

HEPTopTagger Parameter Optimization

David E. Sosa

D. Sosa from A. Schöning

A thesis submitted for the degree of Master of Sciences

Physikalisches Institut

Universität Heidelberg

October 2012

Abstract

At the LHC many top quarks are produced with a large transverse momentum. These top quarks have great potential for new physics. The HEPTopTagger is an algorithm designed to reconstruct the four-momenta of hadronically decaying top quarks with $p_T > 200$ GeV, on an event-by-event basis. It clusters jets with the Cambridge/Aachen algorithm with a large distance parameter $R=1.5$. The performance of the HEPTopTagger is optimized by tuning internal parameters of the algorithm for the largest signal-to-background ratio and highest signal efficiency in a sample containing pairs top quarks decays in which one W boson decays into a muon and a neutrino. Optimal parameter configurations that improve relevant figures of merit are found. Using these parameter configurations, the HEPTopTagger is then applied to Monte Carlo simulated signal and background events and compared to 2011 LHC data, showing good agreement. A multivariate analysis is performed, further improving figures of merit.

Abstrakt

Am Large Hadron Collider werden viele Top-Quarks mit hohem Transversalimpuls (p_T) produziert. Das Top-Quark bietet grosses Potential für Suchen nach Neuer Physik. Der HEPTopTagger Algorithmus wurde entwickelt um den Viererimpuls von hadronisch zerfallenden Top-Quarks mit $p_T > 200$ auf Ereignis-per-Ereignis Basis zu rekonstruieren. Der Algorithmus arbeitet auf C/A Jets, rekonstruiert mit einem Abstandsparameter von $R = 1.5$. Die Leistungsfähigkeit des HEPTopTagger kann durch verändern der internen Parameter für ein besseres Signal-zu-Untergrund-Verhältnis oder höhere Signaleffizienz optimiert werden. Parameterkonfigurationen welche diese Kenngrößen optimieren werden vorgestellt. Der HEPTopTagger Algorithmus wird auf Monte Carlo simulierte Signal- und Untergrundereignisse angewandt und mit 2011 am Large Hadron Collider aufgenommenen Daten verglichen wobei sich gute Übereinstimmung zeigt. Eine multivariate Analyse wird durchgeführt um die Kenngrößen weiter zu verbessern.

Acknowledgments

I would like to thank Prof. Dr. André Schöning and his ATLAS Heidelberg Group for the opportunity to collaborate with them. I feel truly lucky for having ended up in such a wonderful group. Particularly, I would like to thank Gregor Kasieczka without whom this thesis would not have been possible. His continuing assistance and mentoring have been and still are invaluable to me. Lastly, I want to thank all of my family in Mexico, my dad Jorge, my mom Imelda, my brother Luis and my sister Andrea. You will never know how much your unwavering support and encouragement has allowed me to keep going.

Contents

1	Introduction	2
2	Large Hadron Collider and the ATLAS Detector	5
2.1	The Large Hadron Collider	5
2.2	The ATLAS Detector	9
3	Theoretical Context	21
3.1	The Standard Model of Particle Physics	21
3.2	The Electromagnetic Interaction	22
3.3	The Strong Interaction	24
3.4	The Weak Interaction	25
3.4.1	The Electroweak Theory	26
3.5	The CKM Matrix	27
4	Beyond the Standard Model	27
4.1	Supersymmetry	29
4.2	Z' and Kaluza-Klein Gluons	30
5	The Top Quark in the Standard Model	31
5.1	Introduction	31
5.2	Top Pair Quark Production	31
5.3	Top Quark Decay	32
6	The HEPTopTagger	34
6.1	Background	34
6.1.1	Sequential Recombination Algorithms	34
6.1.2	Boosted Physics and Jet Substructure	35
6.2	The HEPTopTagger Algorithm	36
6.2.1	Steps	38

7	ATLAS Event Simulation	43
8	Cut-based Optimization	44
8.1	Event Selection	44
8.2	HEPTopTagger Parameter Tuning	45
8.3	Monte Carlo Validation with LHC Data	53
8.4	Low mass peak investigation	56
8.5	Dynamic Fat Jet Distance Parameter	58
9	Multivariate Analysis	60
9.1	Introduction	60
9.2	The Toolkit for Multivariate Analysis Package	60
9.3	Multivariate Methods	61
9.3.1	Simulated Annealing Optimized Cuts	61
9.3.2	Artificial Neural Networks	62
9.3.3	Boosted Decision Trees	63
9.4	Multivariate Analysis	64
9.4.1	Multivariate Analysis Preparation	64
9.4.2	Multivariate Optimization	73
10	Conclusion	81

1 Introduction

The Standard Model of particle physics attempts to describe all the known fundamental particles and their interactions. Built over the last half century, it encapsulates our best understanding about the nature of fundamental particles and their interactions with each other. Many predictions of the SM, such as the existence of the W and Z boson and the top quark, have been confirmed in particle accelerator experiments. Despite the many successes of the Standard Model, many fundamental questions remain to be answered. What are dark matter and dark energy? What is the origin of the matter-antimatter asymmetry? Do extra dimensions exist? Why is gravity much weaker than the strong, electromagnetic and weak forces? By which mechanism do particles get their mass? The Standard Model does not offer answers to these questions. Therefore there must exist new unexplained phenomena beyond the scope of the Standard Model. The Large Hadron Collider at CERN is giving and will give answers to many of these questions. As the most notable example, the ATLAS and CMS collaborations have made great progress in answering the question about where the mass of particles comes from. On July 4th 2012, both collaborations presented conclusive evidence for existence of a neutral boson compatible with the Standard Model Higgs boson, an excitation of the Higgs field which gives mass to particles.

The Large Hadron Collider is a circular proton-proton accelerator of 27 km circumference which guides the proton beams to collide in four different locations to search for interesting physics. Four detectors have been placed at these locations to examine the debris out of these high energy collisions. ATLAS is one of the general purpose particle detector that will search for indications of new physics. A detailed study of the top quark will greatly aid to answer many of these questions and will shed light into possible new physics. Because it is the heaviest known elementary particle and it is the only fermion with mass of the order of the electroweak scale, the top quark is an ideal window to look for new physics phenomena. These new phenomena would be observed as a new heavy resonances that would appear in the $t\bar{t}$ mass spectrum or a cascade of decays coming

from previously undetected particles. Close to 8,000,000 top quark pairs per year are expected LHC runs at designed luminosity and center-of-mass energy [1]. With such a high number of top quarks produced, very detailed measurements on its mass, spin, forward-backward asymmetry etc. will be possible. In the Standard Model, a top quark decays 99.8% to a W boson and to a b quark. Therefore, a $t\bar{t}$ pair decays (for practical purposes) always into two W bosons and two b quarks. The $t\bar{t}$ has three decay modes, all determined by the W boson: both W bosons decay to leptons (leptonic channel), one W boson decay to leptons and the other to quarks (lepton + jets) and both W boson decaying into quarks (all-hadronic). The all-hadronic channel was previously thought inaccessible for physics searches. However, in recent years new methods and algorithms that can find hadronically decaying top quarks (t_{had}) have arisen. Most of these top taggers focus in regimes where the top quark has a transverse momentum $p_T > 500$ GeV (highly boosted) [2]. The HEPTopTagger was designed to identify and reconstruct t_{had} 's with p_T range of 200-500 GeV (moderately boosted) in busy large multi-jet backgrounds.

The HEPTopTagger has parameters that control how much underlying event and pile-up are rejected, while still maintaining a good signal efficiency and acceptable background rejection. This thesis first focuses on investigating how different values for these parameters affect the performance of the HEPTopTagger against figures of merit and in different pile-up conditions. Secondly, the thesis investigates if it is possible to improve even further on figures of merit by taking relevant discriminant variables of the HEPTopTagger and performing a multivariate classification.

The Large Hadron Collider and the ATLAS detector are briefly described in Section 2. In Sections 3,4 and 5, overviews of the theoretical context in which this thesis develops, a brief description of the Standard Model, the top quark and physics beyond the Standard Model are given respectively. Section 6 introduces sequential recombination algorithms and their role in the HEPTopTagger algorithm. Particles decaying with a high transverse momentum (boosted objects), are introduced as a promising probe for new physics. The section goes on to introduce the rationale behind the implementation of the HEPTopTagger and to explain the algorithm's steps. Section 7 briefly describes

the ATLAS event simulation framework used to simulate the detector response. Section 8 investigates how different parameter configurations of the HEPTopTagger can deliver the best fraction of signal over background. Section Section 9 investigates how a multivariate analysis of relevant discriminant variables in the HEPTopTagger, can improve the mentioned figures of merit. Improvements are found with a multivariate analysis using optimized cuts, decision trees and artificial neural networks.

2 Large Hadron Collider and the ATLAS Detector

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [3] is a two-ring-superconducting accelerator and proton-proton (pp) collider¹ designed for high energy particle physics research. It is located at the European Center for Nuclear Research (CERN), at the Franco-Swiss border near Geneva, Switzerland. The largest and most powerful ring-accelerator in the world, the LHC was designed to reach center-of-mass (\sqrt{s}) collision energies of up to 14 TeV and a luminosity of up to $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. It is situated in a 26.7 km circular tunnel originally constructed for the Large Electron-Positron Collider that reaches a depth underground of up to 175 meters (Figure 2.1).

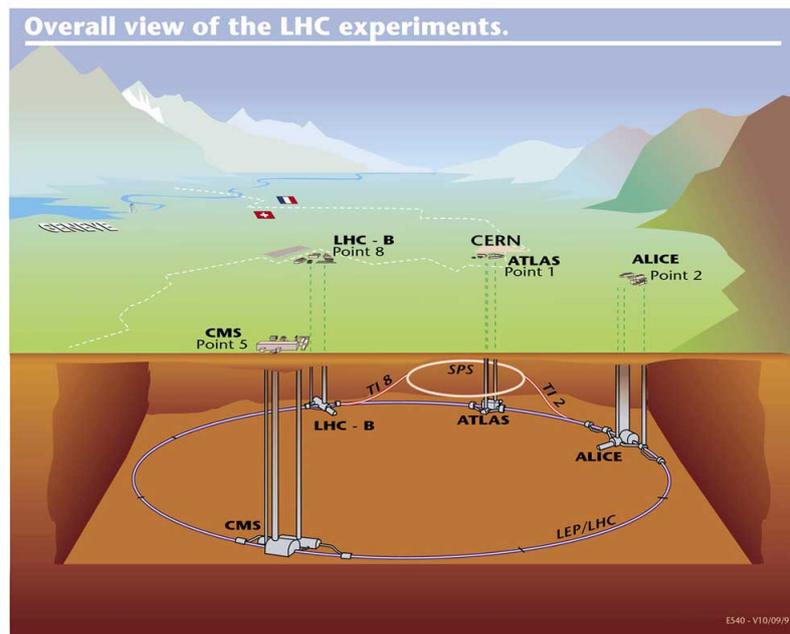


Figure 2.1: The LHC and the four main detectors: ALICE, ATLAS, CMS and LHCb [4].

The LHC consists of two adjacent parallel beamlines which intersect at four points. Through the beamlines, the proton beams travel in opposite directions and meet at the

¹Occasionally heavy ions, such as lead, are also collided for heavy ion research.

interaction points. The four main LHC experiments, ALICE [5], ATLAS [6], CMS [7], and LHCb [8] are located at these interaction points. Protons need to undergo a pre-acceleration process before entering the LHC. First, hydrogen gas is feed into a duoplasmatron to ionize it. Next, the bare protons go through a linear accelerator (LINAC2), the PS Booster, and the Proton Synchrotron and the Super Proton Synchrotron where they are accelerated to 50 MeV, 1 GeV, 26 GeV and 450 GeV, respectively (Figure 2.2). The LHC can maintain a stable beam at a minimum energy of 450 GeV. Protons are accelerated in the ring for about 20 minutes before reaching their maximum energy of 7 TeV. To maintain the proton beam on a circular path, 1232 dipole magnets are used. They are cooled down to 1.8 °K by a cryogenic system using liquid helium and provide a magnetic field of up to 8.33 T. 502 quadrupole magnets along the straight sections of the beam keep the beams focused to maximize the rate of pp interactions. With these magnets the protons are accelerated close to the speed of light and guided to collision. The beams have a lifetime of approximately 10 hours. During this time, collisions take place and data are taken inside the four LHC experiments. As the intensity of the beam decreases below a certain level, it is "dumped" and directed to collide with a metal block. The field strength of the magnets is then decreased to 0.54 T for 20-40 minutes. The beam injection is then repeated and the magnets field strength again increased to 8.3 T for another cycle. This beam cycle is called a "fill".

The first beams on the LHC were circulated successfully on September 10th 2008. However on of September 19th, an electrical fault caused a magnetic quench. This in turn, caused a helium gas explosion that damaged over 50 superconducting magnets and contaminated the vacuum pipe halting the operation almost for a year. On November 2nd 2009, proton beams were again successfully circulated and the first collision recording took place at $\sqrt{s} = 450$ GeV. Later on March 2011, the LHC set a world record for the highest energy on man-made particle collision, at $\sqrt{s} = 7$ TeV. It continued running at this energy until the end of 2011. On March 2012, \sqrt{s} was increased to 8 TeV and it will continue running at this energy until December 2012.

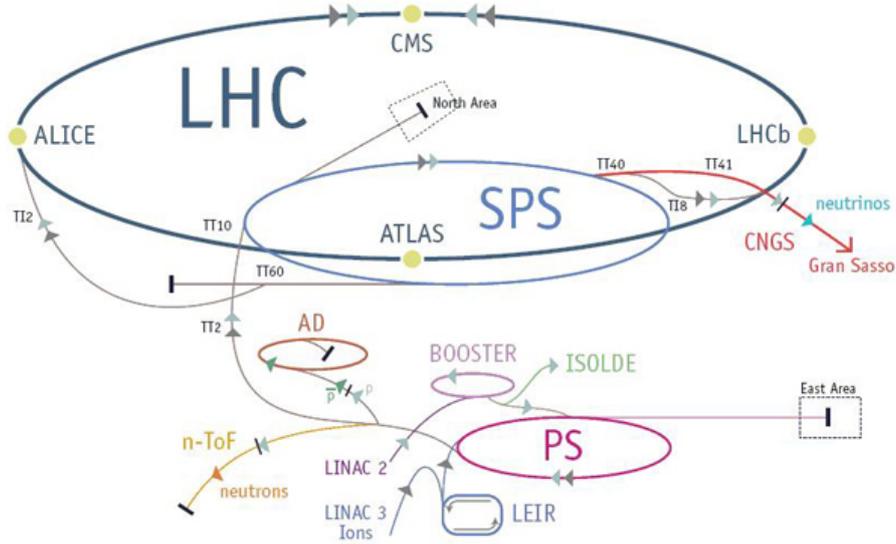


Figure 2.2: The accelerator complex at CERN [9].

Luminosity and Pile-Up at the LHC: At the LHC, the pp beams are not continuous, but they come in "bunches". Each proton beam consists of 2808 bunches each containing around 1.15×10^{11} protons. Currently, collisions take place at discrete intervals every 50 ns. The instantaneous luminosity, L is an important parameter that measures the performance of a particle collider. It indicates the rate pp collisions. The machine luminosity is determined entirely by beam parameters as given by,

$$L = f \frac{n_1 n_2}{4\pi\sigma_1\sigma_2} \quad (2.1)$$

where f is the crossing rate of the bunches (currently, 20 MHz), n_1, n_2 is the number of protons in each bunch and σ_1, σ_2 are the standard deviations of the Gaussian shaped beam. The goal of constructing more powerful accelerators is to increase the luminosity and have a chance to observe more rare and exotic processes (i.e. low cross section) as seen in the formula,

$$\frac{dN_{events}}{dt} = \sigma_{events} \cdot L. \quad (2.2)$$

where σ_{events} is the the cross section of a particular process and N_{events} is the number of events at a certain \sqrt{s} . The luminosity at the LHC is increased by reducing their transverse size of the beam ("squeezing" the protons) or by increasing the number of protons per bunch. The instantaneous luminosity in ATLAS is measured by specialized detectors located in the forward region of the detector [10, 11]. Usually, the number quoted to measure the progress in data taking of an accelerator is the integrated luminosity,

$$\int \frac{dN_{events}}{dt} dt = \sigma_{events} \cdot \int L dt, \quad (2.3)$$

It is estimated that ATLAS will collect 100 fb^{-1} of data per year when running at design luminosity. However this figure will not be reached until 2014. In 2010 and 2011, 48.1 pb^{-1} and 5.61 fb^{-1} of data were collected respectively. From April 4th to mid-October 2012, a total of 20.63 fb^{-1} have been collected.

As the instantaneous luminosity increases, the additional energy in an event coming from interactions other than the primary vertex and its underlying event (known as pile-up) increases as well. Increasing pile-up worsens measurements and new methods must be devised to compensate for this additional, unrelated energy. There are two types of pile-up: extra pp interactions within the same bunch crossing (in-time pile-up) and pp interactions coming from a different bunch crossings (out-of-time pile-up). The number of reconstructed vertices (N_{PV}) can be used to estimate the in-time pile up. Another measure of pile-up sensitive to both, in-time and out-of-time pile-up, is the mean number of pp collisions per bunch crossing at the time of the recorded event [12]

$$\langle \mu \rangle = \frac{L_{bunch} \times \sigma_{inel}}{f_{LHC}} \quad (2.4)$$

where L_{bunch} is instantaneous luminosity per bunch, σ_{inel} is the pp inelastic cross section and f_{LHC} is the revolution frequency of the protons at the LHC. At the beginning of the LHC operation in 2010, with a maximum instantaneous luminosity of $2 \times 10^{32} \text{ cm}^2 \text{ s}^{-2}$, pile-up was negligible. In 2011, luminosity increased to $3 \times 10^{33} \text{ cm}^2 \text{ s}^{-2}$ and $\langle \mu \rangle$

increased to 6. For 2012, the instantaneous luminosity is $7 \times 10^{33} \text{ cm}^2 \text{ s}^{-2}$ and $\langle \mu \rangle = 20$ (see Figure 2.3). For this thesis, the effects of pile-up were investigated for $\langle \mu \rangle = 20$.

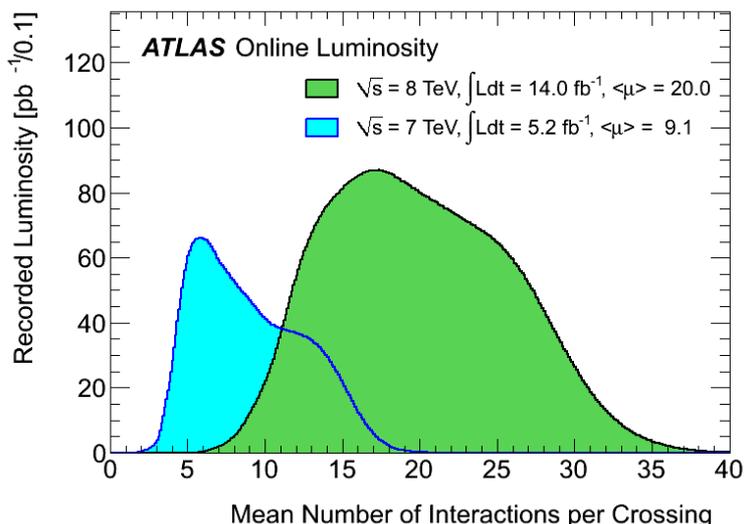


Figure 2.3: Recorded luminosity as function of the mean number of interaction per bunch crossing $\langle \mu \rangle$ for the full 2011 data and 2012 data collected between April 4th and September 17th 2012 [12].

2.2 The ATLAS Detector

The ATLAS (A Toroidal LHC Apparatus) detector is one of two multipurpose particle physics experiments located at a collision point of the LHC [6]. The layout of the detector with its major subsystems is shown in Figure 2.4. The ATLAS detector is 25 m high and 44 m long and weighs approximately 7000 tons. It has an almost 4π coverage with nominal forward-backward symmetry with respect to the interaction point. The ATLAS collaboration is formed by more than 2900 physicists and engineers from 172 institutions around the world. Four major components or sub-detectors compose the ATLAS detector: The magnet systems, the inner detector, the calorimeters and the muon spectrometer. A sophisticated trigger system is implemented to select only events that contain signature with potential interesting physics.

