüning, Oliver Sim and Collier, Paul and Lebrun, P and Myers, Stephen and Ostojic, Ranko and Poole, John and Proudlock, Paul, *LHC Design Report Vol. 1.* Geneva: CERN, 2004.

[18] G. Aad *et al.*, "The ATLAS Experiment at the CERN Large Hadron Collider," *JINST*, vol. 3, p. S08003, 2008.

[19] M. Lamprecht, "Studien zu Effizienz und AKzeptanz des ATLAS-Myontriggers mit simulierten Messdaten," 2007.

[20] A. Collaboration, "Object selection for the $H \to WW$ search with the ATLAS detector at $\sqrt{s} = 8\,\text{TeV}$," Tech. Rep. ATL-COM-PHYS-2012-861, CERN, Geneva, Jun 2012.

[21] "Expected electron performance in the ATLAS experiment," Tech. Rep. ATL-PHYS-PUB-2011-006, CERN, Geneva, Apr 2011.

[22] B. Resende, "Muon identification algorithms in ATLAS," Tech. Rep. ATL-PHYS-PROC-2009-113, CERN, Geneva, Sep 2009.

[23] M. Cacciari, G. P. Salam, and G. Soyez, "Anti-KT jet clustering Algorithm," *JHEP*, vol. 0804, p. 063, 2008.

[24] S. A. et al., "GEANT 4, A Simulation Toolkit," *Nucl. Instrum. Meth.*, vol. A506, p. 250, 2003.

[25] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations," *JHEP*, vol. 0206, p. 029, 2002.

[26] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, "ALPGEN, a generator for hard multiparton processes in hadronic collisions," *JHEP*, vol. 0307, p. 001, 2003.

[27] B. P. Kersevan and E. Richter-Was, "The Monte Carlo event generator AcerMC version 2.0 with interfaces to PYTHIA 6.2 and HERWIG 6.5," 2004.

[28] T. Binoth, M. Ciccolini, N. Kauer, and M. Kramer, "Gluon-induced W-boson pair production at the LHC," *JHEP*, vol. 0612, p. 046, 2006.

[29] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1," *Comput.Phys.Commun.*, vol. 178, pp. 852–867, 2008.

[30] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX," *JHEP*, vol. 1006, p. 043, 2010.

[31] A. Collaboration, "Search for the Standard Model Higgs boson in the $H \to WW^{(*)} \to \ell\nu\ell\nu$ decay mode using Multivariate Techniques with $4.7\,\mathrm{fb}^{-1}$ of ATLAS data at $\sqrt{s} = 7\,\mathrm{TeV}$," 2012.

[32] "Measurements of the properties of the Higgs-like boson in the $WW^{(*)} \to \ell\nu\ell\nu$ decay channel with the ATLAS detector using 25 fb^{-1} of proton-proton collision data," Tech. Rep. ATLAS-CONF-2013-030, CERN, Geneva, Mar 2013.

Acknowledgements

Ich möchte mich bei folgenden Personen bedanken.

- Prof. Dr. Dorothee Schaile für die Möglichkeit diese Masterarbeit in der experimentellen Teilchenphysik zu erstellen.
- Dr. Johannes Elmsheuser für die gute Betreuung, Ratschläge und Geduld.
- Dr. Johannes Ebke für Hinweise und Tipps zu BDTs.
- Bonnie Chow, Christian Meineck und Friedrich Hönig für das gute Arbeitsklima im Büro.
- Meiner Familie für die stetige Unterstützung.

Erklärung

Hiermit erkläre ich, die vorliegende Arbeit selbständig verfasst zu haben und keine anderen als die in der Arbeit angegebenen Quellen und Hilfsmittel benutzt zu haben.

München, den 11.4.2013

_____

(Philipp Heimpel)