Boosted decision tree performance studies in the $t\bar{t} \rightarrow \tau + \mu$ analysis at ATLAS

Allison MacDonald

Supervised by Dr. Dugan O'Neil

Department of Physics, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6

September 17, 2011

Abstract

The cross section for top quark pair production with $\tau + \mu$ final states in proton-proton collisions at $\sqrt{s}$=7 TeV is measured at the ATLAS detector. We use a multivariate technique called a boosted decision tree to separate signal from background and extract the cross section by fitting signal and background templates to the output of this boosted decision tree. We study the performance of the boosted decision tree output distribution under various conditions to verify whether the shape of the signal or background distributions changes significantly.

1 Introduction

In order to probe increasingly smaller distances in our universe, we need to increase the energy scale at which we are working. As this was done in the past, scientists were able to discover substructure in the atom, then within the nucleus, and finally within the nucleons (protons and neutrons). Today, we believe the fundamental matter particles are leptons and quarks. Forces are interactions between the quarks and leptons which are mediated by a third type of particle known as gauge bosons.

The Standard Model (SM) of particle physics is a very well-tested theory describing these fundamental particles and the interactions between them. It describes three of the four fundamental forces (strong, weak and electromagnetic) as the exchange of a gauge boson between two interacting matter fermions, with the interaction length being inversely related to the mass of the force-carrying boson. Hence, massless bosons carry forces which have an infinite range, while the range of the interaction of a massive boson is limited. It has also succeeded in unifying two of these three forces (electromagnetic and weak). So far, all the particles predicted by the standard model have been observed except for the Higgs boson. The Higgs boson is a consequence of the Higgs mechanism which allows the weak force-carrying W and Z bosons to couple to the Higgs field, thereby giving them mass. If this is true, then the Higgs field should also be able to couple to itself, creating a massive spin-0 boson. Without this mechanism, the standard model predicts that all elementary particles should be massless, which is contrary to experimental observations.

One of the most important quests in particle and high energy physics today is to confirm or exclude the existence of the Higgs boson. Since we don't know its precise mass, this is quite challenging, but with the start-up of the Large Hadron Collider (LHC), we can produce particles whose mass would be in the required energy range. If the Higgs boson does not exist, we must turn to other theories which would explain the origin of mass in our universe such as super symmetry (SUSY). In these cases, we look for signatures of such beyond-the-standard-model physics (BSM) in the events that we observe at the LHC.

1.1 The Large Hadron Collider

The LHC at the European Organization for Nuclear Research (CERN) is the world's most energetic particle collider, currently accelerating and colliding protons at centre-of-mass energies of 7 TeV. The LHC is a circular collider, with the protons traveling along a 27 km ring 100 m underground crossing the French-Swiss border near Geneva. The protons collide at one of four collision points along this ring every 25 ns, and these points are the locations of the major experiments: A Large Ion Collider Experiment (ALICE), LHCb, the Compact Muon Solenoid (CMS), and A Toroidal LHC ApparatuS (ATLAS). ALICE investigates the nature of strong interactions through heavy ion (Pb-Pb) collisions as well as proton collisions, in the hopes of observing phase transitions to quark-gluon plasma. LHCb studies the asymmetries between matter and antimatter by observing decays of B mesons, bound states of bottom quarks and other quarks, which are produced when the quarks and gluons inside the protons collide. CMS and ATLAS are both large-scale general particle detectors with many layers which look to detect signatures of interesting physics such as the Higgs boson, super symmetry (SUSY) or dark matter. In addition to these four large experiments, there are two smaller experiments, LHCf and TOTEM, which are located a short distance away from the ATLAS and CMS detectors, respectively, and study the structure of the

protons and their interactions.

1.2 Structure of the ATLAS Detector

The ATLAS detector consists of several layers, moving outward from the beam axis, which detect particles in different ways. Particles interact with different materials in the detector depending on their interactions (electromagnetic, strong, weak) and lifetimes. The innermost layer consists of a tracking system, which is composed of three sub-detectors: pixel detectors, the silicon SemiConductor Tracker strip detectors and the Transition Radiation Tracker [1, 2]. These three sub-detectors work together to record the tracks of charged particles as they are bent by a 2 Tesla magnetic field. The granularity is highest closest to the beam axis, in the pixel detectors, and decreases outwards to the TRT detectors which provide a more continuous measurement through long straw-like sensors which contain a wire and a mixture of Xe, CO_2 and O_2 gases. The momentum and charge of these particles can be measured here through the curvature of their tracks.

The calorimeters measure the energy of most strong- and electromagnetically-interacting particles. The electromagnetic (EM) calorimeter is located towards the centre of the detector and consists of metal plates and liquid argon (LAr). When a photon or an electrically charged particle hits the plates of the EM calorimeter, it causes a cascade of electrons and photons to be released. These electrons and photons are then detected by the LAr. The hadronic calorimeter is situated outside the EM calorimeter and is meant to detect strongly interacting particles such as hadrons, which were not stopped in the EM calorimeter. The hadronic calorimeter consists of steel plates and scintillating plastic tiles which emit photons when a particle passes through.

The final part of the detector is the muon spectrometer. Muons are able to traverse all parts of the inner detector and calorimetry without being absorbed so the muon spectrometer is located outside the other elements. It consists of large superconducting toroidal magnets which bend the muons and sensors which are similar to the TRT straws in the inner detector. This allows the tracks of the muons to be measured precisely.

1.3 Data Acquisition and Reconstruction

The LHC proton beams consists of bunches of 10^{11} protons, which are separated by approximately 25 ns. At a beam luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$, this amounts to 40 million bunch crossings per second and approximately 1 billion collisions per second [2]. This is far too much data to save, store and process so ATLAS (as with the other detectors) uses a multi-level trigger system to decide which collision events are of interest. Triggering cuts on events are made based on the type of signals that are desired. This varies from analysis to analysis so not all collected data is used by all analyses. After three levels of triggering, 200 events per second are stored permanently for future reconstruction and analysis by individual groups.

Once the event data is collected, it is reconstructed with software called Athena. Athena converts electronic signals from the ATLAS detector to physical objects and kinematical information which can be used in physics analyses.