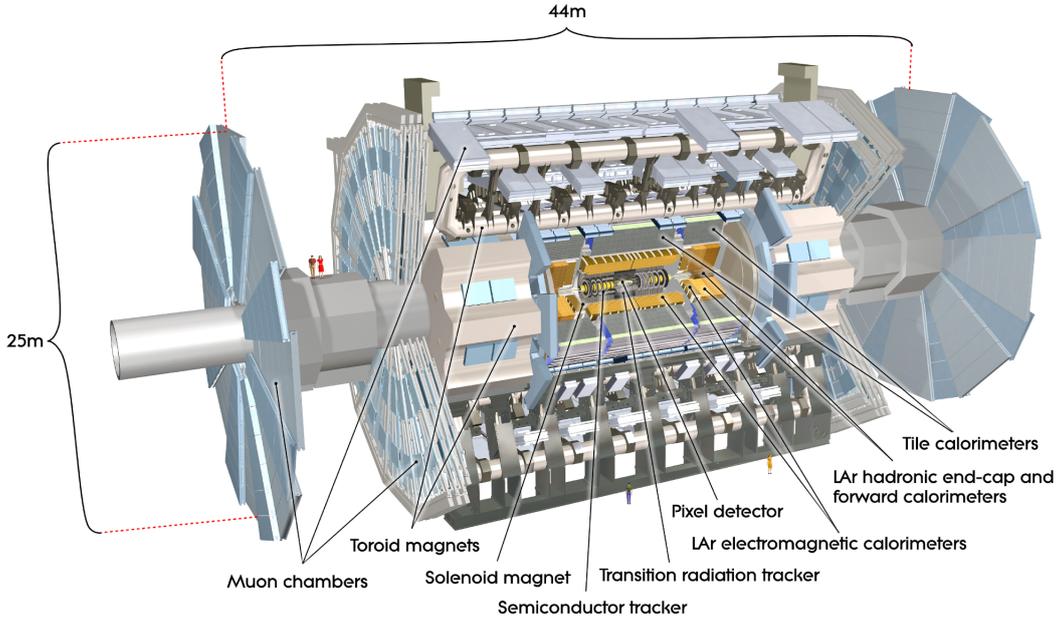


Figure 2.4: A layout of the ATLAS Detector and its main components [6].

ATLAS Coordinate System: ATLAS uses a Cartesian right-handed coordinate system, with the pp interaction point as the origin. The x -axis points to the center of the LHC ring, the z -axis points to the beam direction and the y -axis points upwards (see Figure 2.5). The azimuthal angle ϕ is measured from the positive z -axis. The polar angle θ is measured from the positive y -axis. The rapidity is defined as $y = 0.5 \times \ln[(E + p_z)/(E - p_z)]$ and the pseudorapidity as $\eta = -\ln(\tan(\frac{\theta}{2}))$. For massless objects the rapidity and the pseudorapidity are equivalent. These variables are extensively used in collider physics because of their invariance to Lorentz boosts along the z axis. In a pp collision, the boost along the z -axis is not known because the partons² that collide and give rise to new particles, carry an unknown fraction of the proton's momentum, as determined by the proton distribution function. Therefore, variables that are invariant to boosts along the z -axis are preferred in collision experiments. Likewise, the transverse momentum, p_T , transverse energy, E_T and transverse missing energy, E_T^{miss} (coming from undetected particles) are all defined with respect to the x - y plane. A

²A parton can be a quark or a gluon.

distance between two points in the $\eta - \phi$ space is defined as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

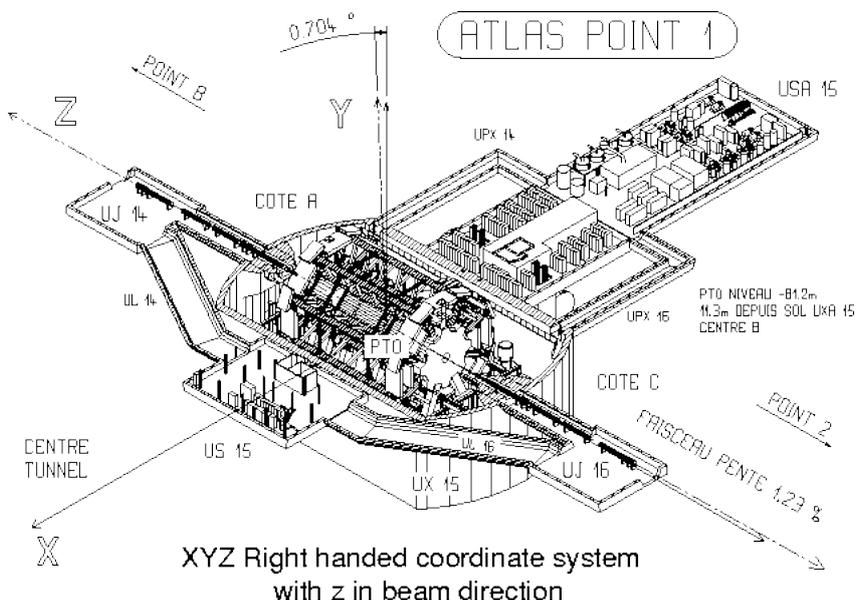


Figure 2.5: The ATLAS detector coordinates [13].

Magnet System Two superconducting magnets are used to bend charged particles (see Figure 2.6). The inner solenoid provides a 2 Tesla magnetic field surrounding the Inner Detector. It has a longitude of 5.8 m and a diameter of 2.5 m. It operates with a nominal current of 7730 A. It is cooled down using liquid helium down to a temperature of 4.5 K. The outer toroidal magnetic field is produced by very large air-core superconducting barrel loops and two end-caps air toroidal magnets, all situated around the muon system. It provides a magnetic field of approximately 0.5 and 1.0 for the muon detectors in the central and end-cap regions respectively. The outer toroidal magnet measures 22 meters in diameter and 26 m long in length. With a stored energy of 1.6 GJ it provides the magnetic field over a volume of approximately 12,000 m³.

Inner Detector: The inner detector (ID) starts a few centimeters from the beamline and continues radially up to 1.2 meters. It spans seven meters in length along the beamline. Its basic function is to track charged particles created from the pp collision. Tracking reveals the paths of charged particles. The curvature of the particle's track

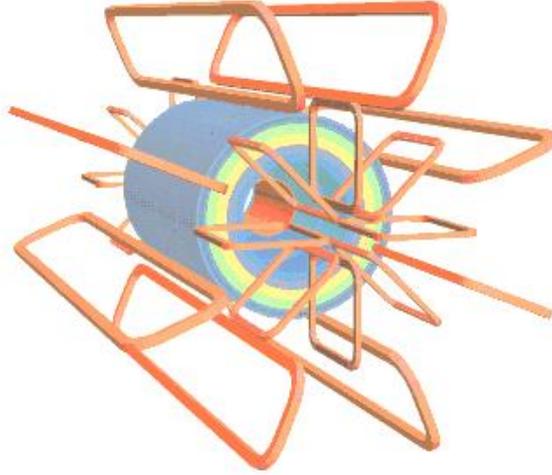


Figure 2.6: A computer generated image of the ATLAS Magnet System. The eight barrel toroid coils with the end-coil and inner solenoidal magnet are visible [6].

(caused by the presence of a magnetic field) is recorded and this allows for a measurement of its momentum. The orientation of the track gives the particle's charge. The inner detector is also capable to identify secondary vertices coming from long-lived particles . The designed momentum resolution is [6]

$$\sigma_{p_T}/p_T = 0.05\% p_T \text{ GeV} \oplus 1\%. \quad (2.5)$$

The ID is composed of three parts: the Pixel Detector, the Semi-Conductor Tracker, and the Transition Radiation Tracker. The Pixel Detector (PD) is the inner most part of the inner detector (see Figure 2.7). It has a coverage up to $\eta = 2.5$ and complete ϕ coverage. It provides three high resolution 3D space points with a spatial resolution of $10 \mu\text{m}$ in the $r - \phi$ direction and $110 \mu\text{m}$ in the z direction. The 1744 modules of the detector are built from $250 \mu\text{m}$ thick silicon pixels connected to a read-out electronics. The pixels have dimensions of $50\mu\text{m} \times 400 \mu\text{m}$. With this size, it is possible to achieve a very high granularity close to the interaction point. This high granularity is needed to detect primary vertices and secondary vertices coming from long lifetime decaying

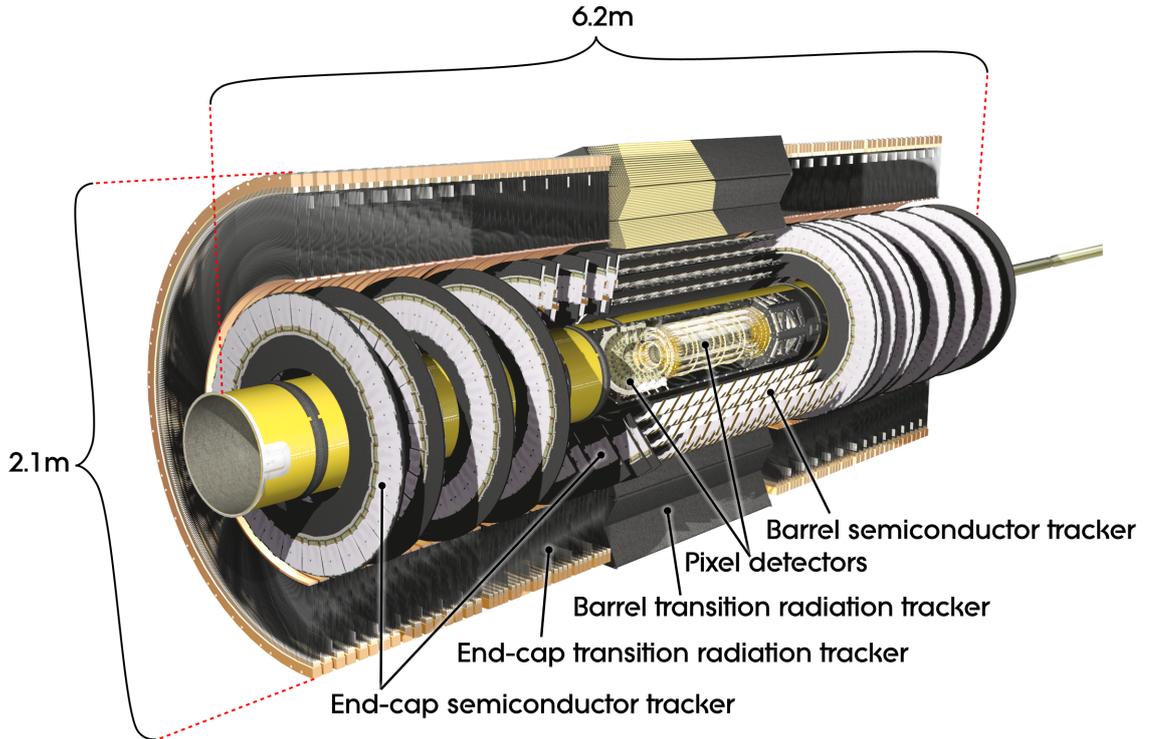


Figure 2.7: Detailed view of the Inner Detector [6].

particles such as b quarks and τ leptons.

The Semi-Conductor Tracker (SCT) is the middle component of the ID. It has the same coverage as the PD. The design concept of the SCT is similar to the PD. However, due to a lower particle density, it is possible to use strips instead of pixels. It provides eight precision measurements per track and contributes to the measurements of momentum, impact parameter and vertex position.

The Transition Radiation Tracker is the outer component of the ID. It has η coverage up to 2.0 and only provides $r - \phi$ information. It uses gaseous straw tubes interspaced with transition radiation material. Charged particles crossing the straw tubes create transition radiation³ that is then absorbed by the xenon gas within. In the center of the straw there is a $30 \mu\text{m}$ gold covered tungsten wire that in turn collects the electrons

³Transition radiation is created when a highly relativistic particle crosses a material with different index of refraction. The amount of energy radiated is proportional to $\gamma = E/m$. For a given energy, electrons radiate 250 times more than pions.

coming from the xenon gas. The capability of transition radiation from the TRT provides stand-alone electron identification. The TRT provides around 36 measurements allowing a precise reconstruction of the particles trajectory. It contributes the most to the momentum measurement in the inner detector.

Calorimeters: The purpose of the calorimeters is to measure the particle's energy by totally absorbing it. The calorimeter was designed to provide good energy resolution for electrons, photons and hadrons. There are two calorimeter systems: the inner electromagnetic calorimeter and the outer hadronic calorimeter. Both are sampling calorimeters, i.e. high-density material is used as an absorber while an active layer is placed in between to sample the shape from the particle shower. They are position sensitive, that is, they can measure energy depositions depending on their location. Large particle energies at the LHC, make the calorimeter an indispensable tool. As seen from Equation 2.5, the resolution of the ID decreases, as the energy increases (p_T). The resolution of a sampling calorimeter is given by

$$\frac{\sigma_E}{E} = \frac{S}{\sqrt{E}} \oplus \frac{N}{E} \oplus C, \quad (2.6)$$

where the S represents the sampling or stochastic term. The choice of the absorber, active material and thickness of sampling layers among other factors contribute to this term. It affects the calorimeter resolution mostly in the range of 10 - 100 GeV. N represents the noise term. It includes the electronic noise and the signal pile-up. Its contribution is significant at low energies. The constant term C takes into account the depth of the detector, detector non-uniformities and dead material among other factors. It dominates at high energies.

The electromagnetic calorimeters absorb mostly particles that interact electromagnetically. They have a resolution of $\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.7\%$. The barrel covers $|\eta| < 1.475$ and the two coaxial end-cap wheels cover the region $1.375 < |\eta| < 3.2$. The energy-absorbing materials are lead and stainless steel with liquid argon (LAr) as the active medium. Full ϕ coverage is ensured by the calorimeter accordion shape. Liquid argon

provides a good resistance to radiation and uniformity that translates into spatial uniformity in the energy measurement. Charged particles crossing the calorimeter ionize the LAr and the electrons drift towards the electrodes in the read out cell following the principle of a drift chamber. For argon to stay liquid it needs to be maintained at a temperature of 88 K. Therefore the EM calorimeter and the solenoid magnet share the same cryostat vessel in order to minimize inactive material. The EM calorimeter is designed to completely contain an electromagnetic shower. An important parameter to define the length of a given EM shower, is the radiation length, X_0 of a material. It is defined as the mean distance over which an electron loses all but $1/e$ of its energy. The particles going through the ATLAS EM calorimeter transverse $22 X_0$ to $33 X_0$.

The hadronic calorimeter Because the hadronic interaction length λ_{int} , is larger than the EM radiation length X_0 , hadronic calorimeters need to be deeper to completely absorb hadrons and measure their energy. Hadronic showers are more complex than their electromagnetic counterparts. As an example, for a 5 GeV proton in a lead-scintillator calorimeter, the energy that goes through the hadronic calorimeter is distributed as follows [14]:

- Ionization of charged particles (p, π , μ) (40%)
- Electromagnetic showers ($\pi^0 \rightarrow \gamma\gamma, \eta^0 \rightarrow \gamma\gamma, e$) (15%)
- Neutrons (10%)
- Protons from nuclear de-excitation (6%)
- Non-detectable energy (nuclear binding, neutrinos) (29%)

The resolution of the ATLAS hadronic calorimeter is $\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%$. The main reason for the lower resolution, is that the deposited energy that is absorbed in nuclear breakups and excitations (invisible energy) cannot be detected. The intrinsic response to EM showers and hadronic showers is not equal i.e. ($e/h \neq 1$). Therefore a compensation needs to be implemented to account for these different responses. The source of these different energy responses, is the invisible energy. This ratio is brought to one with

compensation. This can be done with hardware (modifying thickness of sampling or absorbing material, etc.) or with software (adjusting the response, etc.). Once $e/h = 1$, the calorimeter is compensated.

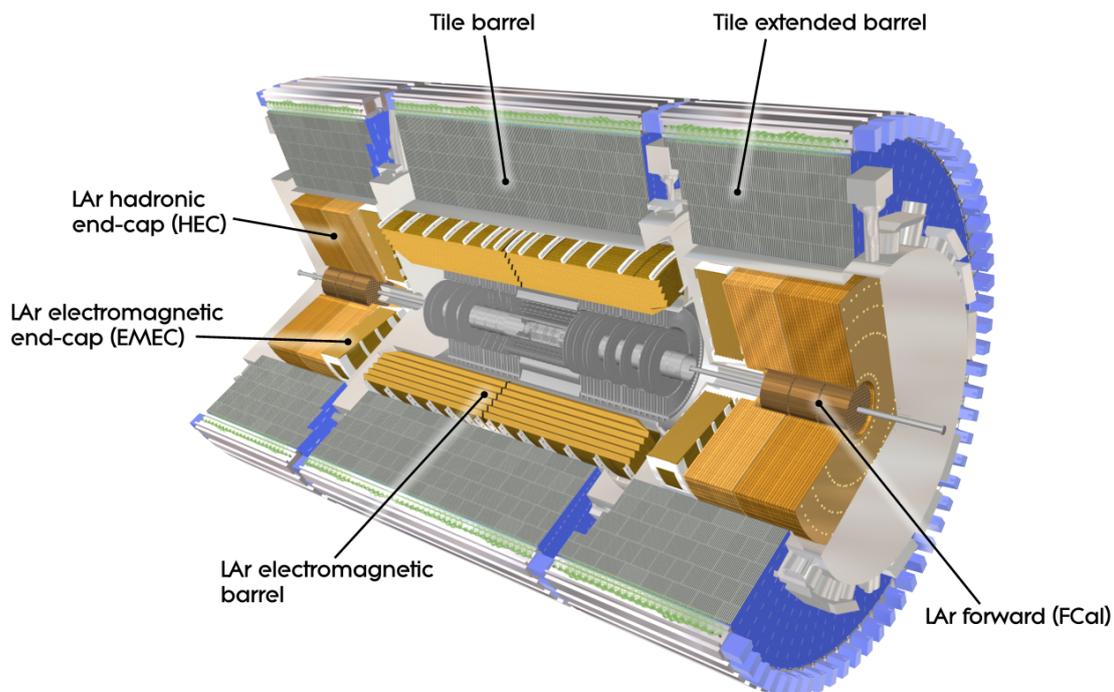


Figure 2.8: A computer generated cut-away of the ATLAS calorimeters [6].

Muon Spectrometer: The muon spectrometer (MS) starts at a radius of 4.25 meters around the calorimeters and ends at 11 m radial distance. Muons penetrate easily the inner detector and the calorimeters and reach mostly undisturbed the muon chambers. It occupies a volume of around 16000 m³. Its main function is to trigger on muons and identify and measure their momentum. Such dimensions are required for accurate momentum measurements. The resolution on transverse momentum p_T is $\sigma_{p_T}/p_T = 10\%$ at $p_T = 1$ TeV. The momentum reconstruction resolution is optimal around 100 GeV with 3%. This resolution can be improved to 2% if track measurements coming from the MS are combined with those from the inner detector. Triggering is achieved with coarse

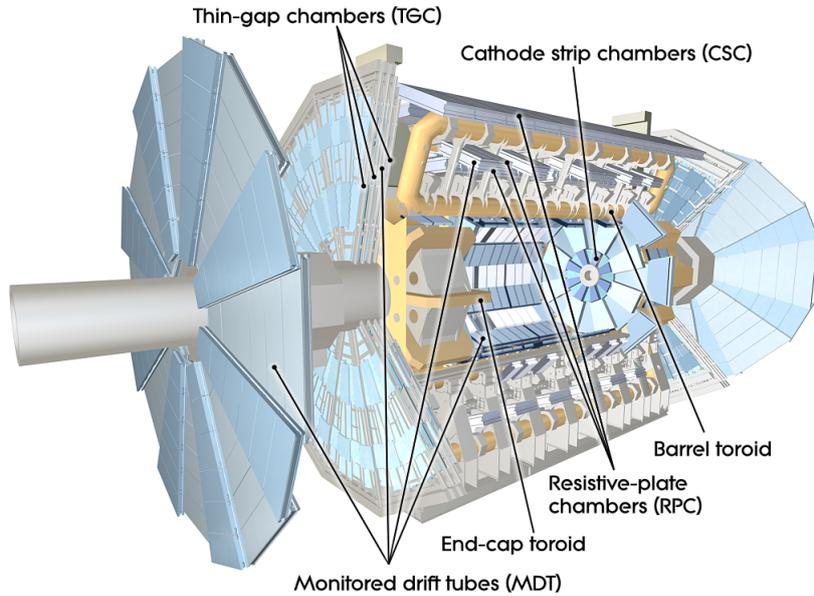


Figure 2.9: A computer generated image of the ATLAS Muon System [6].

chambers: the Resistive Plate Chambers and the Thin Gap Chambers. Momentum precision measurements are done with the Monitored Drift Tube and the Cathode Strip chambers. The chambers provide full coverage up to $|\eta| = 2.7$, except for $|\eta| \approx 0$ where all the services from the other detectors are routed out of ATLAS. Because the magnetic field bends the particles only on a plane of constant azimuth, ϕ , that passes through the beam axis, no precise ϕ information is needed to reconstruct the muon momentum. Therefore, the precision chambers are built to measure the coordinate of interest η . Because of the long latency of the precision chambers, they cannot be used as trigger. This is why they are complemented with coarse trigger chambers. The end-caps of the MS consists on 4 disks which provide a pseudorapidity range of $1.0 < |\eta| < 2.7$. They are placed 7, 13 and 21 m away from the interaction point along the beam direction. This guarantees almost full $|\eta|$ coverage. The different parts of the MS are shown in Figure 2.9.

Together, the tracking system, the calorimeters and the muon spectrometer are able to distinguish and reconstruct different types of particles as shown in Figure 2.10.

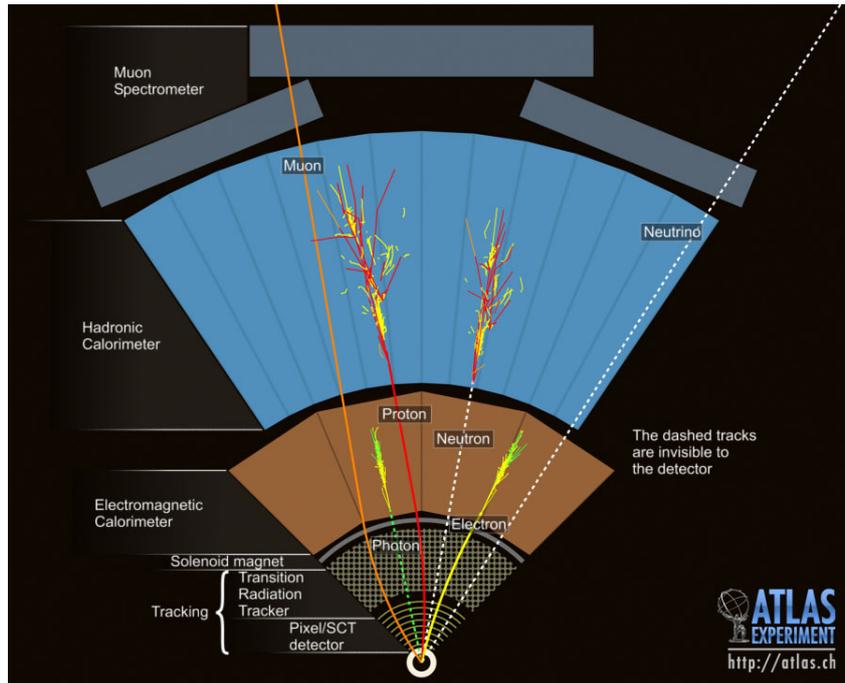


Figure 2.10: A slice of the ATLAS showing how the particles would be identified in each subdetector [15].

Forward Detectors: Forward detectors are located collinear to the beamline far from the interaction point. Their purpose is to provide additional important measurements such as beam monitoring and luminosity measurement. The main two forward are LUCID and ALFA. LUCID (Luminosity Measurement using Cerenkov Integrating Detector) and ALFA (Absolute Luminosity for ATLAS) determine the luminosity delivered to ATLAS. LUCID is designed to cope with luminosities of $10^{27} \text{ cm}^{-2} \text{ s}^{-1}$ up to $4 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$. The luminosity is measured bunch by bunch. ALFA was installed in 2011 and its goal is to measure using dedicated runs of low luminosity (from 10^{27} to $10^{28} \text{ cm}^{-2} \text{ s}^{-1}$), the total pp cross section and absolute luminosity thus providing a calibration point for LUCID. With ALFA in place, absolute luminosity measurements will have an accuracy of about 3% in 2012.

Trigger and Data Acquisition At the LHC, a rate of pp interactions of 10^9 Events/sec is expected. To cope with this unprecedented rate, a very sophisticated trigger system

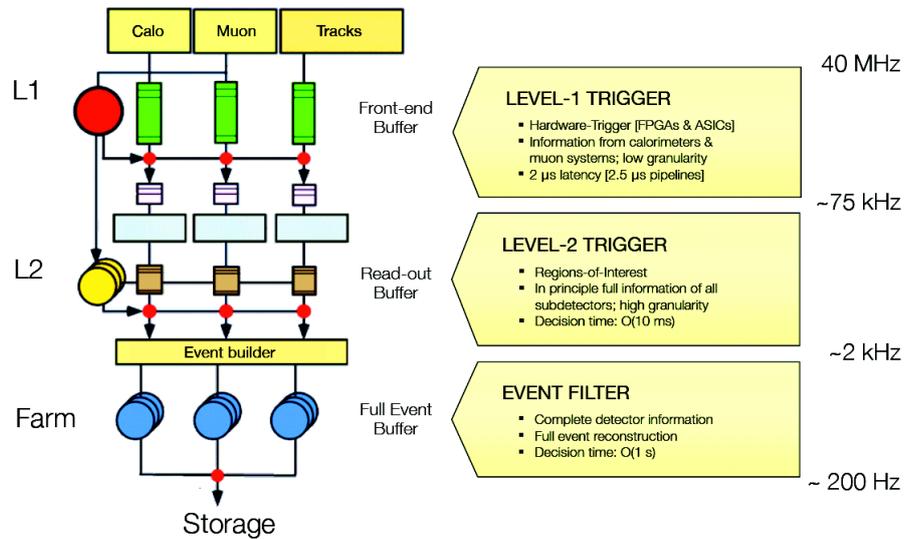


Figure 2.11: The ATLAS trigger levels [6].

is implemented. This means that selecting only interesting events or triggering on them is needed. The ATLAS trigger system is organized in 3 levels:

- **Level-1 Trigger:** The first-level trigger (L1) works with a subset of information from the calorimeter and muon detectors. It reaches a decision within $2 \mu\text{s}$. L1 is implemented on hardware. All information is stored in pipeline memories until the L1 trigger decision is available. It is designed to accept a maximum of 75000 events/s, effectively reducing the event rate from 40 MHz to 75 kHz.
- **Level-2 Trigger:** Data from Level-1 trigger is transferred to read-out buffers until a L2 trigger is available. L2 is implemented on software. The L2 trigger selects the areas of interest identified by the L1 trigger and then refines this selection, using the full-granularity information from all the detectors, including the ID which is not used on L1. A latency of around 10 ms is expected from the L2 trigger.
- **Event Filter:** The event filter is designed to reduce the event rate from 3.5 kHz to the 200-400 Hz. At this rate, events can be written to disk. It has a similar reconstruction algorithms than the offline one but with looser selection criteria. This

reconstruction takes about 4 seconds per events with an event size of approximately 1.3 Mbyte.

Computing LHC Data Even with triggering, around 15 petabytes of data per year at the LHC are expected. In view such amounts of data, the LHC Computing Grid has been developed. It is a network of computers, which analyze blocks of data and send it back to a centralized computer. This approach is called grid computing. Tier 0 at CERN first processes and divides the data and divides it for its distribution. Twelve Tier 1 sites located in different countries take this data and further process it. The data is then distributed to Tier 2 sites located in more than a 100 universities around the world. This approach allows for a very versatile distribution and analysis of LHC data.

3 Theoretical Context

3.1 The Standard Model of Particle Physics

The Standard Model (SM) is a $SU_C(3) \times SU_L(2) \times U_Y(1)$ gauge theory that describes the fundamental matter particles and their interactions. Developed between the 1960s and 1970s, it explains all the known fundamental particles and their interactions, except for gravity. Since its formulation, it has survived stringent experimental tests that have validated it with a high degree of accuracy. Though it has some limitations and open questions remain, the amount of experimental results it has predicted and explained make it the best theory to explain the interactions of fundamental particles. Unless stated otherwise the bibliographic references used in the chapter come from [16,17].

In the SM there are two types of particles which constitute the building blocks of matter: quarks and leptons. There are six quarks (known as quark flavors: up, down, charm, strange, top, bottom) and six leptons (electron, electron neutrino, muon, muon neutrino, tau, tau neutrino). In total, there are twelve spin- $\frac{1}{2}$ particles or fermions and they are classified into three families (see Figure 3.1). Each family consists of 2 leptons (one charged and one neutral) and two quarks. For all of these fundamental particles, a corresponding antiparticle exists. The first generation of quarks and leptons forms all stable matter. The second or third generation can only be generated in high energy particle collisions or in cosmic ray events because they are unstable and decay immediately to the first generation particles. Neutrinos of the three generations do not decay but they can oscillate between generations. Quarks form bound states called hadrons with integer electrical charge. The top quark cannot hadronize because it decays before hadronization is possible. Hadrons exist either as baryons, made from three quarks, e.g. protons and neutrons, or mesons made from a quark anti-quark pair. The interactions of the particles with each other are determined by particle mediators of spin 1, so called gauge bosons. Quantum electrodynamics (QED), quantum chromodynamics and the weak interaction theory explain how the gauge bosons interact with other particles. The weak interaction theory and QED are unified into a single theory called the electroweak

theory. They are the theory pillars of the the Standard Model.

Three Generations of Matter (Fermions)				
	I	II	III	
mass→	2.4 MeV	1.27 GeV	171.2 GeV	0
charge→	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	0
spin→	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
name→	u up	c charm	t top	γ photon
Quarks	4.8 MeV	104 MeV	4.2 GeV	0
	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	0
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
	d down	s strange	b bottom	g gluon
Leptons	<2.2 eV	<0.17 MeV	<15.5 MeV	91.2 GeV
	0	0	0	0
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
	ν_e electron neutrino	ν_μ muon neutrino	ν_τ tau neutrino	Z⁰ weak force
	0.511 MeV	105.7 MeV	1.777 GeV	80.4 GeV
	-1	-1	-1	± 1
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
	e electron	μ muon	τ tau	W[±] weak force

Figure 3.1: The three families of the Standard Model [18].

3.2 The Electromagnetic Interaction

Electromagnetic interactions take place between electrically charged particles. Together with gravity, it is the most tangible in everyday life. The force carrier of the electromagnetic interaction is the photon (γ). It is a massless particle, electrically neutral and therefore does not interact with itself.

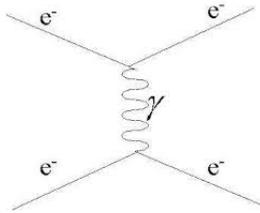


Figure 3.2: Photon exchange by two electrons.

Quantum Electrodynamics (QED) is the relativistic quantum field theory of the electromagnetic interaction. The theory of QED has achieved a remarkable level of compatibility with experimental observations [19]. QED is based on a local symmetry

called U(1). As with any quantum field theory the kinematics and dynamics of the theory can be deduced from the Lagrangian (\mathcal{L})

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \quad (3.1)$$

The QED Lagrangian describes the couplings between charged fermions field ψ to the boson field A^μ . The covariant derivative D_μ and the field strength $F_{\mu\nu}$ are given by:

$$D_\mu = \partial_\mu - ieA_\mu \quad (3.2)$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (3.3)$$

The \mathcal{L}_{QED} is invariant under local U(1) gauge symmetry ($\psi \rightarrow e^{i\theta(x)}\psi$). This gauge invariance implies that the electrical charge is conserved locally. The addition of the mass term of type $m^2 A_\mu A^\mu$, leads to a violation of the gauge invariance. Therefore, the QED gauge boson needs to be massless and it can be directly associated with the photon. The elementary charge e corresponds to the elemental charge and is given by:

$$e = \sqrt{4\pi\alpha_{QED}} \quad (3.4)$$

where α_{QED} is the electromagnetic coupling constant. It is a fundamental parameters of the theory and it determines the strength for the EM interaction. In QED, observables are usually expressed as function of α_{QED} . Using perturbation theory to calculate these observables, one encounters divergences involving the Feynman diagrams with loops with virtual particles. A technique called renormalization is used to get rid of these divergences. The renormalization technique consists in redefining measurable observables at a given energy scale (called renormalization scale μ_0) to include virtual particle corrections, absorbing in this way the infinities. Imposing the independence of the physical observable from μ_0 reveals that α_{QED} depends on the energy scale (Q^2) at which one observes the process. $\alpha_{QED}(Q^2)$ increases as energy increases, from 1/137 at $Q^2 = 0$ to

1/127 at energies corresponding to the mass of the Z boson.

3.3 The Strong Interaction

The strong interaction is responsible for holding hadrons together. The strong interaction is described by a quantum field theory called Quantum Chromodynamics (QCD). QCD is represented by the non-abelian asymmetry SU(3). In this representation color is the charge of QCD. Gluons are the elementary particles that act as the exchange particles for the strong force between quarks, being analogous to the photons in QED. They are massless and electrically neutral. However, as opposed to the photons, they do interact with themselves. That is to say, they are not color neutral. They carry color charge and this fact makes the strong interaction different and more complex from the electromagnetic where the photon has no electrical charge. The strong color charge comes in 3 types: red, green and blue. Anti-quarks have the corresponding anti-color. Leptons do not carry color charge and do not participate in the strong interaction.

The strong coupling constant decreases logarithmically with increasing energies. Hence, quarks and gluons behave as quasi-free particles at high energies (short distances), while at low energies (large distances) quarks are confined into hadrons. These behaviors are called asymptotic freedom and color confinement, respectively.

- **Asymptotic freedom:** This property causes the bonds between strongly interacting particles to become asymptotically weaker as energy increases (or distance decreases). This makes perturbation theory calculations possible for QCD at high energies. At the LHC, energy is sufficiently high to describe pp collision and production of particles i.e. $t\bar{t}$ production, using perturbative QCD. However, once the protons have collided and new particles are created, the quarks and gluons lose energy, the strong coupling constant increases and perturbation theory is no longer applicable (see Figure 3.3).
- **Color Confinement:** When quarks reach a separation distance of around 10^{-15} m, the running coupling constant is so large that new quark-antiquark pairs are

produced from the radiated gluons. That is why quarks and gluons cannot exist in isolation in nature, i.e. confinement. These and anti-quarks join together in myriad combinations to make the mesons and baryons actually observed in collision. In all this debris, there is an unmistakable footprint left behind by the original quark-antiquark pair as sprays of collimated hadrons or jets emerge along the direction of the primordial quarks/gluons.

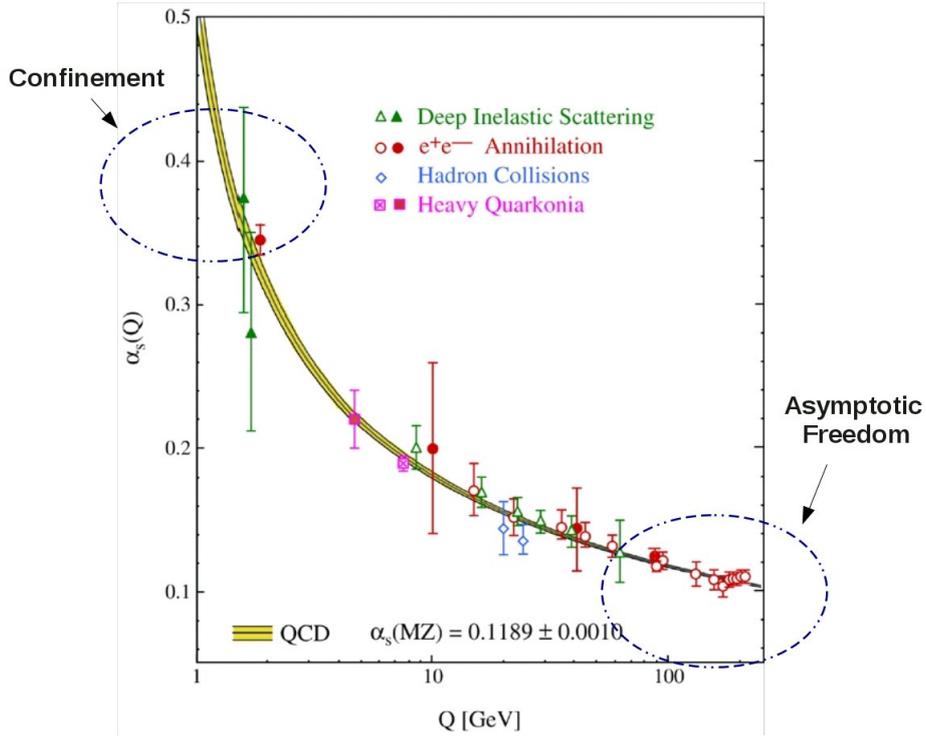


Figure 3.3: The dotted circles shows the energy scale at which asymptotic freedom and color confinement take place [20]

3.4 The Weak Interaction

The weak interaction theory was initially devised to explain beta decays. It affects all fermions. Unlike the electromagnetic and strong mediators, the weak force mediators are massive. The Z^0 and W^\pm bosons and have a mass of 91.2 and 80.4 GeV respectively. and mediate the neutral and charged weak currents, respectively. As they are massive

mediators, the weak force has a very short range of interaction. Because the life of a particle is proportional to the inverse square of the coupling constant of the force causing the decay, the lifetime of particles decaying through the weak force is large. The weak interaction is the only one capable of changing the flavor of a quark or a lepton. Moreover, it also breaks parity-symmetry since W^\pm bosons couple only to left-handed particles (e.g. particles with spin and momentum of opposite direction) right-handed antiparticles.

3.4.1 The Electroweak Theory

The weak and electromagnetic interaction were successfully described as different manifestations of the same fundamental interaction by Glashow, Weinberg and Salam in the 60s. The gauge theory describing both interactions is called unified electroweak theory and is based on the $SU_L(2) \times U_Y(1)$ symmetry group. The local gauge invariance requirement leads to the existence of for bosons: $W_\mu^i (i = 1, 2, 3)$ from $SU(2)$ and B_μ for $U(1)$. g and g' are the coupling constants associated to $SU(2)$ and $U(1)$, respectively. They are related to $e = g \sin \theta_W = g' \cos \theta_W$. θ_W is known as the weak mixing or Weinberg angle.

The photon like the gluon, are massless because of the exact conservation of the corresponding symmetry generators: the electric charge and the eight color charges. The fact that the weak bosons are massive indicates that the corresponding symmetries are broken. In 1964, Brout, Englert and Higgs proposed a the mechanism to explain the breaking of the electroweak gauge symmetry, now called Brout-Englert-Higgs mechanism. It predicts the existence of a spin 0 particle, known as Higgs boson. On July 4th 2012, the ATLAS and CMS collaborations presented evidence for a neutral boson with a measured mass of $126.0 \pm 0.4(\text{stat}) \pm 0.4(\text{sys})$ [21] and $125.3 \pm 0.4(\text{stat}) \pm 0.5(\text{sys})$ [22], respectively. Both results are compatible with the SM Higgs boson. However, more data is needed to confirm if the neutral boson has all the properties ascribed to the SM Higgs boson.

3.5 The CKM Matrix

As mentioned in the previous section, weak charged currents are the only interactions that change the flavor of fermions. The mass eigenstates of fermions are not identical to the weak eigenstates. The transformation between them is described by a 3×3 unitary matrix, the Cabibbo-Kobayashi-Maskawa (CKM) matrix which describes the mixing of the quark eigenstates. The probability for a quark of flavor i to be transformed to a quark of flavor j and emitting a W boson is proportional to $|V_{ij}|^2$ as given by [23]

$$V_{CKM} = \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97428 & 0.2253 & 0.00347 \\ 0.2252 & 0.97345 & 0.0410 \\ 0.00862 & 0.0403 & 0.999152 \end{pmatrix} \quad (3.5)$$

The CKM matrix elements are free parameters of the SM and need to be determined experimentally.

4 Beyond the Standard Model

So far there is no experimental evidence that contradicts SM predictions. However, there are unresolved issues with the theory.