1.4 Physics Object Definitions

Different physics objects interact differently with the ATLAS detector and so are reconstructed differently

by the ATHENA software. Electrons are generally leave tracks in the inner detector system and are stopped in the electromagnetic calorimeter. Muons have a greater penetration depth and are reconstructed by looking at tracks and energy deposits in the muon spectrometer as well as tracks in the inner detector. Taus, with a mass of 1776.84±0.17 MeV, are much heavier than the other charged leptons and consequently decay before they can be directly detected in ATLAS, with an average lifetime of $290.5\pm1.0\times10^{-15}$ seconds [3]. Taus can decay either leptonically or hadronically. However, leptonically decaying taus are very difficult to distinguish from primary electrons or muons so the ATLAS reconstruction is limited to hadronically decaying taus. These are separated into taus with one charged decay daughter (τ_1) and taus with 3 charged decay daughters (τ_3). Similarly, quarks and gluons cannot be directly detected due to colour confinement, which prevents particles with a colour charge from being isolated. Instead, a single quark or gluon undergoes hadronization and is detected in ATLAS as a narrow shower of hadronic particles. Finally, many decays produce neutrinos, the neutral partners of the charged leptons. Neutrinos are so weakly interacting that they pass completely through the ATLAS detector and their presence can only be inferred from missing energy, E_T^{miss}.

1.5 Monte Carlo Simulations

An important tool in understanding the performance of the detector components and the processes we expect to occur are simulations of the detector. In ATLAS, we use Monte Carlo (MC) simulations to create pp collision events and then allow the products to decay with the distributions of energy, location, etc. that we expect from theory. We then simulate the detection of these particles in the detector system including, for example, the detection efficiencies and any known problems with the detector. The simulated detections are then reconstructed with the same reconstruction software used on the real data.

Because the detected particles are created in a simulation, we can store truth information in the MC samples about the particles generated. For example, we can trace the parent and daughter(s) of a particle as well as access its true kinematic properties (energy, location in the detector, momentum) using the truth properties. We can then compare this information with the reconstruction information to determine how well we can separate signal from background among the reconstructed particles.

2 $\mathbf{t\bar{t}} \rightarrow \tau + \mu$

When a pair of top quarks ($\mathrm{t\bar{t}}$) is produced in a proton-proton collision, each top decays almost exclusively to a W boson and a bottom quark (b or $\bar{\mathrm{b}}$). The W then decays further, either hadronically to a quark-antiquark pair ($W \rightarrow \mathrm{q\bar{q}}$), or leptonically to a charged lepton (electron, muon or tau) and its associated neutrino ($W \rightarrow l\nu$). We investigate the channel in which one top decays to a tau and the other to a muon. This channel is a promising venue to observe new physics. For example, the minimal supersymmetric standard model (MSSM) requires the existence of a charged Higgs boson ($\mathrm{H}^{\pm}$) [4]. If the mass of this Higgs is smaller than the mass of the top quark less the mass of the b quark ($\mathrm{M}_H+\mathrm{M}_b < \mathrm{M}_t$), then it is possible for the top to decay to the Higgs and a b quark ($t \rightarrow H^{\pm} + b$). In MSSM, the $H^{\pm}$ would couple more strongly to the tau than to the other charged leptons, resulting in an increase in the branching ratio for $\mathrm{t\bar{t}}$ pairs to final states containing τs, including the $\mathrm{t\bar{t}} \rightarrow \tau + \mu$

process. The cross section times the branching ratio ($\sigma(t\bar{t})\times BR(t \to \tau\nu + b)$) of this channel has been measured before, and found to be within the expectations of the standard model; however the most precise measurement (at the Tevatron) carried an uncertainty of 25% [5]. The high energy of the LHC results in an increased cross section for $t\bar{t}$ pair production which will allow the cross section of $t\bar{t} \to \tau + \mu$ to be measured with higher precision.

We select our signal from the data by identifying at least one b-jet and applying the following pre-selection cuts to events [6]:

- The event has a primary vertext with at least 5 tracks
- All jets with p_T >20 GeV must pass jet quality tests
- Exactly one identified muon
- At least one loose τ candidate
- At least two jet candidates, which do not fall within R=0.4 of the τ candidate
- E_T^{miss} >30 GeV
- $\Sigma|E_T|$ >200 GeV
- The τ and μ have opposite sign. The data are split into opposite sign (OS) and same sign (SS) samples which are used for reducing quantum chromodynamical (QCD) background

Loose τ candidates are defined by the following criteria:

- E_T >20 GeV
- The leading track associated with the τ candidate must have p_T >4 GeV
- The τ candidate must fall in the pseudorapidity range $|\eta|$ <2.3 where pseudorapidity is defined as $\eta = -\ln(\tan(\frac{\theta}{2}))$ with θ being the angle between a particle's momentum and the beam axis
- The total charge of the tracks has absolute value 1 or 2

B-jets can be identified using a b-tagging algorithm. Since hadrons which contain b quarks are long-lived, b-jets contain secondary vertices from these b hadrons that can be separated from the primary vertex of the jet. B-tagging takes advantage of this by identifying secondary vertices associated with this decay.

The effects of these preselection criteria are illustrated in Tables 1 and 2. We can see that after applying these cuts and the *b*-tagging requirements, the primary backgrounds are $t\bar{t}$ processes which contain final states other than $\tau + \mu$. These are primarily a result of jets faking taus and electrons faking taus. We then exploit tau identification (TauID) to separate these backgrounds from our signal.

Cut	$t\bar{t}(\mu,\tau)$	$t\bar{t}(l$+jets)	$t\bar{t}(ll')$	W+jets	Z+jets	Single top	Diboson	Total	Data
Trigger	2693±10	20880±40	4679±10	5680000±6800	867000±630	7692±40	8212±40	6591500±6900	8872361
Isolated μ	2243±10	13700±20	2064±10	5419300±7000	416800±460	6316±40	6453±40	5866900±7000	8113657
$\geq 1\tau$ candidate	497±5	2042±10	131±2	74340±660	13740±80	470±10	680±10	91900±670	154513
$N_{jet} \geq 2$	401±4	1941±10	103±2	7330±90	1516±30	227±10	169±10	11690±100	16385
E_T^{miss} >30 GeV	351±4	1562±10	93±2	5237±80	654±20	178±10	117±103	8191±80	8930
H_T >200 GeV	346±4	1553±10	92±2	4332±60	514±20	170±5	102±5	7109±60	7304
$\geq$ b-jet	235±3	979±10	56±1	163±10	17±3	87±3	6±1	1543±10	1593

Table 1: Cutflow table for 1-prong taus with 1.08 fb^{-1} of ATLAS data [6].