Theoretical limitations:

- The Standard Model describes all of the known forces but the gravitational force. At the Plank scale, $M \approx 10^{19}$, gravity becomes important at the level of fundamental particles. It is believed that the validity of the SM stops at these energies.
- It predicts massless particles. To give mass to particles spontaneous symmetry breaking is included in an unnatural way. The SM offer no explanation for this mechanism.
- The hierarchy problem is also another indication of the incompleteness of the Standard Model. Due to divergent loop corrections to the Higgs mass, renormalization

needs to be applied to calculate its mass. At first order the Higgs mass is,

$$m_H^2 = (m_H^2)_0 - \frac{\lambda_f^2 \Lambda^2}{8\pi^2} \quad (4.1)$$

where the first term is just the bare Higgs mass and the second term is the one-loop quantum correction at first order involving a fermion. λ_f^2 corresponds to the Yukawa coupling. The size of the correction depends on Λ , the scale of the process. If the SM is valid up to the Plank scale and $m_H = 125$ GeV there has to be an ad-hoc fine-tuning to balance the correction term with respect to the first term. This fine-tuning makes the universality of the theory at all energies doubtful. The hierarchy problems arises because of the great difference between the strengths of the electroweak and the gravitational force. Asymptotic freedom and color confinement.

Experimental limitations:

- Gravitational effects on visible matter, radiation, and the large scale structure of the universe, point to the existence of a new type of matter referred to as dark matter. Dark matter accounts for for 84% of the mass of the Universe and it is believe to be composed by previously unseen massive, weakly interacting, stable particles. Visible matter and dark matter account for about 30% of the mass-energy in the Universe. The rest is composed by an unknown energy "dark energy", believed to be responsible for the acceleration of the expansion of the Universe. Neither dark matter nor dark energy are explained by the SM.
- In the SM, the neutrinos are massless and do not oscillate between generations. However in recent years, experiments have shown that neutrinos do oscillated between generations, indicating that they have a finite mass.

These and other gaps have lead physicists to propose extensions to the Standard Model. These theories are commonly known as Beyond the Standard Model (BSM) theories. The LHC is in a position to test the validity of such theories and will do so in

the coming years. The top quark plays a special role in BSM searches. The fact that the top is the only fermion with a weak-scale mass and the only quark which decays before hadronization, makes the top quark a unique window to look for new physics. In this section a brief overview theories for the BSM theories for which the HEPTopTagger will become an important tool.

4.1 Supersymmetry

Supersymmetry (SUSY) [24,25] offers a solution to the hierarchy problem through the stabilization of the mass scale in the Higgs potential. SUSY is a symmetry that allows transformation between fermions and bosons. In the simplest version of supersymmetry, the Minimal Supersymmetric Standard Model (MSSM), each SM fermion has a supersymmetric partner boson (a particle with the same mass and quantum number but different spin, differing by 1/2). There are two main theoretical motivations for SUSY: the hierarchy problem and the unification of forces. Supersymmetry is free of quadratic divergences. So the scale of the theory can be extended without introducing the hierarchy problem. Because SUSY has not been observed yet, the super-partners should be much heavier than their SM counterparts. Therefore, if supersymmetry is valid it must be a broken symmetry. This broken symmetry would introduce new particles at in the TeV range that the LHC could detect. The second motivation is that with SUSY the coupling constant of the electromagnetic, weak and strong interaction are unified at an energy scale of around 10^{16} GeV.

The most important experimental motivation for SUSY is, that it offers an natural candidate for dark matter. SUSY predicts a light supersymmetric particle, which would be massive, neutral and would only interact via the weak force. This would fulfill the requirements of a weakly interacting massive particle as stipulated in the Cosmological Model. Evidence for SUSY would come in a long decay cascade of particles which include leptons and jets with very high multiplicities and also very high missing transverse energy and momentum as the lightest supersymmetric particle will escape the detector, leaving an unbalanced momentum in the transverse plane. An example of a supersymmetric

event in which two stops quark are produced is shown at the top in Figure 4.1. These two stops would decay into a top or b quark and neutralinos which would escape undetected.

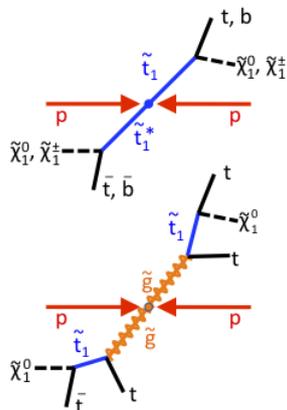


Figure 4.1: Hypothetical production of stop quark pair and gluinos in a pp collision.

4.2 Z' and Kaluza-Klein Gluons

The Z' is an hypothetical gauge boson that arises from various models of physics beyond the Standard Model [26]. It is basically a placeholder for hypothetical heavy gauge bosons, having the same couplings as a SM Z bosons. Topcolor, Little Higgs, and Kaluza-Klein models also predict one or several Z' bosons. The main characteristic of the Z' in these models would be its heavy mass. This makes the $t\bar{t}$ an ideal channel to look for this particle. Kaluza-Klein (KK) excited states are predicted by theories with extra dimensions [27, 28]. Gluons, electroweak gauge bosons and gravitons are predicted to exist in KK excited states. KK-gluons would have a very high production rate at the LHC. Because of their expected high mass on the order of TeV, KK gluons are also expected to decay predominantly into $t\bar{t}$.

5 The Top Quark in the Standard Model

5.1 Introduction

The top quark is the heaviest known elementary particle. With a mass five orders of magnitude greater than the first generation quarks, it is close to the electroweak symmetry breaking scale. This and the fact that previously unknown heavier particles will decay into top quarks make it a very important window for BSM searches. The existence of a third generation of quarks was postulated in 1973 by Makoto Kobayashi and Toshihide Maskawa to explain the CP violations in kaon decay. After its prediction it was actively sought and finally discovered in 1995 at Tevatron. Since then, the measurements of the top quark mass have reached an accuracy in the range of 1 GeV,

$$m_t^{\text{Tevatron}} = 173.2 \pm 0.9 \text{ GeV [29, 30]}$$

The production of the top quark in the SM framework as well as its decay modes is described in the following sections.

5.2 Top Pair Quark Production

After the shutdown of Tevatron in April 2012, the LHC became the only accelerator able to create sufficiently high energies to produce top quarks. Several QCD processes contribute to the $t\bar{t}$ production at hadron colliders (see Figure 5.1). At the LHC the main production mechanism is the gluon fusion accounting to 85% percent of the the top total top quark cross section. The rest is produced in quark-antiquark annihilation.

	$\sigma_{NLO}(\text{pb})$	$q\bar{q} \rightarrow t\bar{t}$	$gg \rightarrow t\bar{t}$
Tevatron ($\sqrt{s} = 1.96 \text{ TeV}, p\bar{p}$)	$6.77 \pm 9\%$	85%	15%
LHC ($\sqrt{s} = 14 \text{ TeV}, pp$)	$833 \pm 15\%$	10%	90%

Table 5.1: Cross sections at next-to-leading order for $t\bar{t}$ production via the strong interaction at the LHC and Tevatron [31, 32]. Errors in the cross section come mostly from parton distribution function uncertainties.

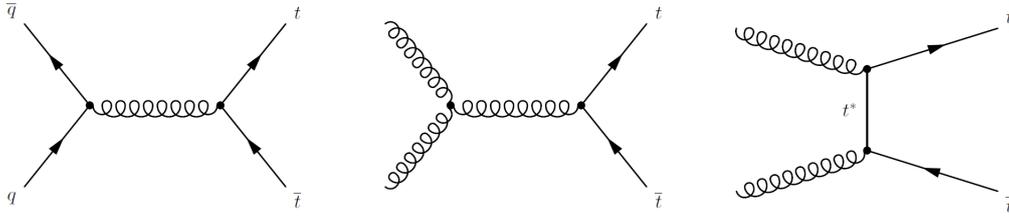


Figure 5.1: The leading order Feynman diagrams for top quark production. The left diagram shows quark-antiquark annihilation. The center and left ones show gluon-gluon fusion.

The LHC has been appropriately called a top factory. It produced around 800,000 $t\bar{t}$ pairs in 2011. At full luminosity it will collect 10 times this quantity [1]. This will allow very detailed studies on the properties of the top quark such as mass, couplings and asymmetries.

5.3 Top Quark Decay

The top quark has a life-time of approximately 5.0×10^{-25} s. With such a low lifetime, it is the only quark to decay before hadronization, which has a time scale of around 3×10^{-24} s. In the Standard Model and as prescribed by the CKM matrix, the only possible decays for the top quark, are $t \rightarrow bW^+$, $t \rightarrow sW^+$ and $t \rightarrow dW^+$. The probability of these decays is given by $|V_{tq}|^2$ with $q = b, s, d$, respectively. As given by the CKM matrix, approximately 99.8% of the top quarks will decay into a W boson and a b-quark, with the other decays being negligible. The top decay final states are therefore determined by the decay of the W boson (see Figures 5.2 to 5.4). About 33% of the times the W boson decays into a charged lepton and a neutrino (leptonic decay) and about 67% into a quark-antiquark pair (hadronic decay). As a result, there are three kinds of top pair decays: hadronic, lepton+jets, and dileptonic (see Figures 5.5 and 5.6). The branching ratios of the top quark decay follow from the individual branching ratios of the W boson decay modes.

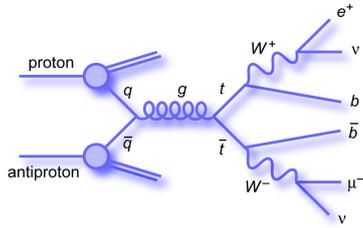


Figure 5.2: Top pair decay in the leptonic channel [43].

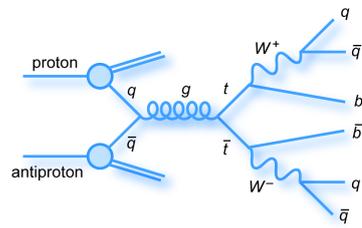


Figure 5.3: Top pair decay in the lepton+jets channel [43].

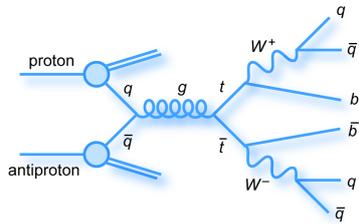


Figure 5.4: Top pair decay in the all-hadronic channel [43].

Top Pair Decay Channels

$c\bar{s}$	electron+jets	muon+jets	tau+jets	all-hadronic	
$u\bar{d}$	electron+jets	muon+jets	tau+jets		
$\tau^+\tau^-$	dileptons	dileptons	dileptons		tau+jets
$e^+\mu^-\tau^+$	dileptons	dileptons	dileptons		muon+jets
W decay	e^+	μ^+	τ^+	$u\bar{d}$	$c\bar{s}$

Figure 5.5: Top pair channel decays [43].

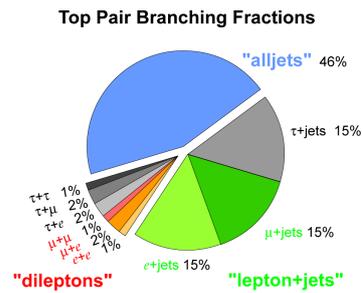


Figure 5.6: Top pair branching ratios [43].

6 The HEPTopTagger

6.1 Background

6.1.1 Sequential Recombination Algorithms

As mentioned in Section 3.3, quarks and gluons are not directly seen in the detector due to confinement. Due to hadronization, they appear in the detector as collimated sprays of hadrons. By analyzing the tracks and energy deposits of these hadrons it is possible to figure out the characteristics of the original quark or gluon. Because partons have divergent probabilities in perturbative QCD, a jet is not universally defined. The definition of jet depends of the prescription on how to group particles and how to assign momentum to the resulting jet, i.e. a jet algorithm. There are several ways to construct jet algorithm. However, a jet algorithm should always be infrared and collinear safe. This means that jets found in an event, should remain unchanged by a collinear splitting or the addition of a soft emission. Sequential recombination algorithms fulfill these requirements. They have become the standard jet finder and jet reconstruction algorithms in ATLAS and CMS. These types of algorithms sequentially merge pairs of objects i and j by adding their four-momenta. A sequential recombination algorithm first creates a list of all objects (either hadrons, topo-cluster tracks) in an event. Then, two distances are obtained for each of these objects: one between the object and its closest neighbor, as defined in Equation 6.1 and the distance between the object and the beam as defined in Equation 6.2 [48].

$$d_{ij} = \min(p_{Ti}^{2n}, p_{Tj}^{2n}) \frac{\Delta R_{ij}^2}{R^2}, \quad (6.1)$$

$$d_{iB} = p_{Ti}^{2n} \quad (6.2)$$

In Equations (6.1) and (6.2), p_{Ti} is the transverse momentum of the object. In Equation 6.1, n is an integer, $\Delta R_{ij} = \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}$ is a measure of the opening

angle between objects i and j . These two distance are then compared:

- If $d_{iB} < d_{ij}$ then the object is “closer” to the beam than to other objects in the event, so it is defined as a jet and removed from the list.
- If $d_{iB} > d_{ij}$ then the two objects i and j are combined into one (by adding their four momenta), forming a new object. This procedure continues until there are no more objects on the list.

The difference between different jet algorithms comes from the value of n in the exponent of p_{Ti} in Equation 6.1:

- $n = +1$: the k_T algorithm. Objects with smaller p_T tend to be clustered first.
- $n = 0$: the C/A algorithm. Objects are combined based only on their angular separation from one another and the beam.
- $n = -1$: the anti- k_T algorithm. Objects with higher p_T tend to be clustered first.

ΔR_{ij} is the angular distance between objects i and j . The jet distance parameter R controls the size of the jets in $y - \phi$ space, and can be roughly referred to as the jet “radius”. However, what this parameter does exactly is to ensure that particles separated by $\Delta R < R$ at a given clustering stage are not combined and that an object can only be promoted to a jet if there are no other objects within $\Delta R < R$ [44].

6.1.2 Boosted Physics and Jet Substructure

At the LHC, the heaviest particles of the Standard Model (W^\pm , Z^0 , Higgs boson, top quark) and new possible particles (predicted by BSM theories in the same mass range or even heavier) can be produced with a transverse momentum greatly exceeding their rest mass i.e. they are boosted. When boosted particles decay, they exhibit a highly collimated topology in the detector (see Figure 6.1). A complete overview of boosted objects at the LHC is given in the report from BOOST 2011 [44].

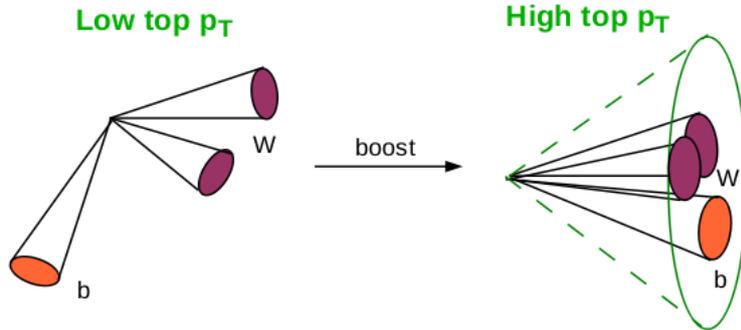


Figure 6.1: Comparison between top quark decays, with low and high p_T .

The energy cones coming from boosted particles will overlap each other. For a boosted particle, the higher its p_T , the closer its decay products will be as given by

$$R \sim \frac{2m}{p_T} \quad (6.3)$$

This is an indication that standard reconstruction techniques are not adequate to reveal the precedence of a boosted heavy particle. Many efforts have been performed to identify and reconstruct boosted W bosons, Higgs, etc [34–36]. This study focuses on the boosted top quark. In the case of the boosted top quark, standard top identification is not effective: b-tagging is difficult as a consequence of crowded and unresolved tracks, W decay products are not isolated from each other, and the measured top mass may differ from m_t due to an increase in QCD radiation. New tools have been developed to identify and reconstruct boosted top quarks. As previously hinted before, a strong theoretical motivation to study highly boosted top quarks is that heavy s-channel resonances will decay to $t\bar{t}$ pairs. The higher the new physics mass scale is pushed by LHC searches, the more boosted these top quarks will become if these new states exist.

6.2 The HEPTopTagger Algorithm

The HEPTopTagger (Heidelberg-Eugene-Paris) is an algorithm by Plehn et al. [33] designed to tag and fully reconstruct hadronically decaying top quarks. The main features of the HEPTopTagger are the following:

- It focuses on top quarks with p_T in the range of 200 – 500 GeV (moderately boosted) by choosing a jet distance parameter $R_{fatjet} = 1.5$. From Equation 6.3, the lower the p_T of the top, the less collimated its decay products will be. Hence, a larger jet distance parameters necessary to identify and reconstruct top quarks with $p_T > 200$ GeV (see Figure 6.2).

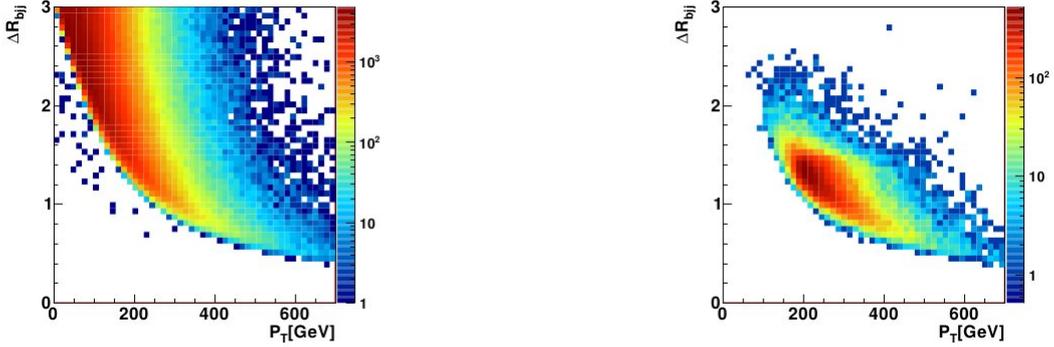


Figure 6.2: Left: Partonic ΔR_{bjj} vs p_T for a Standard Model $t\bar{t}$ sample. Right: the same correlation, but only for tagged top quarks and based on the reconstructed kinematic properties. Taken from [33].

- It uses the C/A algorithm with a mass drop criterion⁴. As explained in subsection 6.1.1, the C/A algorithm cluster particles closest in angle. Therefore, the jet has an "angular-aware" substructure. This improves the mass resolution of the reconstructed object [35]. In addition, QCD-initiated jets processed by this method produce a relatively featureless mass spectrum. However, in the absence of any momentum scale, the last clustering step often involves soft radiation on the edges and therefore it is unrelated to the heavy object. C/A based substructure algorithms must therefore work backwards iteratively through the jet clustering and stop when the subjets meets some specific hardness requirement. A mass drop criterion is applied in the case of the HEPTopTagger, until all objects have a mass lower than a certain parameter. This ensures that wide angle underlying events

⁴The anti-kt jet algorithm clusters high p_T objects first, even if they are geometrically separated by a large distance. This algorithm is therefore not suited for the HEPTopTagger approach where the substructure of the fat jet is analyzed by undoing the last clustering steps.

(UE) and pile-up (PU) are removed from the event.

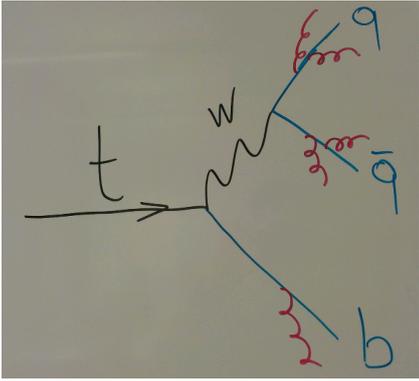
- The mass-drop procedure removes a substantial amount of wide-angle UE/PU. However, as moderately boosted regimes are explored and jet radii become larger ($O(1)$), UE and PU contamination remains a major problem, as they scale to jet mass as R^4 [38]. To further remove unwanted radiation and refine the subjets, a filtering procedure is applied to the jet [39]. The filtering procedure consists on reclustering the substructure constituents with the C/A algorithm, using $R = \min(0.3, \Delta R_{subjets}/2)$. This facilitates the capture of possible gluon radiation in the heavy particle decay, while still eliminating much of the UE/PU.