Cut	$t\bar{t}(\mu,\tau)$	$t\bar{t}(l$+jets)	$t\bar{t}(ll')$	W+jets	Z+jets	Single top	Diboson	Total	Data
Trigger	2693±10	20880±40	4679±10	5680000±6800	867000±630	7692±40	8212±40	6591500±6900	8872361
Isolated μ	2243±10	13700±20	2064±10	5419300±7000	416800±460	6316±40	6453±40	5866900±7000	8113657
$\geq 1\tau$ candidate	500±5	5100±10	437±4	220300±1100	22350±110	1310±20	1412±20	251400±1200	503283
$N_{jet} \geq 2$	371±4	4794±10	331±4	19900±160	2896±40	583±10	336±10	29210±160	44759
E_T^{miss} >30 GeV	326±4	3850±10	297±3	14400±130	1134±20	461±10	230±10	20700±130	23932
H_T >200 GeV	321±4	3823±10	293±3	11860±100	922±20	441±10	200±10	17900±110	19522
$\geq$ b-jet	206±3	2465±10	182±3	535±30	33±4	226±10	13±2	3660±30	4086

Table 2: Cutflow table for 3-prong taus with 1.08 fb^{-1} of ATLAS data [6].

2.1 Tau Identification: Boosted Decision Trees

Since τ candidates are reconstructed from their decay products in the hadronic channel, they closely resemble both jets, which consist of a collection of mesons and baryons in a cone-like shape, and electrons, which can create similar showers of particles in the calorimeter when they are detected. Although there are several variables which can differentiate τs from jets and τs from electrons, we find that none of these provides enough discriminating power to separate the large background from our relatively weak signal. We thus use a multivariate technique called a boosted decision tree (BDT) that combines several of these variables [7].

BDTs consist of several decision trees, each of which works by making a simple cut on a single variable at each decision node in the tree. An individual candidate is then passed to one of two daughter nodes, depending on whether it passed or failed the cut. The passage of a candidate through the tree ends when it reaches a leaf node. Here, a value between 0 and 1 is assigned to the candidate; this is the output of an individual decision tree. In our analysis, values closer to 1 indicate that the candidate is very τ-like while 0 indicates the candidate is more background-like [7].

Each decision tree is trained on a MC sample of truth-matched signal and background. After the samples are passed through the first tree, the performance of the tree is evaluated. The MC samples are then reweighted to put more weight on the candidates that were misclassified. Another tree is created and it is trained on the MC samples with their new weights. This is called boosting.

In our analysis, we use two separate BDTs: one to reject jets (the BDTj), which is trained on samples of real τs and real jets faking τs, and one to reject electrons (BDTe), which is trained on samples of real τs and real electrons faking τs. We also train the BDTjs separately on τ_1 and τ_3 candidates. The variables used in our BDTs are described in Table 3.

2.2 Extracting the Cross Section

We use the BDTs described above in two ways: we cut on the BDTe score to separate τs and jets faking τs from electrons faking τs and then we fit the shape of the BDTj output using templates for our signal and jet background. We then extract the cross section by counting the number of events in the signal region. There is a large QCD background so we use a method which takes advantage of the charge correlation in both gluon and quark jets [6]. An example of the templates used to fit the shape of the BDTj output is shown in Figure 1.

3 BDT Performance Studies

Since our analysis relies heavily on the shape of the BDTj output, it is very important to understand if and how that shape changes with changes in the parameters. To this end, a series of small studies were conducted to examine the effects of the number of reconstructed tracks, an error in the polarization of true taus in the $t\bar{t}$ MC samples and pileup. We also examined the performance of the BDTe. These studies were conducted in MC only.

Variable	Definition	BDTj τ_1	BDTj τ_3	BDTe
Centrality fraction $\Sigma_i E_{T,i}/\Sigma_j E_{T,j}$	i and j run over all clusters within ΔR=0.1 and ΔR=0.4 of the centre, respectively	√	√	√
$p_T^{\text{Lead Track}}/E_T$	Characterizes how much of the total momentum is carried by the leading track (large for τs)	√	√	√
M_{tracks}	Total invariant mass of the tracks	-	√	-
D_{avg}^{tracks}	The average distance of the tracks from the jet axis	√	√	√
Transverse flight path significance, $L_{xy}/\sigma_{L_{xy}}$	Displacement, L_{xy} of the vertex from which the tracks originate divided by the uncertainty $\sigma_{L_{xy}}$	-	√	-
Cluster Invariant Mass	Mass calculated from four vectors of calorimetric clusters contained in the tau cone	√	√	-
Impact parameter significance of the lead track	Impact parameter d_0 divided by its uncertainty	√	-	-
$N_{\text{wide tracks}}$	The number of tracks in the ring $0.2 < \Delta R < 0.4$ of the tau cone axis	√	√	-
R_{calo}	The radius of the cone in the calorimeter	√	√	-
E(lead 3 clusters)/E(all clusters)	Energy of the 3 most energetic clusters over the total energy of all clusters	√	√	-
ΔR_{max}	Largest separation between associated tracks	-	√	-
Hadronic Leak Energy	Ratio of hadronic E_T to p_T of the leading track	-	-	√
Electromagnetic Fraction $\Sigma_{i\epsilon\{EM0-2\}} E_{T,i}^{EM}/\Sigma_{j\epsilon\{all\}} E_{T,j}^{EM}$	Fraction of tau candidate's transverse energy deposited in the EM calorimeter. i, j run over the layers of the calorimeter	-	-	√
$E_{T,max}^{strip,sec}$	Maximum E_T in the strip detector not associated with the leading track	-	-	√
$N_{high}^{TRT}/N_{low}^{TRT}$	Number of high threshold hits in the TRT divided by the number of low threshold hits	-	-	√
Isolation fraction $\Sigma_i E_{T,i}^{EM}/\Sigma_j E_{T,j}^{EM}$	Fraction of E_T in a ring outside the centre of the tau. i runs over $0.1 < \Delta R < 0.2$; j runs over $\Delta R < 0.4$	-	-	√
Hadronic Radius (R_{Had}) $\Sigma_i E_{T,i}^{EM} \Delta R_i/\Sigma_i E_{T,i}^{EM}$	Transverse energy-weighted width of the shower in the hadronic calorimeter. i runs over cells within ΔR =0.4 of the centre in the first three layers	-	-	√

Table 3: List of the variables used in each of our trained BDTs, along with their definitions

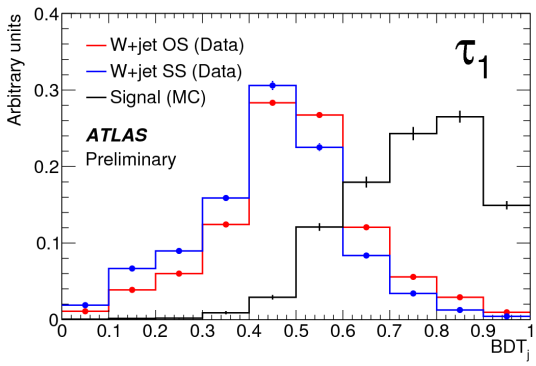


Figure 1: One Prong BDTj templates for signal, OS and SS background [6]

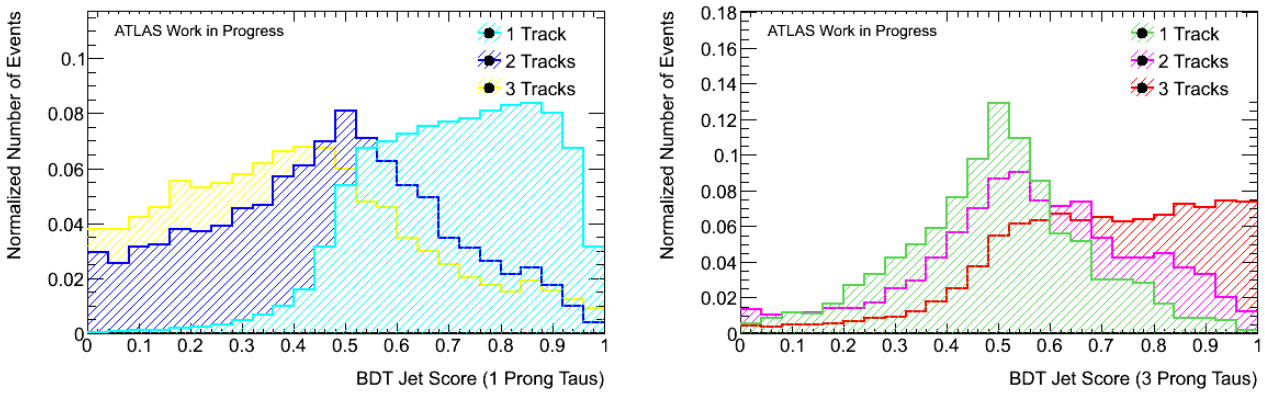
3.1 BDTj Performance on τ_1 and τ_3

The BDTj we use actually consists of two BDTs: one for τ_1 candidates and one for τ_3 candidates. The real τs used in training are chosen such that the reconstructed number of tracks matches the true number of prongs. However, in data, we have no access to truth information so it is important to understand how τ candidates behave both as a function of the number of reconstructed tracks and the number of true prongs. We are especially interested in understanding how τs for which the reconstruction information does not match the truth information behave.

Truth information in a $t\bar{t}$ MC sample was used to select real τs and determine the number of charged decay daughters (1 or 3). The reconstruction variables were used to find the number of charged tracks reconstructed in the detector. Figure 2 shows normalized histograms of the BDTj score for true τ_1 and τ_3 candidates. The BDTj clearly depends on the reconstructed N_{tracks} with those being correctly reconstructed giving rise scores closer to 1 in both cases. Those τ candidates that were misreconstructed have scores near the middle of the spectrum. Figure 3 plots candidates reconstructed with the same number of tracks (1, 2 or 3), separated by truth information. We see that those candidates reconstructed with two tracks look very similar regardless of the truth information.

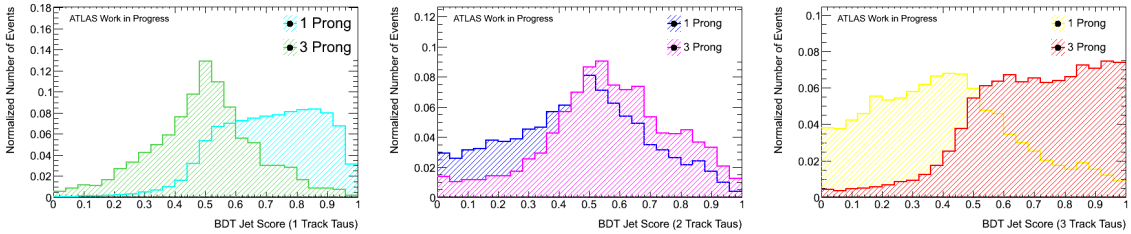
# Reconstructed Tracks	1 Prong (Truth)	% τ_1	3 Prong (Truth)	% τ_3	Total
1	107064	86.92%	2278	7.15%	109342
% 1 track	97.92%	-	2.08%	-	-
2	6978	5.67%	3563	11.18%	10541
% 2 track	66.20%	-	33.80%	-	-
3	9132	7.41%	26017	81.67%	35149
% 3 track	25.98%	-	74.02%	-	-
Total	123174	-	31858	-	155032

Table 4: Number of taus which are truly 1- and 3-prong, and the number of taus that are reconstructed with 1, 2 and 3 tracks.



(a) N_{Prong}=1 (b) N_{Prong}=3

Figure 2: The BDTj score of of real taus reconstructed with 1, 2 and 3 tracks plotted separately for τs which are τ_1 and τ_3 in truth.



(a) N_{track}=1 (b) N_{track}=2 (c) N_{track}=3

Figure 3: The BDTj score for real taus plotted separately for 1, 2, and 3 reconstructed tracks.

Table 4 shows the number of real τ_1s and τ_3s that are reconstructed with each track number. From this, we conclude that although τ candidates with incorrectly reconstructed tracks do not perform very well in the BDTj, there are relatively few (7.15% τ_3 reconstructed as τ_1 and 7.41% τ_1 reconstructed as τ_3) which are misreconstructed in this way. We also see that τs reconstructed with two tracks are more often τ_1 than τ_3. In our analysis, we use the τ_3 BDTj for these candidates but this effect is not too important because so few real τs are reconstructed with two tracks.