6.2.1 Steps

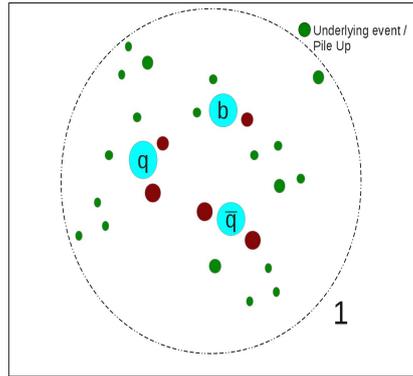
Due to their impact on the HEPTopTagger’s performance, some parameters were investigated more thoroughly than others. In the following introduction of the algorithm, all tunable parameters are highlighted. To illustrate the workings of the algorithm a toy example where the quarks emit gluon radiation, is given after each of the steps.

The algorithm proceeds in the following steps:

1. **Fat jet** Define a fat jet J using the C/A algorithm using $R_{fatjet} = 1.5$ (**fat jet distance parameter**).

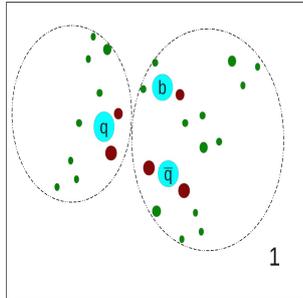


Start with the following topology.

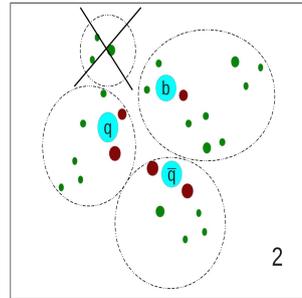


Fat jet clustered with $R_{fatjet} = 1.5$

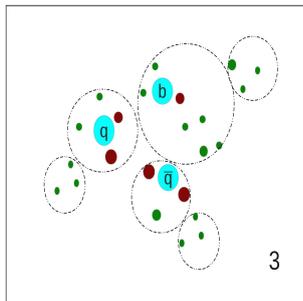
2. **Mass-Drop:** Undo the last clustering step of C/A in the jet J and obtain two subjets j_1 and j_2 with $m_{j_1} > m_{j_2}$. If jet j_1 has 80% of the mass of the original fat jet (m_J) (mass drop criterion) or more, discard jet j_2 . Otherwise keep both. Proceed iteratively with all subjets until all have a mass $m_{cut} < 30$ GeV (**parameter** m_{cut}).



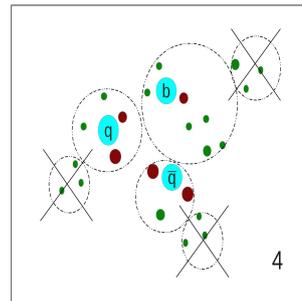
1. Uncluster the jet. Obtain 2 subjets. None below mass drop criterion. Keep both.



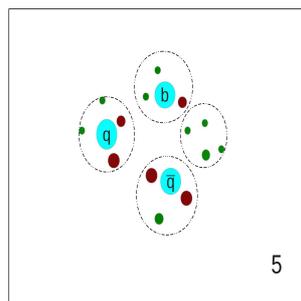
2. Uncluster again. Obtain 4 subjets. One of them below mass drop criterion. Drop.



3. Uncluster again. Obtain 4 subjets.

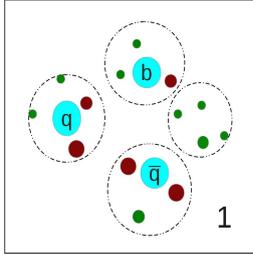


4. Three of them below mass drop criterion. Drop.

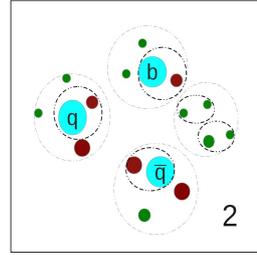


5. Obtain four subjets to be filtered.

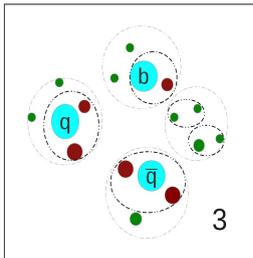
3. **Filtering:** Take all the three-pairings of the previously obtained subjets and filter them (using the C/A algorithm) with a jet distance parameter $R_{\text{filter}} = \min(0.3, \Delta R_{jk}/2)$ (**parameter** ΔR_{jk}). Next use the five (**parameter** N_{filt}) hardest filtered constituents and calculate their jet mass. For less than five filtered constituents use all of them. Finally, select the set of three subjet pairings with a jet mass closest to the top mass (m_t). The filtering step of the involves running the C/A algorithm using a distance parameter than can be different depending on the event topology ("dynamic radius"). In ATLAS, jets are calibrated using the anti- k_T algorithm. There is no calibration using the C/A algorithm. That is why a dedicated calibration using radii 0.2 - 0.5 in steps 0.05 was implemented for jets clustered with the C/A algorithm. If the dynamic distance parameters takes a value in between, it is rounded to the closest calibrated value. This calibration is necessary to obtain high tagging efficiency and background rejection [49].



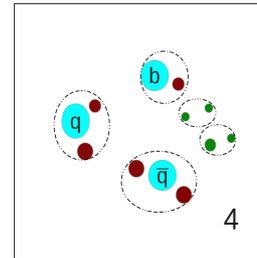
1. Apply filtering to each of these subjets.



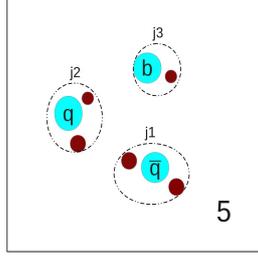
2. Filtering, starts by clustering the closest objects within each subjet previously obtained.



Filtering continues. However for the green objects on the left filtering has stopped because there are no more objects at distance $< R_{\text{filt}}$.



Filtering stops. We obtain 5 filtered subjets. With $N_{\text{filt}} = 5$, select all and take the three jet pairings which are closest to the top mass.



We are left with 3 subjects ordered in p_T to be tested for the following mass relationships.

4. Top mass and W Boson Mass Requirements

The knowledge of the top quark and W boson masses provide two constraints, $m_{123} = m_t$ and $m_{jk} = m_W$ for one (j, k) in the identification of the top quark. An additional mass relation that can be exploited. The three subjects j_k ignoring smearing and assuming $p_i^2 \sim 0$ give

$$m_t^2 \equiv m_{123}^2 = (p_1 + p_2 + p_3)^2 = (p_1 + p_2)^2 + (p_1 + p_3)^2 + (p_2 + p_3)^2 = m_{12}^2 + m_{13}^2 + m_{23}^2 \quad (6.4)$$

which is the surface of a sphere with radius m_t in (m_{12}, m_{13}, m_{23}) space. For a fixed m_{123} , one can choose exactly two variables to describe the kinematics of the event: m_{23}/m_{123} and $\arctan(m_{13}/m_{23})$ which means that m_{12}/m_{123} can be derived as,

$$1 = \left(\frac{m_{12}}{m_{123}}\right)^2 \left(1 + \left(\frac{m_{13}}{m_{123}}\right)^2\right) + \left(\frac{m_{23}}{m_{123}}\right)^2. \quad (6.5)$$

Assuming $m_{123} = m_t$, the condition $m_{12} = m_W \pm 15\%$ reads as Equation 6.6. The selection criteria shown in equations 6.7 and 6.8 are built similarly. In the $m_{23}/m_{123} - \arctan(m_{13}/m_{23})$ space top quark candidates lie in well defined regions that can be separated by selection criteria presented in equations 6.6, 6.7 and 6.8 as seen in Figure 6.3. A more detailed motivation is given in [33]. The selection

criteria are the following,

$$0.2 < \arctan \frac{m_{13}}{m_{12}} < 1.3 \quad \text{and} \quad R_{\min} < \frac{m_{23}}{m_{123}} < R_{\max} \quad (6.6)$$

$$R_{\min}^2 \left(1 + \left(\frac{m_{13}}{m_{12}} \right)^2 \right) < 1 - \left(\frac{m_{23}}{m_{123}} \right)^2 < R_{\max}^2 \left(1 + \left(\frac{m_{13}}{m_{12}} \right)^2 \right) \quad \text{and} \quad \frac{m_{23}}{m_{123}} > 0.35 \quad (6.7)$$

$$R_{\min}^2 \left(1 + \left(\frac{m_{12}}{m_{13}} \right)^2 \right) < 1 - \left(\frac{m_{23}}{m_{123}} \right)^2 < R_{\max}^2 \left(1 + \left(\frac{m_{13}}{m_{12}} \right)^2 \right) \quad \text{and} \quad \frac{m_{23}}{m_{123}} > 0.35 \quad (6.8)$$

The HEPTopTagger algorithm now proceeds to construct exactly three jets subsets j_1, j_2, j_3 from the five filtered constituents, ordered by (p_T). If the subsets masses (m_{12}, m_{13}, m_{23}) satisfy **one** of the selection criteria just mentioned, the four vectors of the 3 subsets are added and the sum is taken as a top candidate.

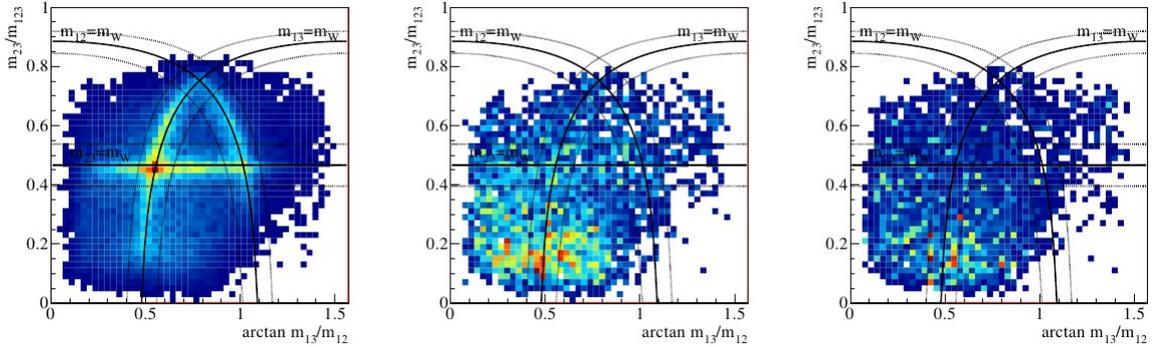


Figure 6.3: The distribution of events in the $\arctan m_{13}/m_{12}$ vs m_{23}/m_{123} . Samples shown are $t\bar{t}$ (left), W +jets (center) and pure QCD jets (right). More densely populated regions of the phase space appear in red [33].

5. Finally, require the combined p_T of the 3 subsets to exceed 200 GeV.

7 ATLAS Event Simulation

The ATLAS collaboration has developed a detailed simulation of the ATLAS detector based on Geant4 (G4) [51, 52]. This simulates the ATLAS detector with an accurate description of the detector material and geometry. For any given hard process, a Monte-Carlo generator will produce a set of events containing lists of final-state particles and their four-momenta with respect to the origin. These four-vectors are feed to the ATLAS G4 Model and then they are propagated through the ATLAS detector. The interactions with the detector material, such as charge deposits in the tracking detectors and showering of particles in the calorimeter material. Interactions between particles and inactive material such as support structures and cabling are also modeled. The energy deposited by particles in the active detector material is converted into detector signals with the same format as the ATLAS detector read-out. The simulated detector signals are in turn reconstructed with the same reconstruction software as used for the data. The Geant4 parameters are adjusted in accordance to match results coming from test-beam analyses.

The Monte Carlo samples used for this analysis are:

- Semileptonic $t\bar{t}$ generated with the MC@NLO generator [46, 47].
- Z' bosons with masses of 1000, 1600 and 3000 GeV generated with the MC@NLO generator.
- W+jets generated with the ALPGEN generator [50].

All these samples run through the complete ATLAS infrastructure simulation. We obtain objects with four-vectors in which the HEPTopTagger can work on. To analyze this objects the HEPTopTagger is implemented in the ATLAS software.

8 Cut-based Optimization

8.1 Event Selection

The primary objects used for this analysis are jets clustered with the Cambridge-Aachen (C/A) algorithm [45], with a large distance parameter $R = 1.5$ ("fat jets"). The parameter optimization is done using a sample containing pairs of top quark decays in which one W boson decays hadronically and other decays into a muon and a neutrino (lepton+jets). To optimize the HEP-TopTagger parameters, it is preferable to select only muons coming from the W boson from the top quark, decays into a muon and a neutrino. Because the muon has a very clear signature in the ATLAS detector, the efficiency with which the HEP-TopTagger finds and reconstructs the hadronic top can be more easily understood and measured. As background, a sample in which W decays into leptons plus jets in the event (W+jets) is used. Although there are other backgrounds that can fake a top signal, around 90% of the background comes only from W+jets. Figure 8.1 shows data and MC mass distributions for C/A jets with $R_{fatjet} = 1.5$ before applying the HEP-TopTagger. Therefore, other backgrounds are ignored for this optimization unless stated otherwise.

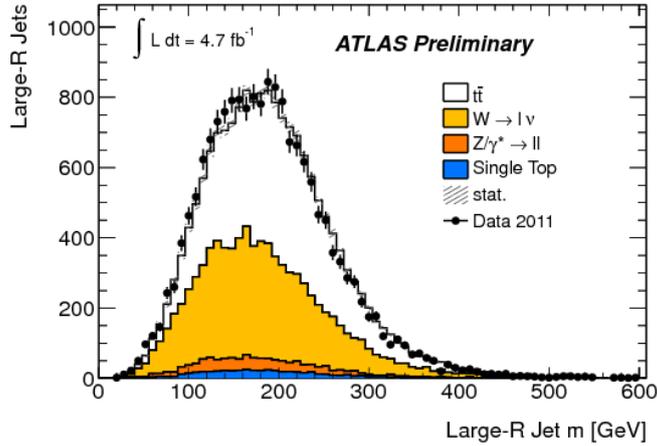


Figure 8.1: Mass distribution for C/A jets with $R_{fatjet} = 1.5$ before applying the HEP-TopTagger [48].

The event selection is the following:

- 1 muon ($p_T > 25$ GeV), 0 electrons.
- Missing transverse energy > 20 GeV.
- ≥ 4 jets (anti- k_T , $R = 0.4$)
- C/A fat jet ($R = 1.5$), $p_T > 200$ GeV, $|\eta| > 2.5$

8.2 HEPTopTagger Parameter Tuning

As explained in subsection 6.2.1, the HEPTopTagger has several internal parameters that can be tuned to optimize its performance. The parameters chosen to optimize the performance are:

- m_{cut} : the mass cut parameter.
- R_{filt} : the filtering distance parameter.
- N_{filt} : the number hardest filtered constituents after the filtering step
- f_W : the width of the window around the W boson mass.

The parameter f_W is related to $R_{min,max}$ from Eq. 6.6 where $R_{min,max} = (1 \pm f_W) \frac{m_W}{m_{top}}$. The parameters and the values tested for the optimization together with the default parameters are presented in Table 8.1.

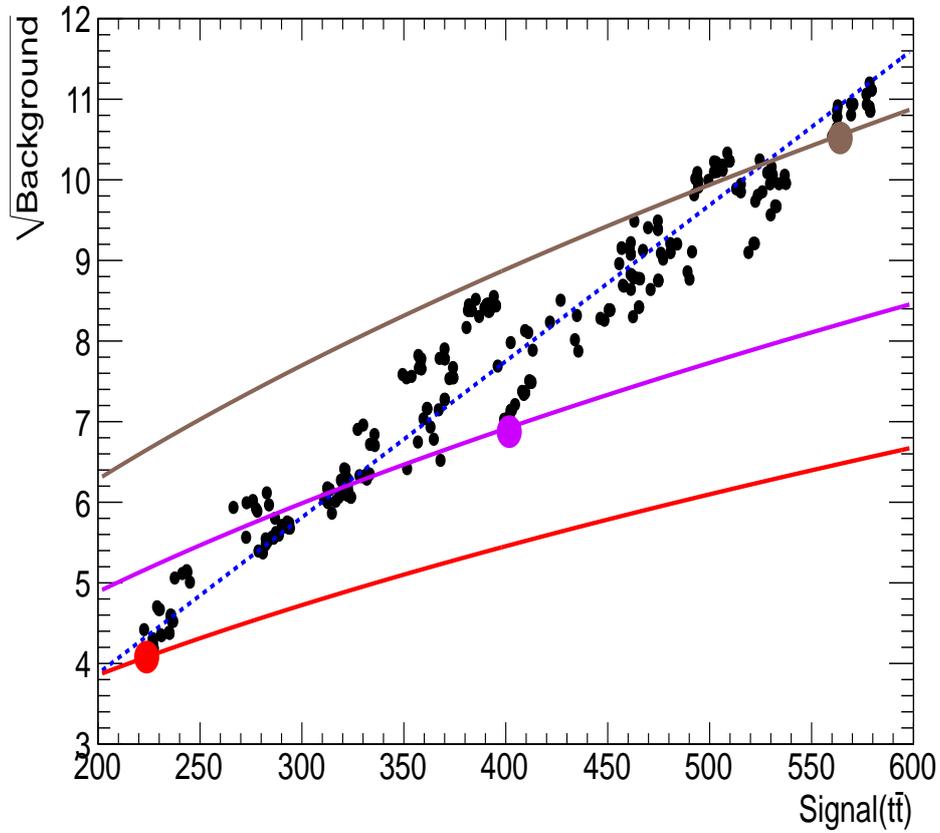
HEPTopTagger Parameters	Optimization Settings	Default Settings
m_{cut} (GeV)	30,40,50,60,70	30
R_{filt}	0.2,0.3,0.4,0.5	0.3
N_{filt}	3,4,5,6,7	5
f_W	10,15,20	15

Table 8.1: HEPTopTagger parameters tuned during the optimization (left column). Possible values the parameters can take (center column). Default parameter values (right column).

To understand the impact which different parameters settings have on the HEPTopTagger's performance, the algorithm is applied to the signal and background samples using the 300 possible parameters setting combinations that can be obtained from the central column values in Table 8.1. To accept an event as a top candidate, the reconstructed top mass is required to be within a range of 120-200 GeV (mass window selection criterion). The number of top candidates in signal and background for the 300 possible parameter configurations is obtained. These numbers are shown in Figure 8.2 as black dots. On the x-axis, the number of top candidates on the signal sample (S) and on the y-axis, the square root of the number of tagged top quarks on the background sample (\sqrt{B}) are shown respectively. Out of the 300 points in Figure 8.2, three points are chosen. These three points are located in areas of high purity (S/B) but low signal efficiency (S) (red point), medium signal efficiency and medium purity (violet point) and high signal efficiency and low purity (brown point). These parameter settings are defined as tight, medium and loose respectively. The parameter settings for these points are obtained and shown in Table 8.2. The mass distributions obtained with the tight, medium and loose HEPTopTagger parameter settings are shown in Figure 8.3, with the corresponding figures of merit obtained for each configuration. Figure 8.2 shows that most parameter configurations cluster around a line of constant S/\sqrt{B} . Hence, there is no tuning of the HEPTopTagger parameters that can increase much further this figure of merit. As S increases with the loosening of the parameters, the purity (S/B) decreases. Because the medium settings are the same as the default settings, from now on default means medium settings as well. Lines of constant S/B are drawn crossing each of the chosen points with the same color as the point, with their corresponding value indicated below the figure.

	Default Settings	Tight Settings	Medium Settings	Loose Settings
m_{cut} (GeV)	30	30	30	70
R_{filt}	0.3	0.2	0.3	0.5
N_{filt}	5	4	5	7
f_W	0.15	10	0.15	0.30

Table 8.2: HEPTopTagger parameter settings in different regions of the $S - \sqrt{B}$ plane in Figure 8.2. Note that the default settings are the same as the medium settings.



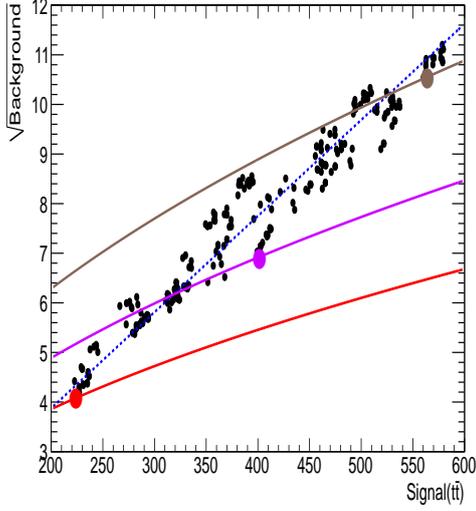
--- $S/\sqrt{B} \simeq 52$

— $S/B = 5$

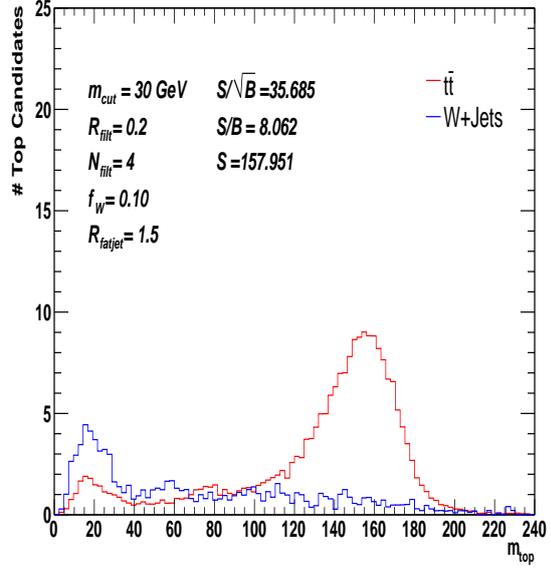
— $S/B = 8$

— $S/B = 14$

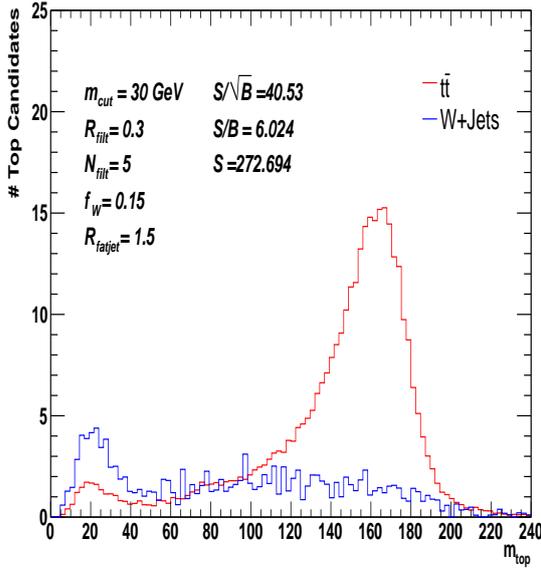
Figure 8.2: Results of 300 HEPTopTagger parameter configurations shown in the $S - \sqrt{B}$ plane. $S=t\bar{t}$, $B=W+\text{Jets}$. The points representing each possible combination of the parameters shown in Table 8.1 cluster around a line of constant S/\sqrt{B} . The color points represent a particular parameter configuration chosen for highest purity (S/B) (low in the plot) in three different regions of the $S - \sqrt{B}$. The crossing color lines represent lines of constant S/B for the chosen points.



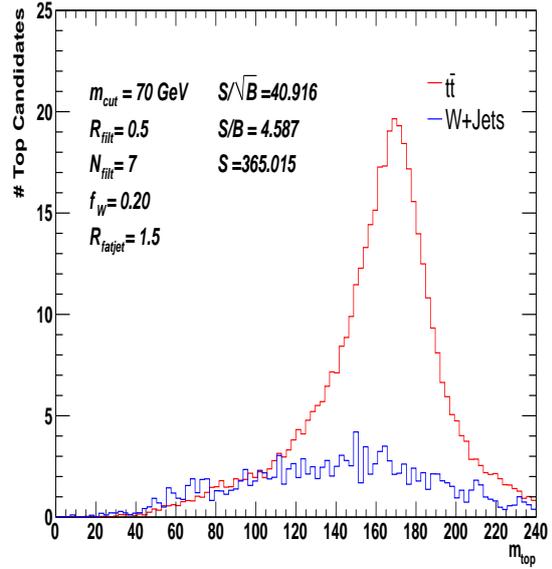
a) Parameter distributions in $S - \sqrt{B}$.



b) Top mass distribution for tight settings.



c) Top mass distribution for default settings.



d) Top mass distribution for loose settings.

Figure 8.3: Mass distributions as reconstructed using the tight, default and loose parameter configurations. Parameters used and obtained S/B and S/\sqrt{B} are shown on the plot. The mass range from 0-240 GeV is shown, however the figures of merit are calculated in the mass range 120 – 200 GeV. Figure 8.2 is shown here to indicate the location of each parameter configuration in the S/\sqrt{B} plane.

Each of the 300 parameter configurations shown in Figure 8.2 were obtained using $R_{fatjet} = 1.5$. As mentioned in subsection 6.2, the top pair quark cross section falls very steeply with increasing transverse momentum. A natural step in this optimization is to increase R_{fatjet} and observe if it is possible to increase the signal efficiency while keeping the background low. As observed in Figure 6.2, the rate of top quarks at low p_T is very high. If R_{fatjet} is increased, the number of signal events will consequently increase. Thus, R_{fatjet} is increased from 1.5 to 1.8 and using the tight, default and loose parameter setting configurations. The mass distribution obtained for signal and background are shown in Figure 8.4. From this one can see that indeed the signal efficiency is increased. However S/B is reduced in comparison with the same settings at $R = 1.5$. Hence, the background increases more than the signal when the fat jet distance parameter is increased. This increase in fat jet distance parameter was also explored by Plehn et. al [41]. The conclusion was the same: there is increase in the signal yield, however S/B decreases. The reason for this is that as we open the fat jet distance parameter, more tops are "captured" , but we also capture much more QCD radiation since this is the p_T range where it is dominant. A comparison of different figures of merit obtained with different parameter settings is given in Table 8.3.

	S	S/B	S/\sqrt{B}
Tight $R = 1.5$	158.0	8.06	35.70
Tight $R = 1.8$	187.5	6.74	35.55
Default $R = 1.5$	272.7	6.02	40.53
Default $R = 1.8$	334.3	5.22	41.74
Loose $R = 1.5$	365.0	4.59	40.13
Loose $R = 1.8$	504.0	3.60	42.65

Table 8.3: Comparison of figures of merit for $R_{fatjet} = 1.5$ and $R_{fatjet} = 1.8$. Looser parameter settings increase signal efficiency while decreasing signal over background. S/\sqrt{B} remains approximately constant for all parameter settings.

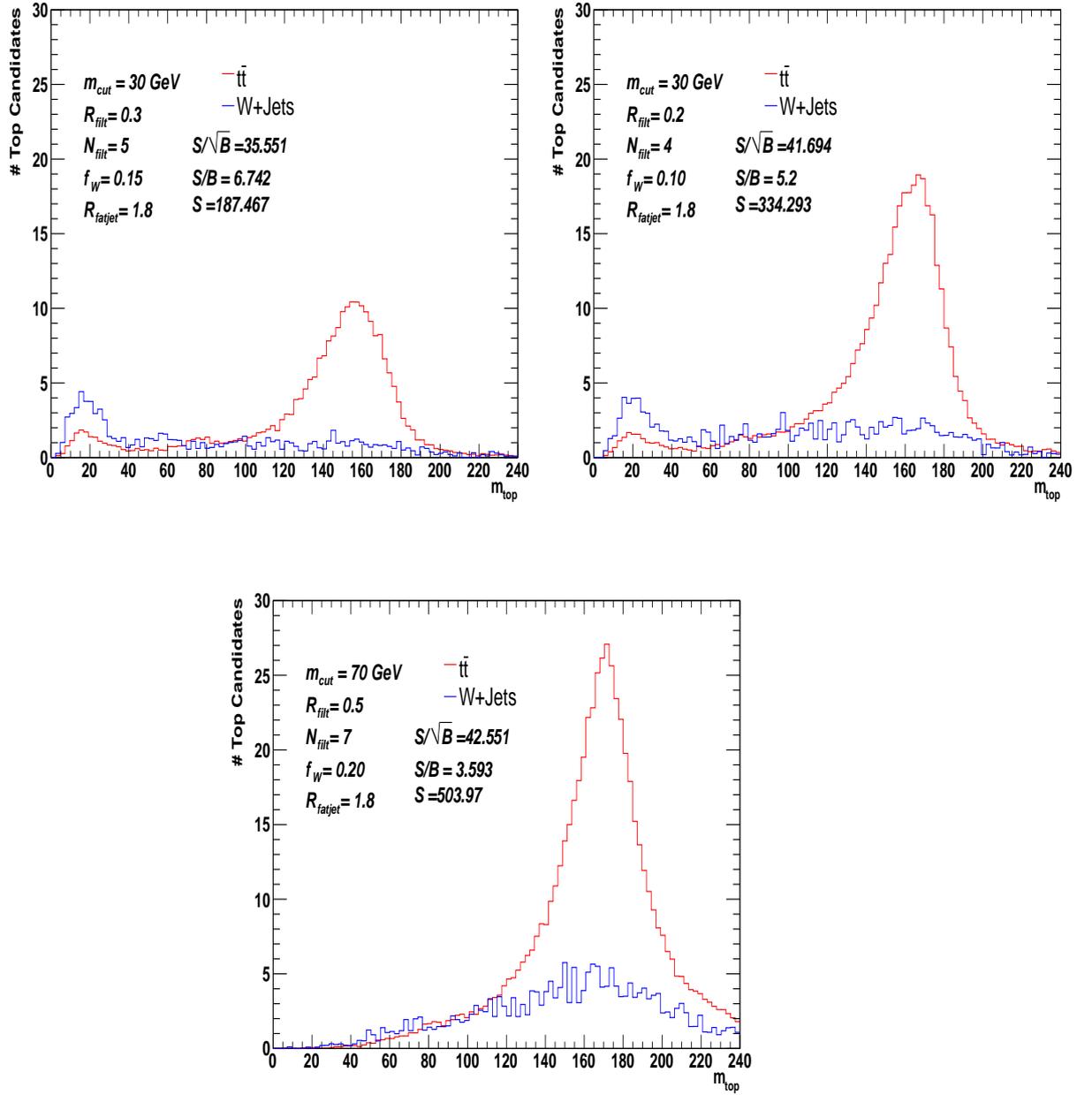
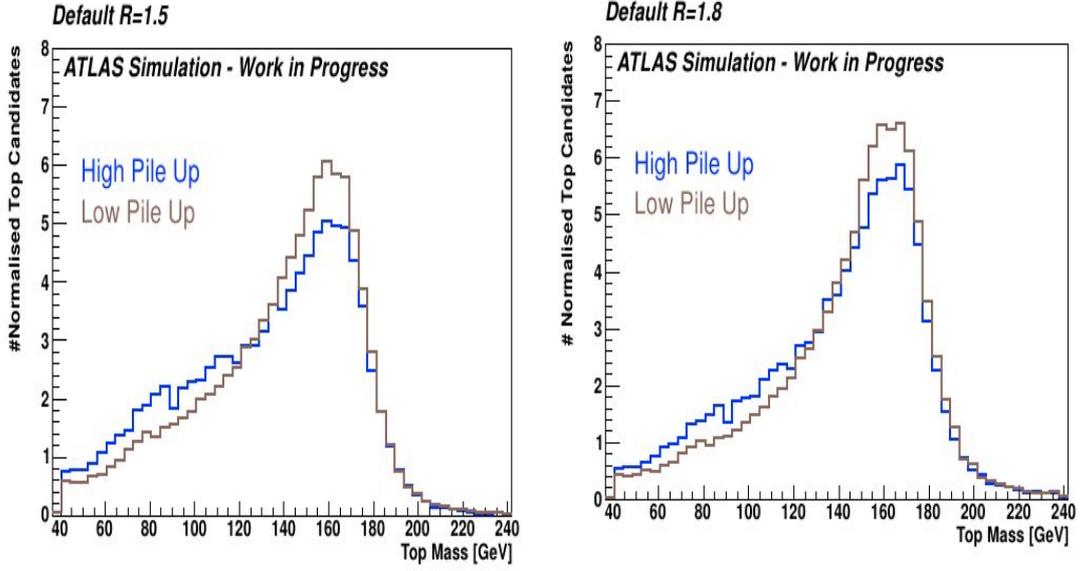
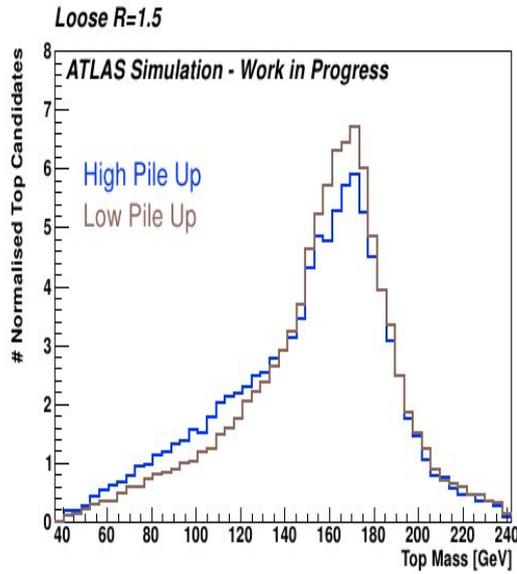


Figure 8.4: Mass distributions as reconstructed using the tight, default and loose parameter configurations with $R_{fatjet} = 1.8$. Parameters used and obtained S/B and S/\sqrt{B} are shown on the plot. The mass range from 0 – 240 GeV is shown, however the figures of merit are calculated with in the range 120 – 200 GeV.



a) default with $R_{fatjet} = 1.5$

b) Default with $R_{fatjet} = 1.8$



c) Loose with $R_{fatjet} = 1.5$.

Figure 8.5: Signal ($t\bar{t}$) and background (W+jets) added mass distributions shown in high pile-up (blue line) and low pile-up (brown lines) conditions for the different HEP-TopTagger configurations. Low pile-up = $\mu < 10$ and high pile-up = $\mu > 10$.

The HEPTopTagger was designed to be resilient against high pile-up conditions. We test this by dividing events into two pile-up regions: low pile-up ($\mu < 10$) and high pile-up ($\mu > 10$). Figure ?? show three different HEPTopTagger parameter configurations for the two different pile up regions. For better visualization signal and background are added into one histogram and then high pile-up and low pile-up histograms are compared. A top quark mass peak is clearly visible in all distributions. The shape becomes more distorted, in high pile-up regions, for looser parameter settings and bigger fat jet distance parameter R. However the top mass peak remains always visible.

8.3 Monte Carlo Validation with LHC Data

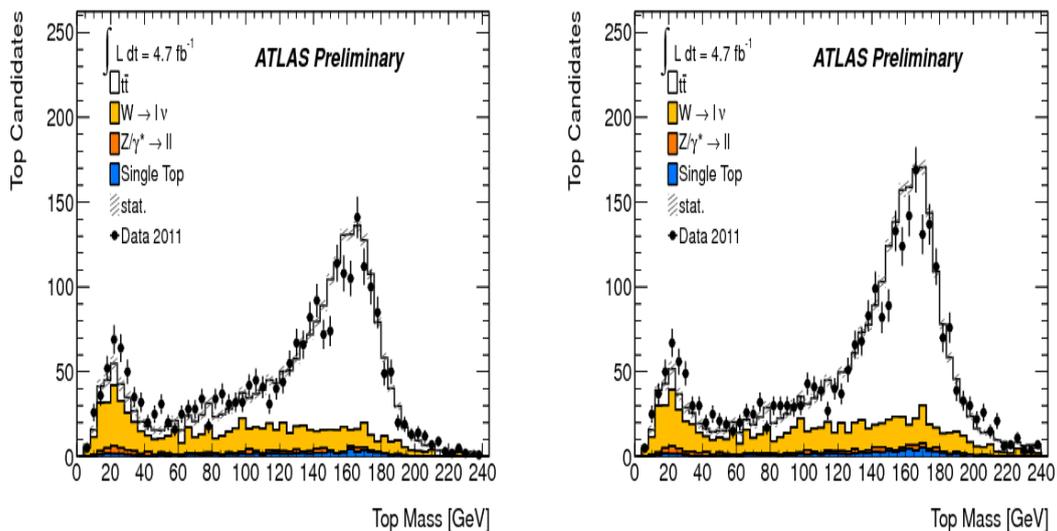
The next step is to validate the ATLAS event reconstruction simulation with actual LHC data. The data set used corresponds to 4.7 fb^{-1} collected until December 2011 at the LHC. The selection criteria previously mentioned are applied to the data and then the HEPTopTagger is applied with the tight, default and loose configurations. A comparison of the tight, default and loose configurations with data is shown on Figure 8.6. For these comparisons, other backgrounds (single top, and $Z/\gamma^* \rightarrow ll$) are included. These backgrounds represent only a small part of the total background. This confirms that is safe to ignore them for optimization purposes. From Figure 8.6, we observe a very good agreement with for all parameter configurations of the HEPTopTagger. In the upper histograms, a peak at low reconstructed top masses is observed. This is due to combinatorial effects passing the HEPTopTagger selection criteria. The peak comes mostly from the W+Jets channel and is very well modeled by Monte Carlo.

Yet another way to test this agreement is the following: One can compare how much each of the parameter configuration agrees with data by computing the following ratio:

$$r_{agree} = \frac{Data - (Signal + Background)}{Data} \quad (8.1)$$

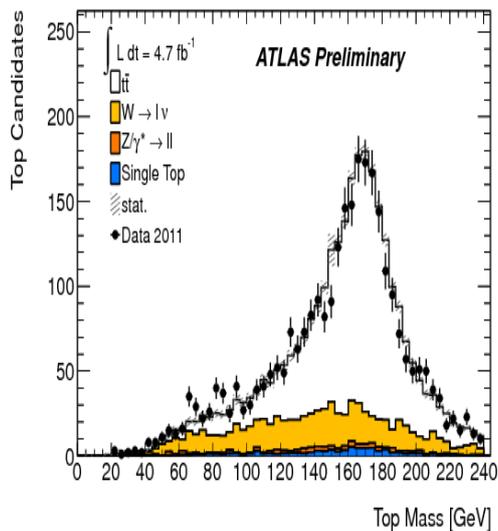
The closer r_{agree} is to zero the better the agreement for a certain parameter configurations. Figure 8.7 shows the 300 parameter configurations as black dots. On the x-axis

the number of tops found on the $t\bar{t}$ MC sample and on the y -axis the ratio r_{agree} . To improve the agreement of this ratio, Z+jets and single top are included in addition to W+jets. This ratio is computed for the 300 parameter configurations and shown in 8.6.



a) Default Settings with $R_{fatjet} = 1.5$

b) Loose Settings with $R_{fatjet} = 1.5$



c) Loose Settings with $R_{fatjet} = 1.5$

Figure 8.6: Comparison of data and Monte Carlo shown for the tight, default and loose configurations. All three HEPTopTagger parameter settings show a good agreement to data [48].

As seen from Figure 8.2, signal efficiency is a measure of the "looseness" of the HEP-TopTagger i.e. as the HEP-TopTagger parameters move towards looser configurations, the signal efficiency increases. An agreement for all parameter settings (even looser ones) is observed to be always within 5%. This is specially noteworthy considering no systematic uncertainties are considered for this optimization and the uncertainty on the $t\bar{t}$ cross section is around 10%.

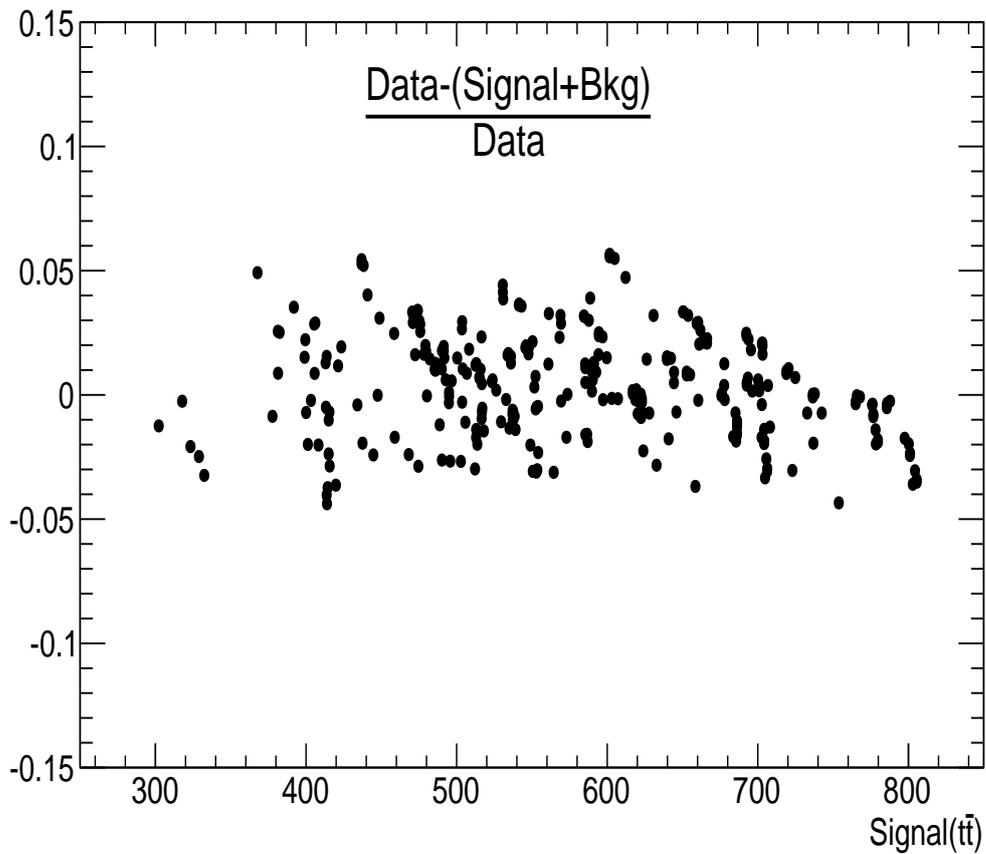


Figure 8.7: Comparison of data and Monte Carlo shown for each of the 300 parameter configurations. All the parameter configurations simulate data within a 5% error.