3.2 Tau Polarization Correction in Monte Carlo

When several of the $t\bar{t}$ MC samples we use to estimate our signal were generated, the polarization of the τ leptons was not properly incorporated into the simulation. There was however, a fix for this problem developed by the ATLAS Tau Working group, which consisted of reweighting each event in the MC samples based on the corrected polarization. This fix looked to be computationally expensive if implemented in our standard analysis so we first investigated its effects. We compared the distributions of several key variables, including the BDTj score (Figure 4), the transverse energy (E_T, Figure 5), η (Figure 6) and ϕ (Figure 7), for a corrected and uncorrected sample. The plots show very little difference between the two samples. Table 5 shows the changes in acceptances between the corrected and uncorrected MC samples. Since this change is very small, a systematic

uncertainty due to this polarization error was assigned instead of performing a full correction.

	Uncorrected MC			Corrected MC		
# of Tracks	Before cuts	After Cuts	Acceptance	Before Cuts	After Cuts	Acceptance
1	1812463	933424	51.50 %	192415	94752	49.24%
>1	1017516	867369	85.24%	112228	95457	85.06%

Table 5: Acceptances of true τ candidates with 1 and multiple tracks for corrected and uncorrected MC samples. The cuts applied are the loose τ candidate cuts, along with a BDTe cut at 0.51.

3.3 Pileup Studies

Both in-time and out-of-time pileup become of greater concern as the LHC moves to higher luminosity. In-time pileup arises when more than one proton-proton pair interacts in a given bunch crossing. This leads to additional vertices in an event from these additional pp interactions, so we can charterize the in-time pileup using the number of vertices in the event, N_{Vertex}.

Our BDTj was trained in bins of N_{Vertex}, to partially account for the anticipated increase in luminosity. We have plotted the BDTj score versus the number of vertices (Figures 8 and 9). These are plotted for different bins of E_T, since both the signal and the background tend toward higher BDTj score values with increasing energy. We see that for both real τs and jets faking τs, the BDTj score increases slightly with increasing number of vertices. This indicates that both the signal and background become more signal-like.

Out-of-time pileup occurs when information from one collision is still present in the detector when the next colllision event occurs. This could be electronic signals in the detector elements that have not yet been read out or residue energy present in the calorimeters and is due to the very frequent bunch crossings. The average number of interactions per crossing, $\langle\mu\rangle$ incorporates the effects of both in- and out-of-time pileup.

Plots of the BDTj score versus the parameter $\langle\mu\rangle$ in bins of E_T for both signal and background (Figures 10 and 11) show that the BDTj score is relatively constant with increasing $\langle\mu\rangle$, although at high E_T, it increases slightly.

To characterize the change in BDTj of the signal with respect to the background, we plot the background rejection versus the signal efficiency. These plots are constructed by making increasingly tight cuts on the BDTj score and counting the number of signal and background events which pass these cuts. The signal and background efficiencies are defined by:

$$\epsilon_{sig} = \frac{\text{\# of real loose } \tau \text{ candidates which pass cut}}{\text{Total \# of real loose } \tau \text{ candidates}} \tag{1}$$

$$\epsilon_{bkgnd} = \frac{\text{\# of loose } \tau \text{ candidates matched to jets which pass cut}}{\text{Total \# of loose } \tau \text{ candidates matched to jets}} \tag{2}$$

In this case the kinematic region is defined by the number of prongs or tracks of the τ candidate as well as the E_T and the value of $\langle\mu\rangle$. We further define the background rejection to be:

$$r_{bkgnd} = \frac{1}{\epsilon_{bkgnd}} - 1. \tag{3}$$

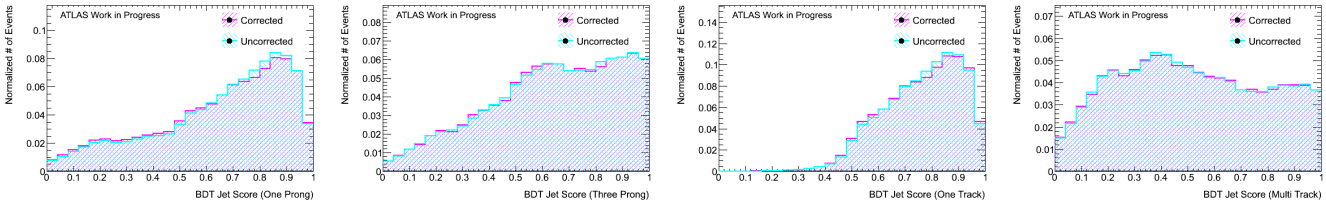


Figure 4: The BDT jet score distributions for both corrected and uncorrected Monte Carlo samples. The true taus are separated by both number of prongs (in truth) and the number of reconstructed tracks. From left to right: 1-Prong, 3-Prong, 1-Track and multi-Track taus.

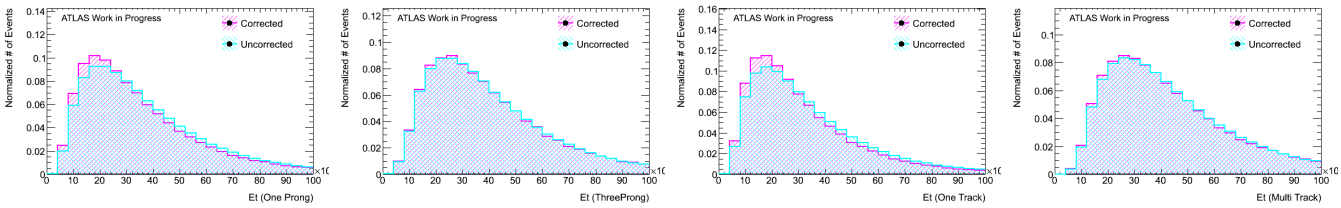


Figure 5: The E_T distributions for both corrected and uncorrected Monte Carlo samples. The true taus are separated by both number of prongs (in truth) and the number of reconstructed tracks. From left to right: 1-Prong, 3-Prong, 1-Track and multi-Track taus.

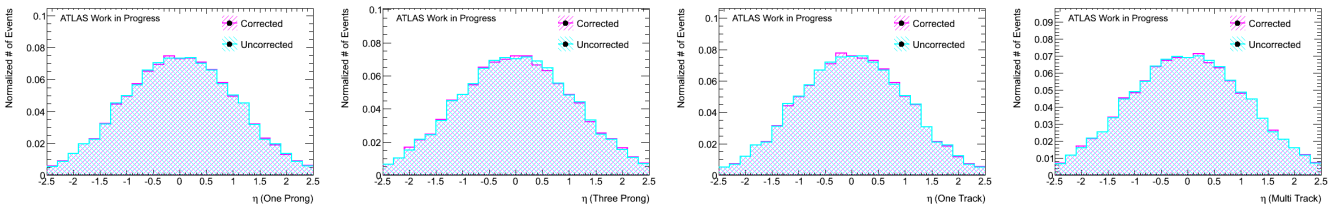


Figure 6: The η distributions for both corrected and uncorrected Monte Carlo samples. The true taus are separated by both number of prongs (in truth) and the number of reconstructed tracks. From left to right: 1-Prong, 3-Prong, 1-Track and multi-Track taus.