8.4 Low mass peak investigation

After verifying that the top mass peak is well described by the simulation, we proceeded to investigate the rest of the top mass distribution. In particular, the peak that is located in the low mass spectrum. This peak appears when $m_{cut} = 30$ GeV but not when $m_{cut} = 70$ GeV. Histograms with different mass cut value starting from $m_{cut} = 30$ GeV are shown in Figures 8.8 - 8.10 . The peak comes from combinatorics, when by accident m_{23}/m_{123} is the m_W/m_{top} ratio but at a much lower scale. This occurs whenever the unclustering goes too far and underlying event and pile up for a combination that by change passes all the HEPTopTagger steps. Increasing the m_{cut} parameter, reduces the "amount" because the subjects reach a hardness criterion faster, thereby avoiding unclustering too far.

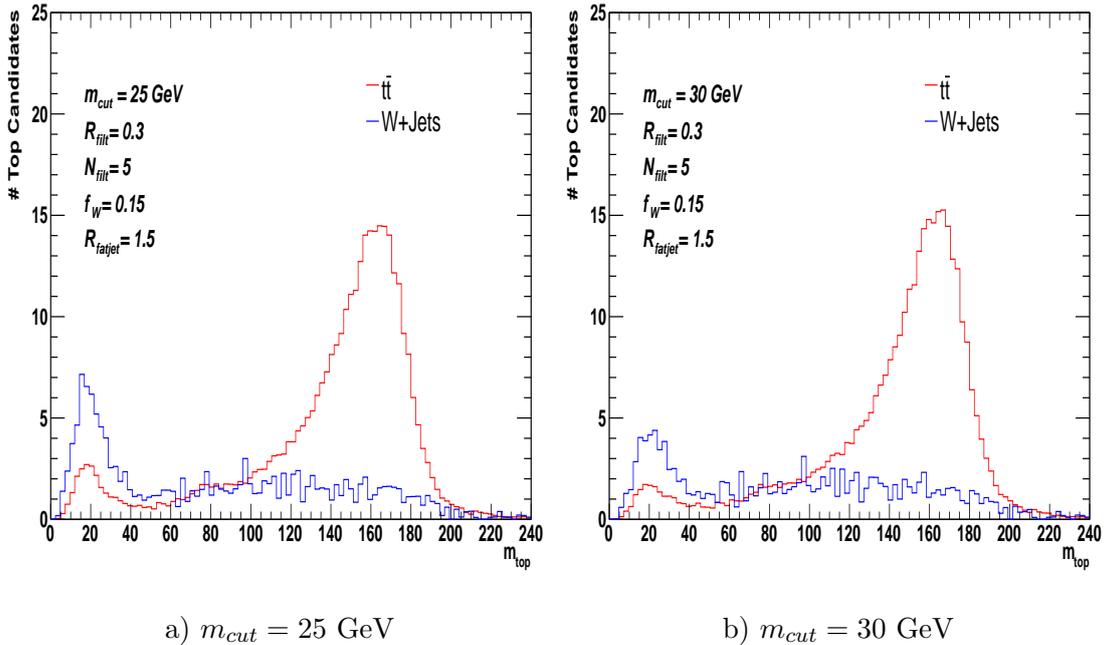
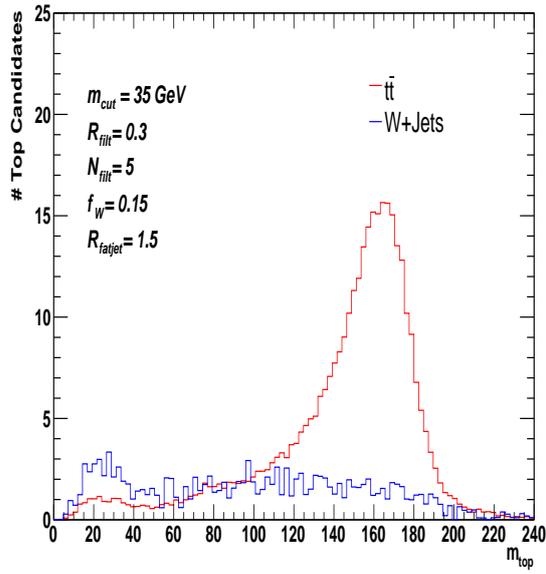
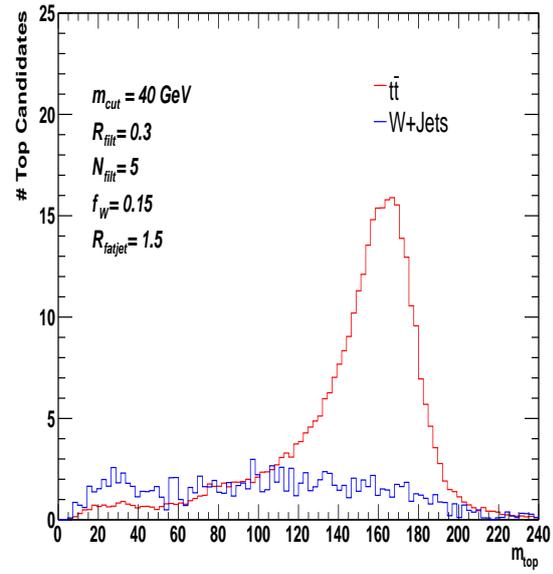


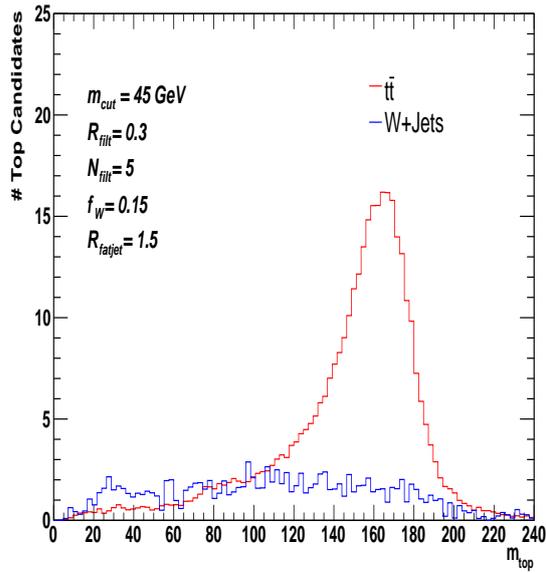
Figure 8.8: Signal and background mass distributions shown with different m_{cut} parameters. As the variable m_{cut} is increased the peak at low masses fades.



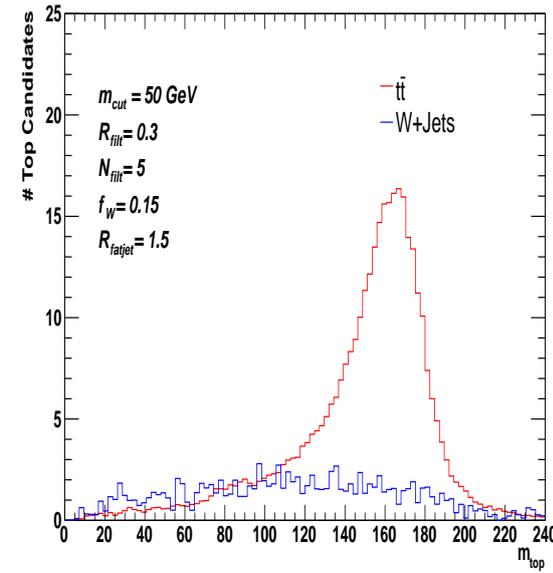
e) $m_{cut} = 35$ GeV



f) $m_{cut} = 40$ GeV



g) $m_{cut} = 45$ GeV



h) $m_{cut} = 50$ GeV

Figure 8.9: Signal and background mass distributions shown with different m_{cut} parameters. As the variable m_{cut} is increased the peak at low masses fades.

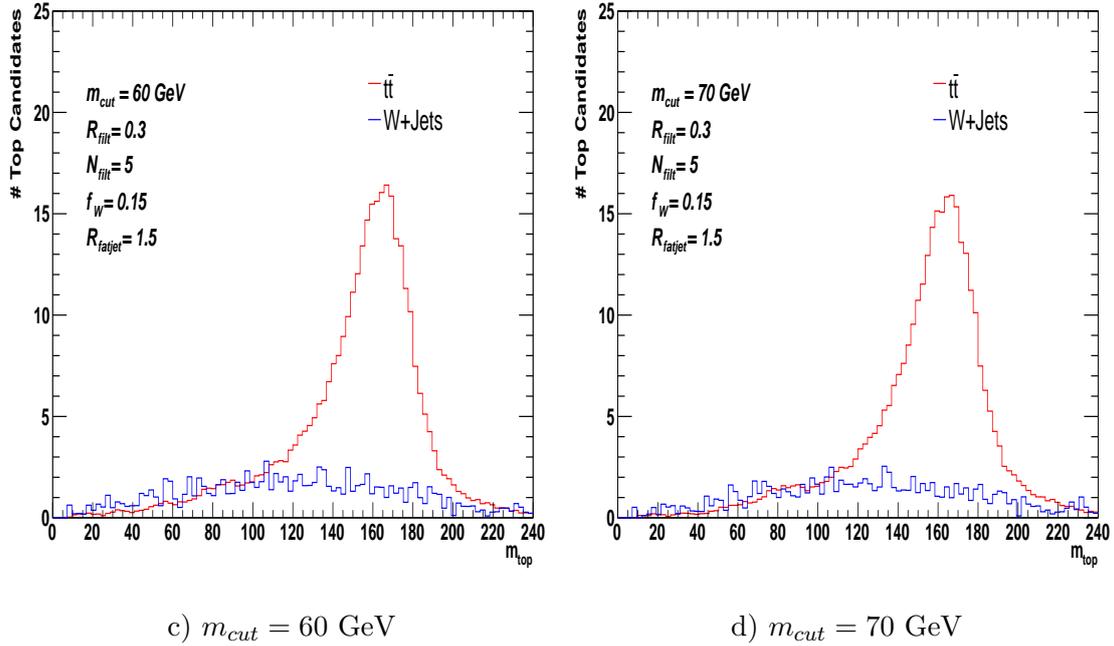


Figure 8.10: Signal and background mass distributions shown with different m_{cut} parameters. As the variable m_{cut} is increased the peak at low masses fades.

8.5 Dynamic Fat Jet Distance Parameter

The idea of increasing the fat jet radius was discussed previously yielding no obvious improvements. The idea behind a dynamic fat jet distance parameter is to adapt the fat jet distance parameter to the top quark transverse momentum. A first step into understanding how this would be work is to investigate how efficient different fat jet distance parameters would be in different p_T regions. Figure 8.11 shows the total efficiency of tagging a top quark in a fat jet as a function of the generated top transverse momentum for different fat jet distance parameters used in the first step of the HEPTopTagger. R1-R9 represents the tested fat jet distance parameter in increasing order from R=0.9 to R=1.7. A possible application of the dynamic fat jet distance parameter would be to use a bigger one for low p_T 's to capture more tops and then make it smaller at higher p_T 's to reduce background while maintaining the same signal efficiency.

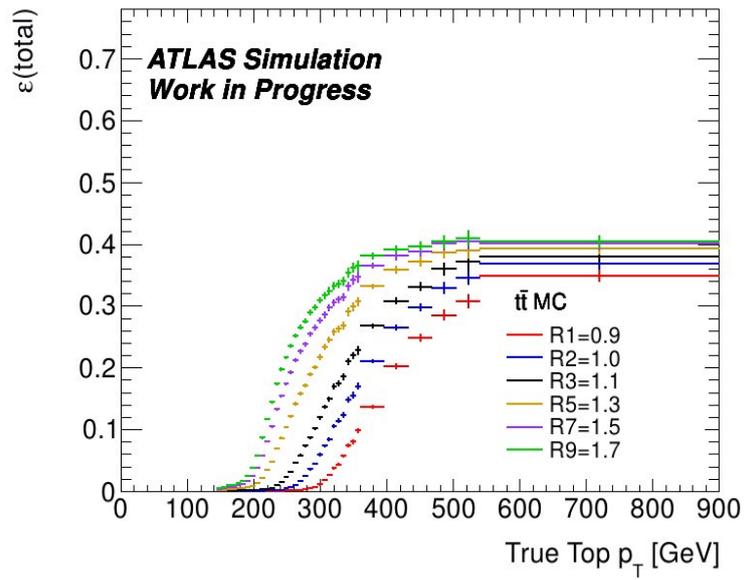


Figure 8.11: Top tagging efficiency as a function of generated top p_T for different fat jet distance parameters.

9 Multivariate Analysis

9.1 Introduction

The goal of a multivariate (MV) analysis is to extract the most possible information from a given set of data. A MV analysis takes place when performing a statistical analysis on more than one statistical variable at the time. Multivariate classification based on machine learning techniques has become a fundamental part in many physics analysis. It consists on identifying to which set of categories a particular events belongs i.e. does this event coming from the signal or the background? This is done by training a set of data containing observations whose membership is known. For the HEPTopTagger, this is a natural step for the optimization because the tagger contain several useful variables that can aid in classification.

9.2 The Toolkit for Multivariate Analysis Package

The Toolkit for Multivariate Analysis (TMVA) provides a ROOT-integrated environment for the processing, parallel evaluation and application of MV classification and regression [40]. All MV techniques in TMVA belong to the family of learning algorithms. Learning algorithms make use of training events, for which the desired output is known, to determine the mapping function that either describes a decision boundary (classification) or an approximation of the underlying functional behavior of the target value (regression). The TMVA classification analysis consists of two different phases: training and application. During the training phase the MV methods are trained, tested and evaluated. During the application phase, the chosen methods are applied to the concrete classification problem they have been trained for. All results shown in this thesis have been produced with TMVA version 3.8.14 and ROOT version 5.14.00. Training/testing and analyzing are separated into two different steps in TMVA. A set of signal and background event candidates, for example Monte Carlo samples are used. The separation of signal-like and background-like events is optimized with respect to a given set of vari-

ables. Secondly, the data are analyzed. The separation rules obtained by the training of the classifiers are applied to each candidate of the data sample. Its probability of being signal or background is returned and can be processed for further treatment.

9.3 Multivariate Methods

9.3.1 Simulated Annealing Optimized Cuts

The goal of a simulated annealing algorithm is to find a point in the space at which a real valued energy function is minimized. Simulated Annealing also aims to a minimization problem with several discrete or continuous, local or global minima. The algorithm is inspired by the process of of annealing which occur in condensed matter physics. When first heating and then slowly cooling down a metal its atoms move towards a state of lowest energy, while for sudden cooling the atoms tend to freeze in intermediate states higher energy.

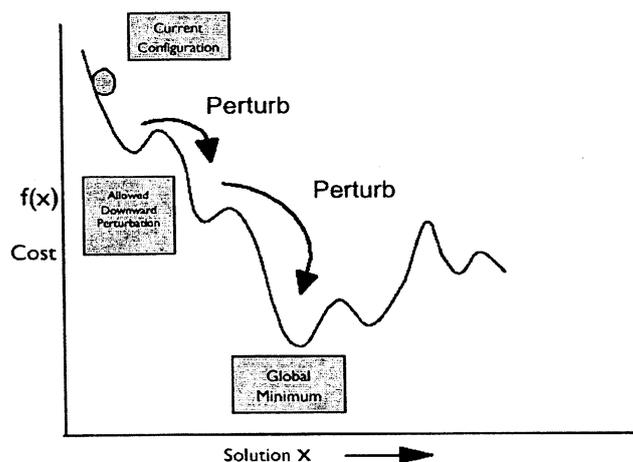


Figure 9.1: Simulated annealing allows escape from local minima with a given probability in order to reach a global minimum.

For infinitesimal annealing activity the system will always converge in its global energy minimum. This physical principle can be converted into an algorithm to achieve

slow, but correct convergence of an optimization problem with multiple solutions. Annealing is a when a material (steel or glass) is heated and then slowly cooled for softening and making the material less brittle. Simulated annealing, therefore, exposes a solution to "heat" and cools it slowly producing a more optimal solution. The probability to escape this local minimum is dependent on the initial temperature and the minimum temperature as given by

$$p(\Delta E) \propto \exp\left(-\frac{\Delta E}{T}\right) \quad (9.1)$$

where: ΔE is the difference between energy levels and T is the initial temperature.

The assumption behind simulated annealing is that the system will accept a worsening of the balance so that the cooling finds a better optimum. The probability of such perturbations to occur decreases with the size of a positive energy coefficient of the perturbation, and increases with the ambient temperature (T).

9.3.2 Artificial Neural Networks

An Artificial Neural Network (ANN) is a mathematical model inspired by biological neural networks. It consists in an group of artificial neurons interconnected in a system capable of solving problems linear computing is not able to. An ANN consists of connected neurons where each connection has its own weight. A number of n input variables leads to n^2 possible neurons. Once an input is presented to the and a corresponding desired or target response it set at the output, an error is composed from the difference of the desired response The error information is fed back to the system which makes all adjustments to their parameters in a systematic fashion (commonly known as the learning rule). This process is repeated until the desired output is acceptable. To speed up the processing, a reduced layout can be used as well, the so-called multilayer perceptron (MLP). In the case of the multilayer perceptron, the complexity of the network is reduced by ordering the neurons into multi-layers and allowing only one-directional connections between intermediate layers. All artificial neural networks implemented in TMVA are based on this method. The HEPTopTagger variables are

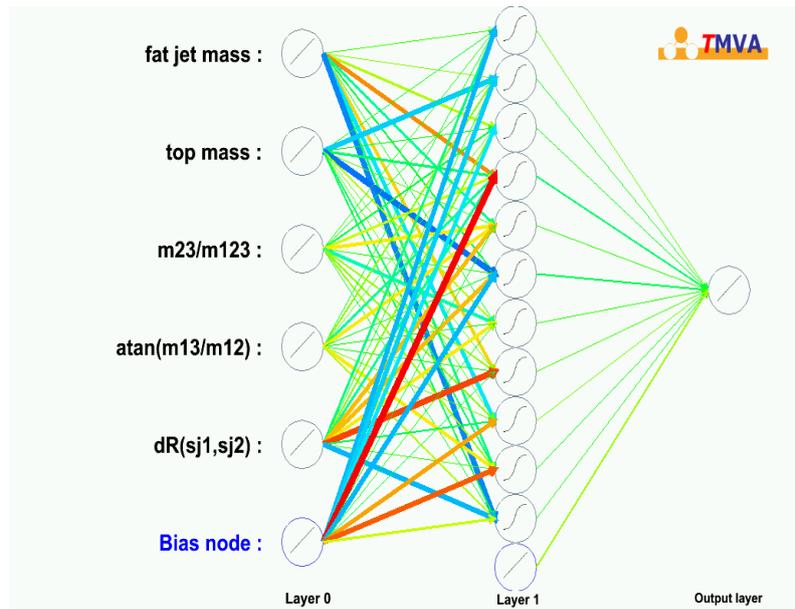


Figure 9.2: Diagram explaining the workings of a neural network with HEPTopTagger variables.