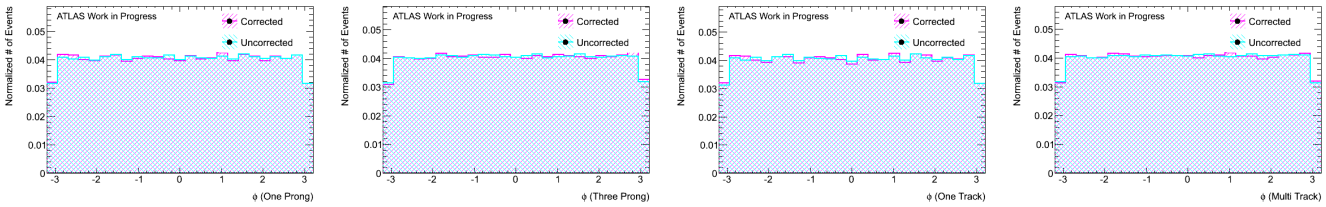


Figure 7: The ϕ distributions for both corrected and uncorrected Monte Carlo samples. The true taus are separated by both number of prongs (in truth) and the number of reconstructed tracks. From left to right: 1-Prong, 3-Prong, 1-Track and multi-Track taus.

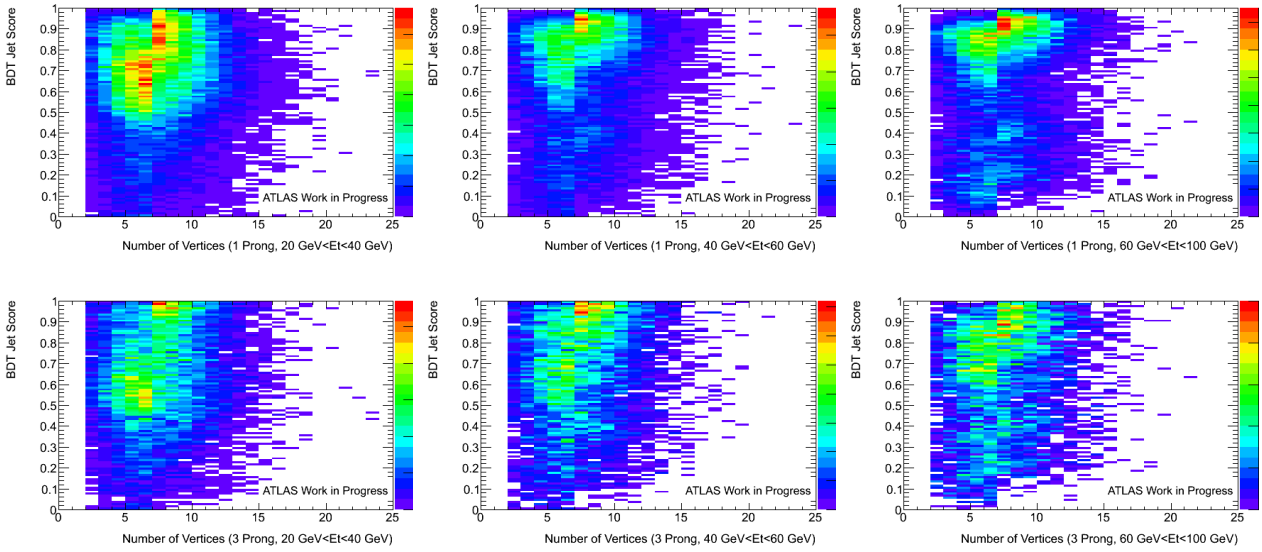


Figure 8: Plots of the BDT Jet Score versus the Number of Vertices for truth-matched taus. The top row are real τ_1s and the bottom row are real τ_3s, in bins of increasing energy: 20 GeV$< E_T \leq$ 40 GeV, 40 GeV$< E_T \leq$60 GeV, and 60 GeV$< E_T <$100 GeV

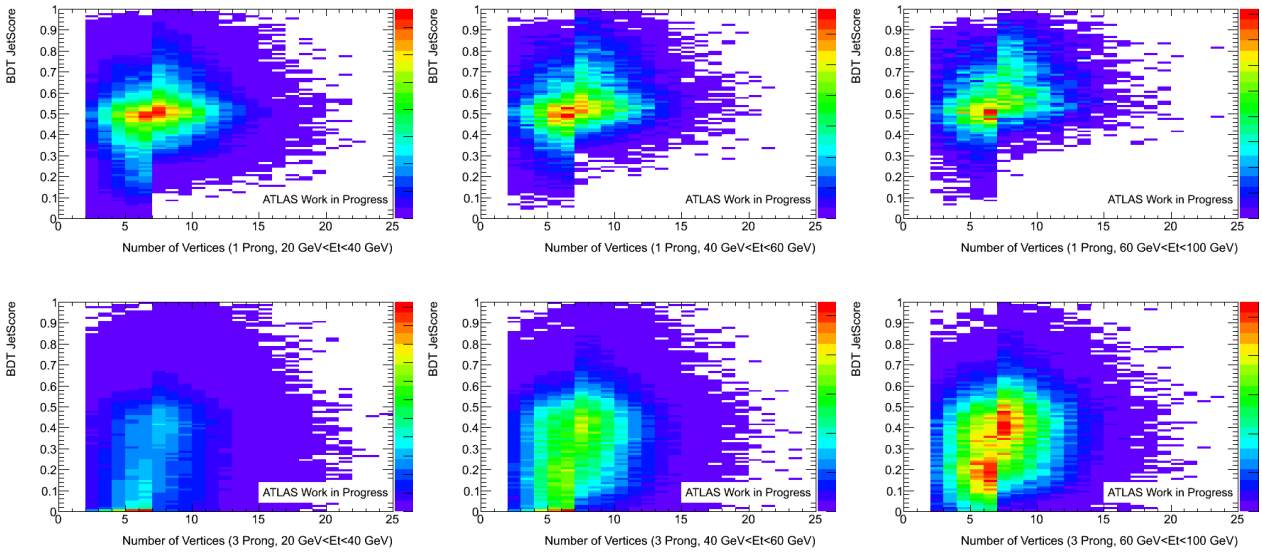


Figure 9: Plots of the BDT Jet Score versus the Number of Vertices for jets faking taus. The top row are τ_1 candidates and the bottom row are τ_3 candidates, in bins of increasing energy: 20 GeV$< E_T \leq$ 40 GeV, 40 GeV$< E_T \leq$60 GeV, and 60 GeV$< E_T <$100 GeV

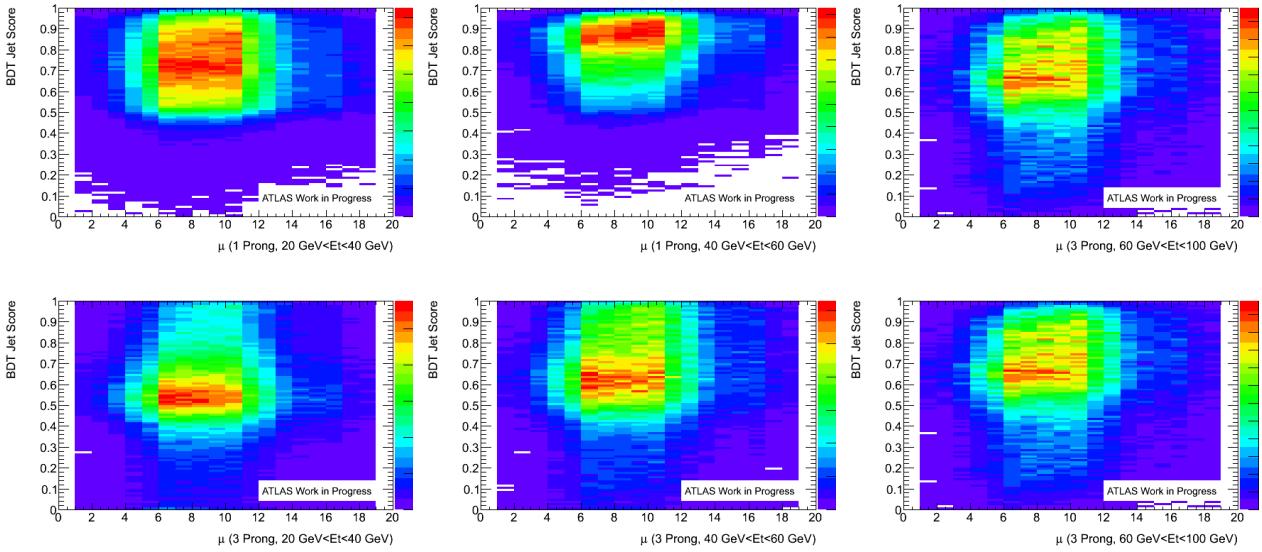


Figure 10: Plots of the BDT Jet Score versus the parameter $\langle\mu\rangle$ for truth-matched taus. The top row are real τ_1s and the bottom row are real τ_3s, in bins of increasing energy: 20 GeV$< E_T \leq$ 40 GeV, 40 GeV$< E_T \leq$60 GeV, and 60 GeV$< E_T <$100 GeV

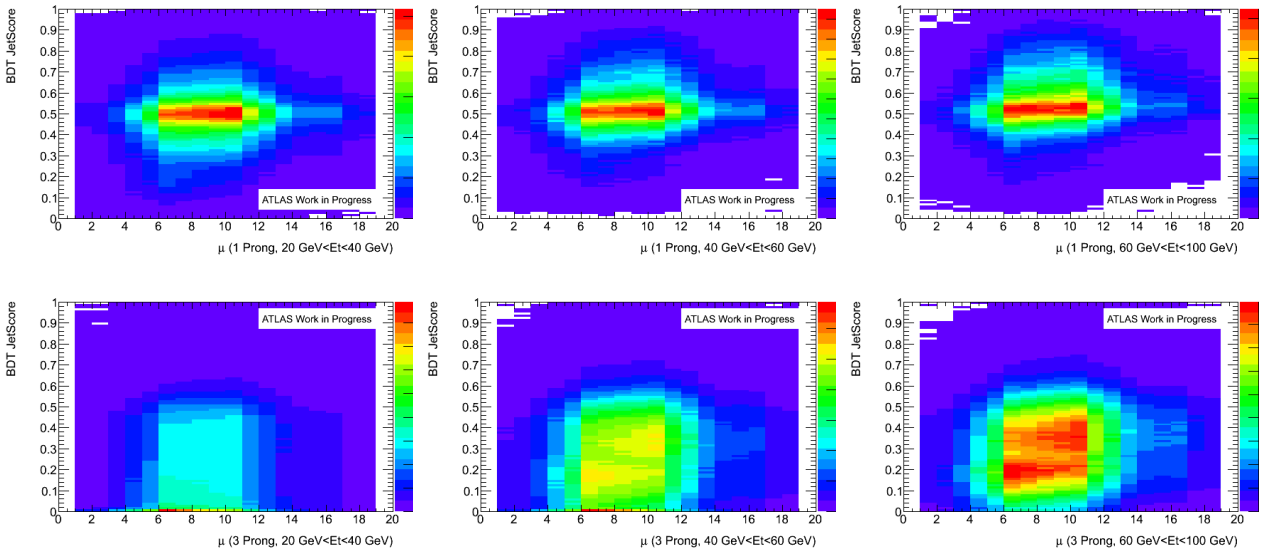


Figure 11: Plots of the BDT Jet Score versus the parameter $\langle\mu\rangle$ for jets faking taus. The top row are τ_1 candidates and the bottom row are τ_3 candidates, in bins of increasing energy: 20 GeV$< E_T \leq$ 40 GeV, 40 GeV$< E_T \leq$60 GeV, and 60 GeV$< E_T <$100 GeV

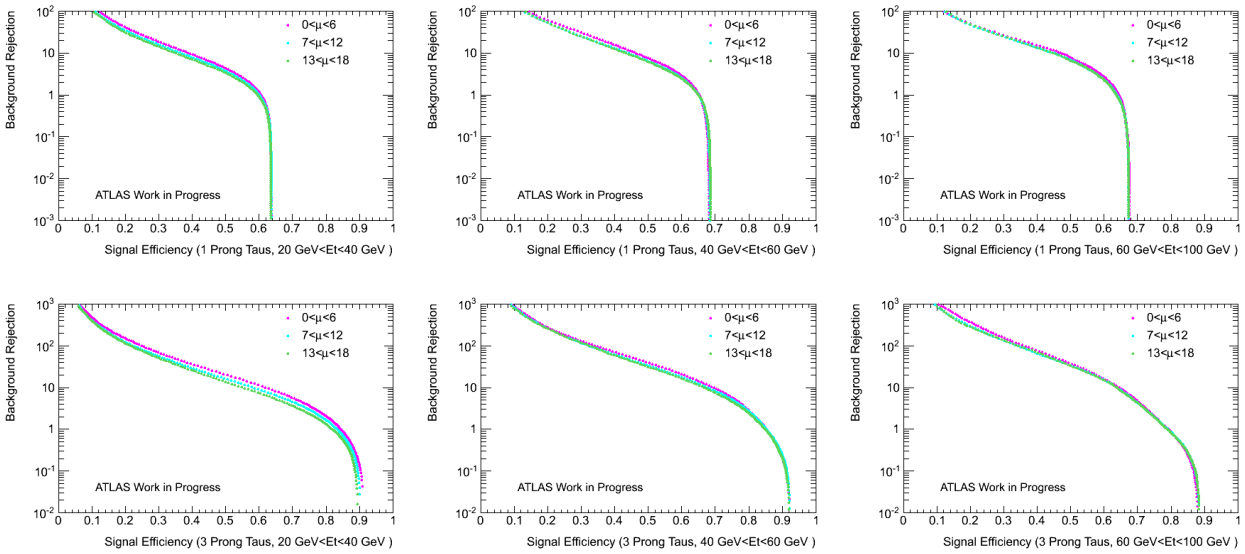


Figure 12: Background rejection versus signal efficiency curves for τ_1 (top row) and τ_3 candidates (bottom row) for the energy ranges 20 GeV$< E_T \leq$40 GeV, 40 GeV$< E_T \leq$60 GeV and 60 GeV$< E_T <$100 GeV.

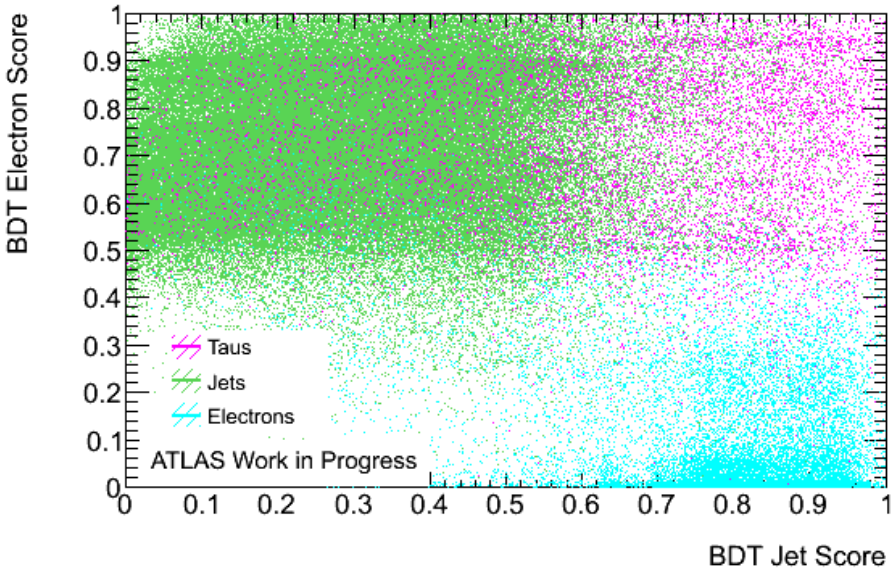


Figure 13: The BDT electron score plotted versus the BDT jet score. Real τs are shown in pink while jets faking τs are in green and electrons faking τs are in blue.

Figure 12 shows plots of the background rejection versus signal efficiency for τ_1 and τ_3 candidates overlaid for candidates falling in different ranges of $\langle\mu\rangle$. Good performance is characterized by both a high signal efficiency and a high background rejection. We see that the BDTj performance is slightly better for τ candidates in the lower $\langle\mu\rangle$ range.

3.4 Performance of the BDTe

Finally, we investigated the performance of the second decision tree used in our analysis, the BDTe by using $t\bar{t}$ MC samples to truth-match reconstructed τ candidates to real MC τs, electrons and jets. Figure 13 shows the BDTe score plotted against the BDTj score. The BDTe performs very well, with τs located at both high BDTe and high BDTj, electrons located at low BDTe and jets located at low BDTj.

4 Conclusions

Two boosted decision trees are used to discriminate real τs from electrons and jets faking τs. We extract the cross section for the $t\bar{t}\rightarrow \tau + \mu$ process by fitting the shape of the BDT jet output using templates for signal and background processes. Several studies in Monte Carlo simulations are conducted in order to better understand how the shape of the signal and background BDT output distributions change under specific conditions. These studies resulted in a systematic error being added to the final uncertainty to account for the incorrect polarization in the $t\bar{t}$ MC samples. The final cross section we measure for the $\tau + \mu$ channel is 142$\pm$21 (statistical) $\pm^{20}_{16}$ (systematic) $\pm$5 (luminosity) pb in 1.08 fb^{-1} of data [6], which is in good agreement with the standard model prediction.

References

[1] ATLAS Collaboration, G. Aad et al, *The ATLAS Experiment at the CERN Large Hadron Collider* (2008).

[2] ATLAS Detector Fact Sheet. http://www.atlas.ch/fact-sheets.html

[3] Particle Data Group, *Particle Physics Booklet* (2006).

[4] G. L. Kane, C. Kolda, L. Roszkowskie, J. D. Wells, *Study of constrained minimal supersymmetry*, Phys. Rev. **D49**, 6173 (1994).

[5] DØ Collaboration, V. M. Abazov et al., *Combination of $t\bar{t}$ cross section measurements and constraints on the mass of the top quark and its decays into charged Higgs bosons*, PRD **80**, 071102 (2009).

[6] ATLAS Collaboration, *Measurement of the top quark pair production cross section in pp collisions at $\sqrt{s}$=7 TeV in μ+τ final states with ATLAS*, ATLAS-CONF-2010-119.

[7] J. Godfrey, *Using boosted decision trees for tau identification in the ATLAS experiment*, Simon Fraser University, Burnaby, BC (2009).