9.3.3 Boosted Decision Trees

The goal of a decision tree is to create a model that predicts the value of a target variable based on several input variables. Each interior node corresponds to one of the input variables. There are edges to children for each of the possible values of that input variable. Left and right decisions are taken on one single variable at a time until a stop criterion is fulfilled (see Figure 9.3). Decision nodes specify a predicate condition. Prediction nodes contain a single number. A decision tree is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility.

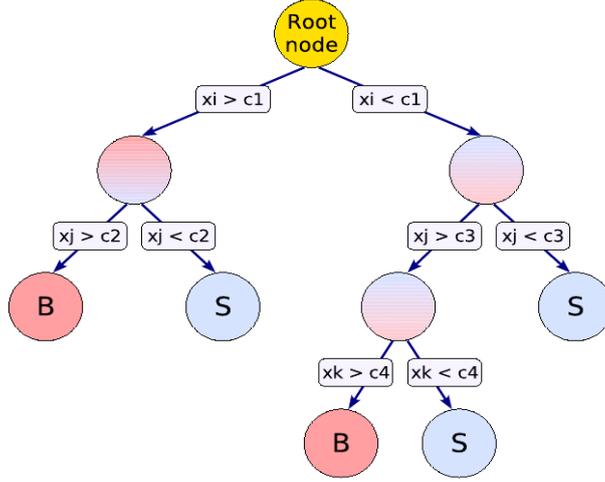


Figure 9.3: Diagram explaining the principle of a boosted decision tree.

9.4 Multivariate Analysis

9.4.1 Multivariate Analysis Preparation

MC samples: Dedicated ntuples were prepared for the MV analysis. Samples of Z' bosons of 1000, 1600 and 3000 GeV (MC@NLO) decaying to $t\bar{t}$ are used as the signal, while for background W +jets (ALPGEN) is used. A standard $t\bar{t}$ sample could have been used as well, but its falling spectrum in p_T does not allow to study large p_T bins. The Z' samples provide a flat spectrum in several regions of p_T up to 1600 GeV. However because statistics on the background are limited for high p_T , the MV optimization is undertaken in in the 200-450 GeV fat jet p_T range. See Figure 9.4.

Fat jet matching: A matching requiring that the fat jet is within $\Delta R = 1.0$ of the generated top quark. This is done to ensure that the fat jet really does come from a hadronically decaying top and not other mis-identified objects.

Open substructure variables: Another requirement for the HEPTopTagger MV analysis was to loosen the substructure requirements from Equations 6.4 to 6.6. This is done in order to avoid a pre-filtering that would not leave much room for the multivariate analysis to optimize. An event is defined as a matched fat jet in the event. It is possible

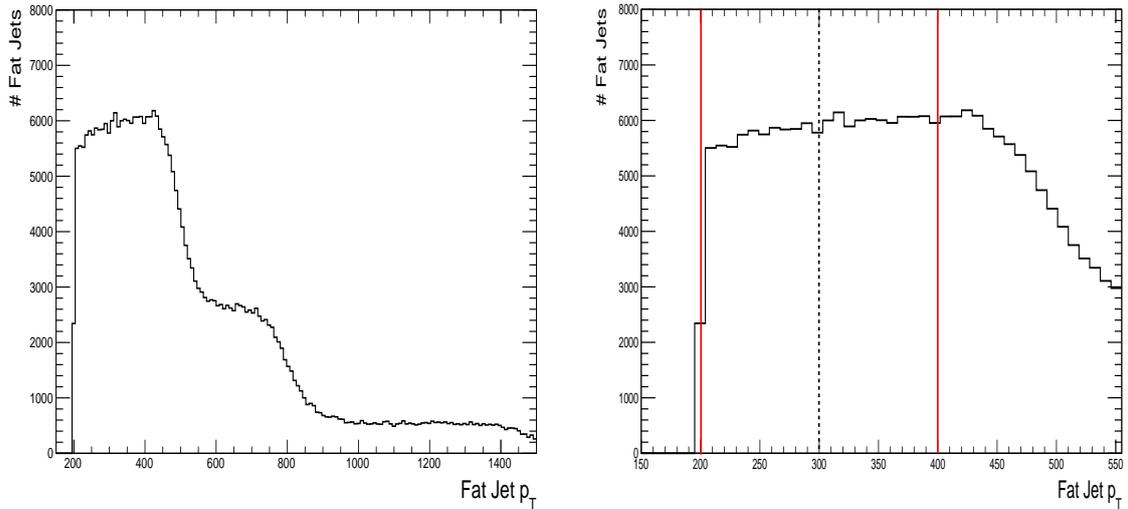


Figure 9.4: Left: p_T spectrum of the three available Z' samples added. The distribution falls at 500 GeV, 800 GeV and 1600 GeV respectively as expected for the Z' samples included. Right: The spectrum with zoom in the studied region, 200-450 GeV. The red lines represent the cut on the studied region while. The dotted line divides the two p_T bins studied.

to have two fat jets coming from the same Z' . These two (matched) fat jets will count as 2 events. Because of the Z' fat jet p_T spectrum, the analysis is performed only in two fat jet p_T bins. To compare the results of the MV analysis, the efficiencies and background rejections are computed for the tight, default, and loose settings as given by

$$\epsilon_s = \text{signal efficiency} = \frac{\# \text{ top candidates}}{\# \text{ matched fat jets}} \quad (9.2)$$

$$r_{bkg} = \text{background rejection} = 1 - \frac{\# \text{ top candidates}}{\# \text{ fat jets}} \quad (9.3)$$

where a top candidate corresponds to a top reconstructed with the HEPTopTagger default parameter settings with a mass in mass window of 140-210 GeV. This is done for two bins of p_T , 200-300 GeV and 300-450 GeV as illustrated by Figure 9.4. The top mass distributions (obtained with the default settings) and fat jet p_T distributions are shown for signal and background in Figure 9.5 and 9.6 respectively for both fat jet p_T bins. The

number ϵ_s quoted on the top mass distributions represents the number of events between the red lines (mass window selection criterion) divided by the number of events in the fat jet p_T distribution in the corresponding range for the signal. The number r_{bkg} is 1 minus the number of events between the red lines (mass window selection criterion) divided by the number of events in the fat jet p_T distribution in the corresponding range for the background. The equivalent histograms for tight and loose settings are not shown. However, their efficiencies are still calculated. They are summarized in Table 9.1 and Table 9.2.

	ϵ_s and r_{bkg} for default settings
tight	$\epsilon_s = 0.04$, $r_{bkg} = 0.999$
default	$\epsilon_s = 0.08$, $r_{bkg} = 0.997$
loose	$\epsilon_s = 0.12$, $r_{bkg} = 0.990$

Table 9.1: Signal efficiency and background rejection in the range **200-300 GeV** in fat jet p_T calculated for the tight, default and loose HEPTopTagger parameter configurations.

	ϵ_s and r_{bkg} for default settings
tight	$\epsilon_s = 0.17$, $r_{bkg} = 0.992$
default	$\epsilon_s = 0.27$, $r_{bkg} = 0.985$
loose	$\epsilon_s = 0.31$, $r_{bkg} = 0.975$

Table 9.2: Signal efficiency and background rejection in the range **300-450 GeV** in fat jet p_T calculated for the tight, default and loose HEPTopTagger parameter configurations.

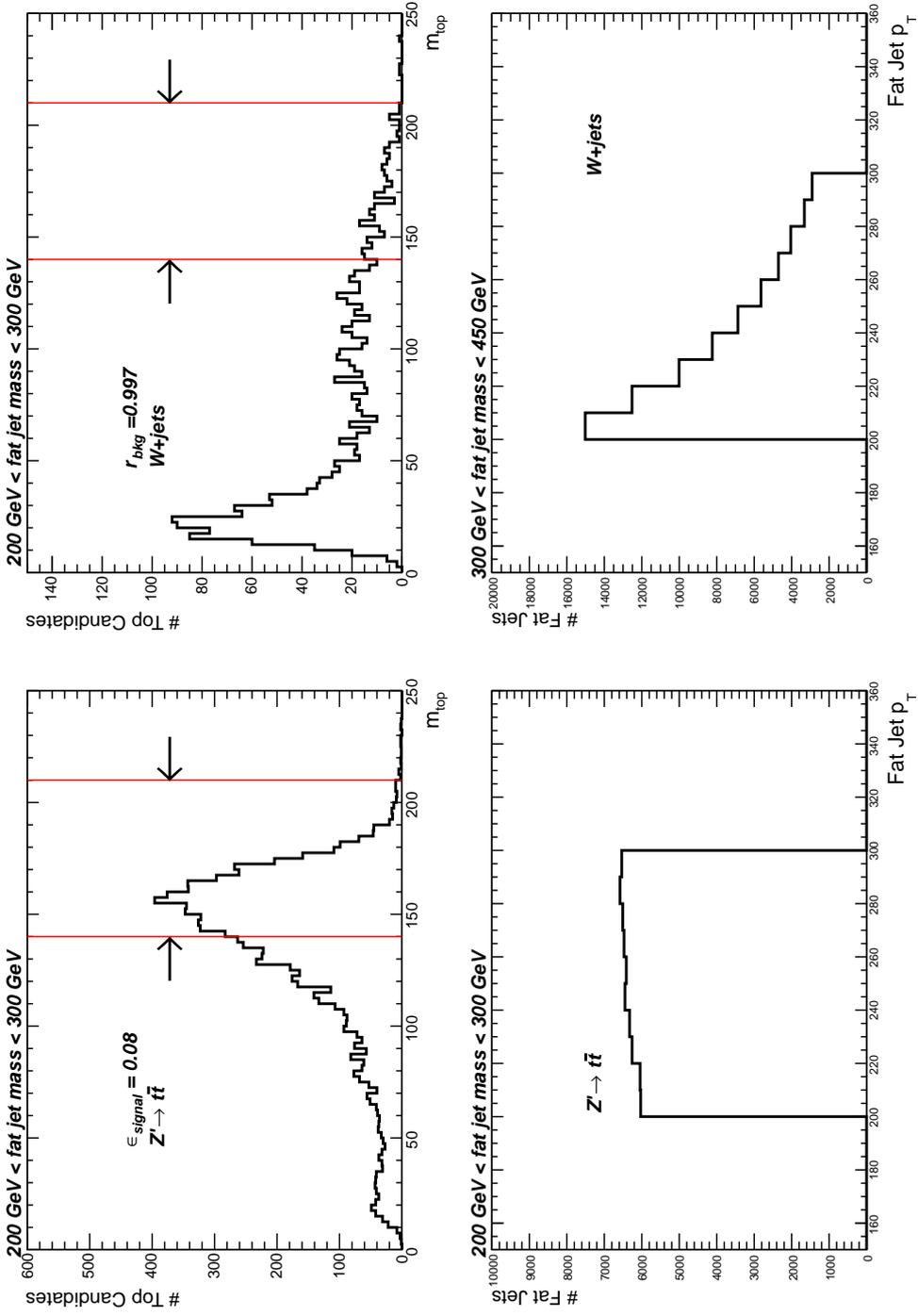


Figure 9.5: Top candidate mass and fat jet p_T distributions for $Z' \rightarrow t\bar{t}$ (left column) and W+jets (right column). A mass window selection criterion from 140-210 GeV is required as indicated by the red lines. The signal efficiency and background rejection are calculated for a fat jet p_T range of **200-300 GeV** . .

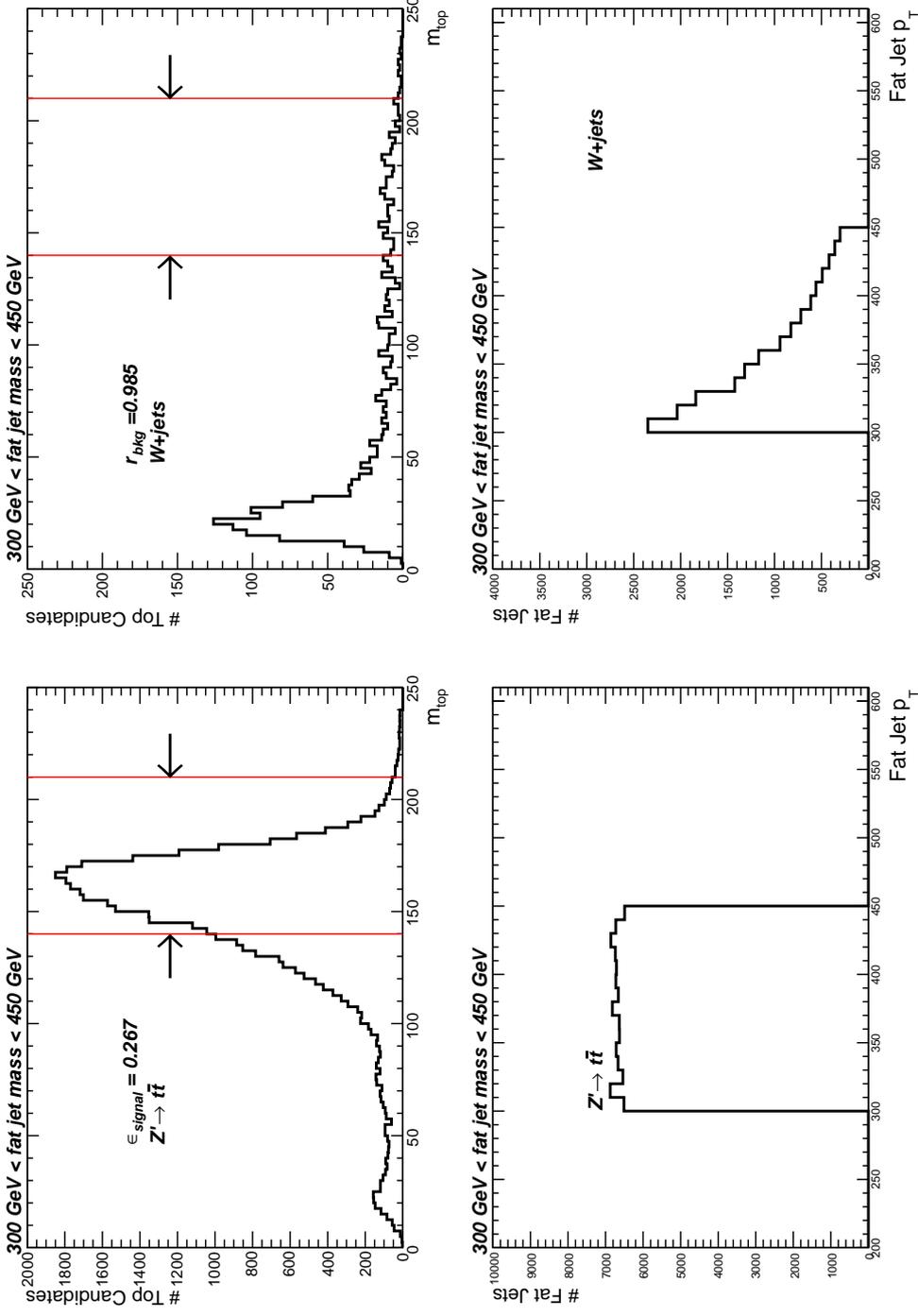


Figure 9.6: Top candidate mass and fat jet p_T distributions for $Z' \rightarrow t\bar{t}$ (left column) and $W+jets$ (right column). A mass window selection criterion from 140-210 GeV is required as indicated by the red lines. The signal efficiency and background rejection are calculated for fat jet p_T range of **300-450 GeV**.

In addition to the top candidate mass and the substructure variables (m_{23}/m_{123} and $\text{atan}(m_{13}/m_{12})$), new variables are included for the MV analysis, to try to improve the relevant figures of merit. The total set of variables available for the MV analysis is

- top candidate mass
- m_{23}/m_{123}
- $\text{atan}(m_{13}/m_{12})$
- fat jet mass
- $\Delta R_{fj,sj1}, \Delta R_{fj,sj2}, \Delta R_{fj,sj3}$
- $\Delta R_{sj1,sj2}, \Delta R_{sj2,sj3}, \Delta R_{sj1,sj3}$

where $\Delta R_{sj_i,sj_j}$ is the angular distance between each subjet and $\Delta R_{fj,sj_i}$ is the distance from the subjet i to the fat jet. As mentioned in subsection 6.2, after the filtering step only these subjets remain. These subjets are ordered in p_T , sj_1 being the subjet with the highest p_T and sj_3 the subjet with the least. It is expected that they will give an additional handle to distinguish between signal and background. The fat jet mass is also expected to help improve the classification. All distributions for the possible discriminant variables are shown in Figure 9.7 and 9.8 as processed by the TMVA package. Their correlations are shown in Figure 9.9 and 9.10.

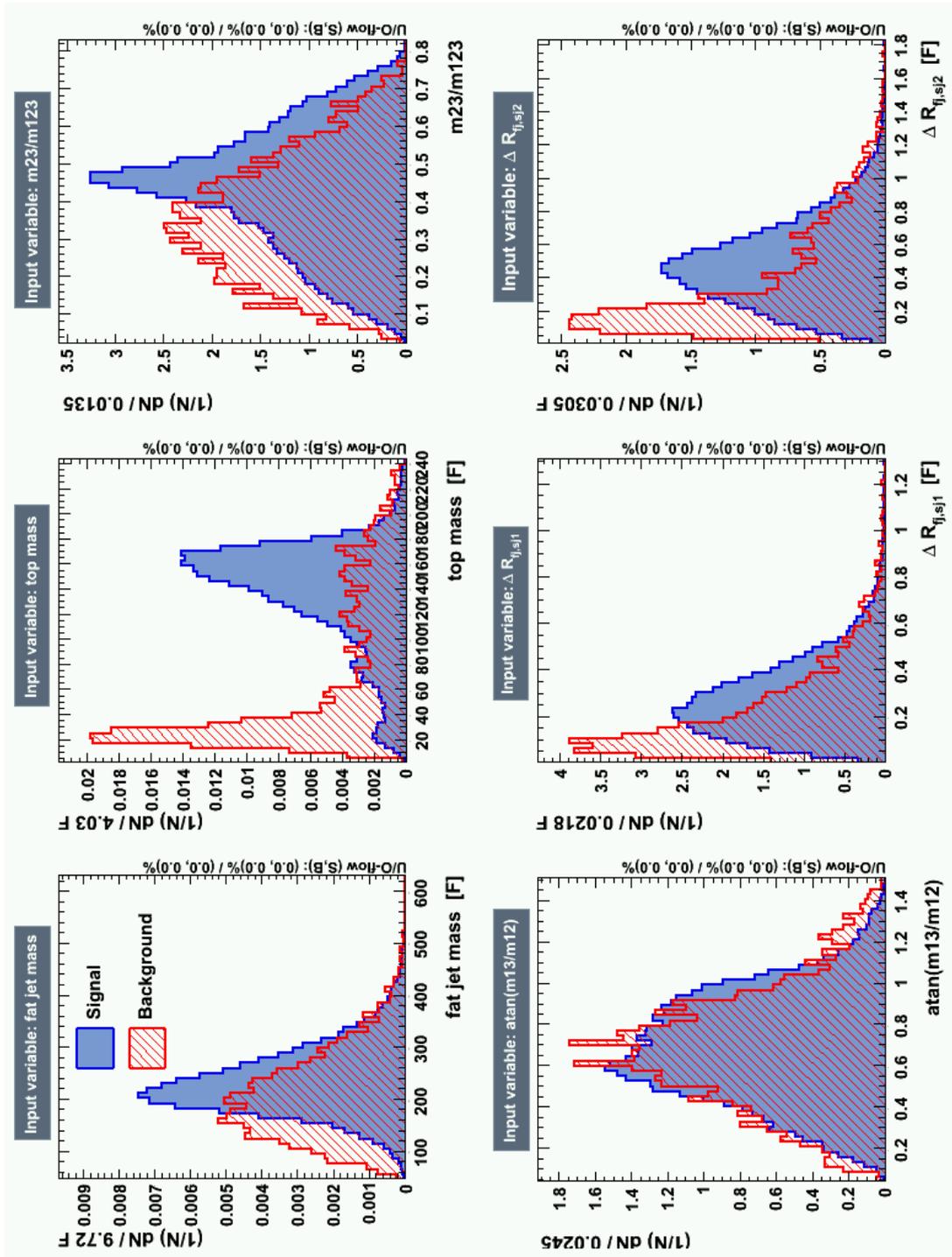


Figure 9.7: Possible input variables for the MV analysis as processed by the TMVA package (1).

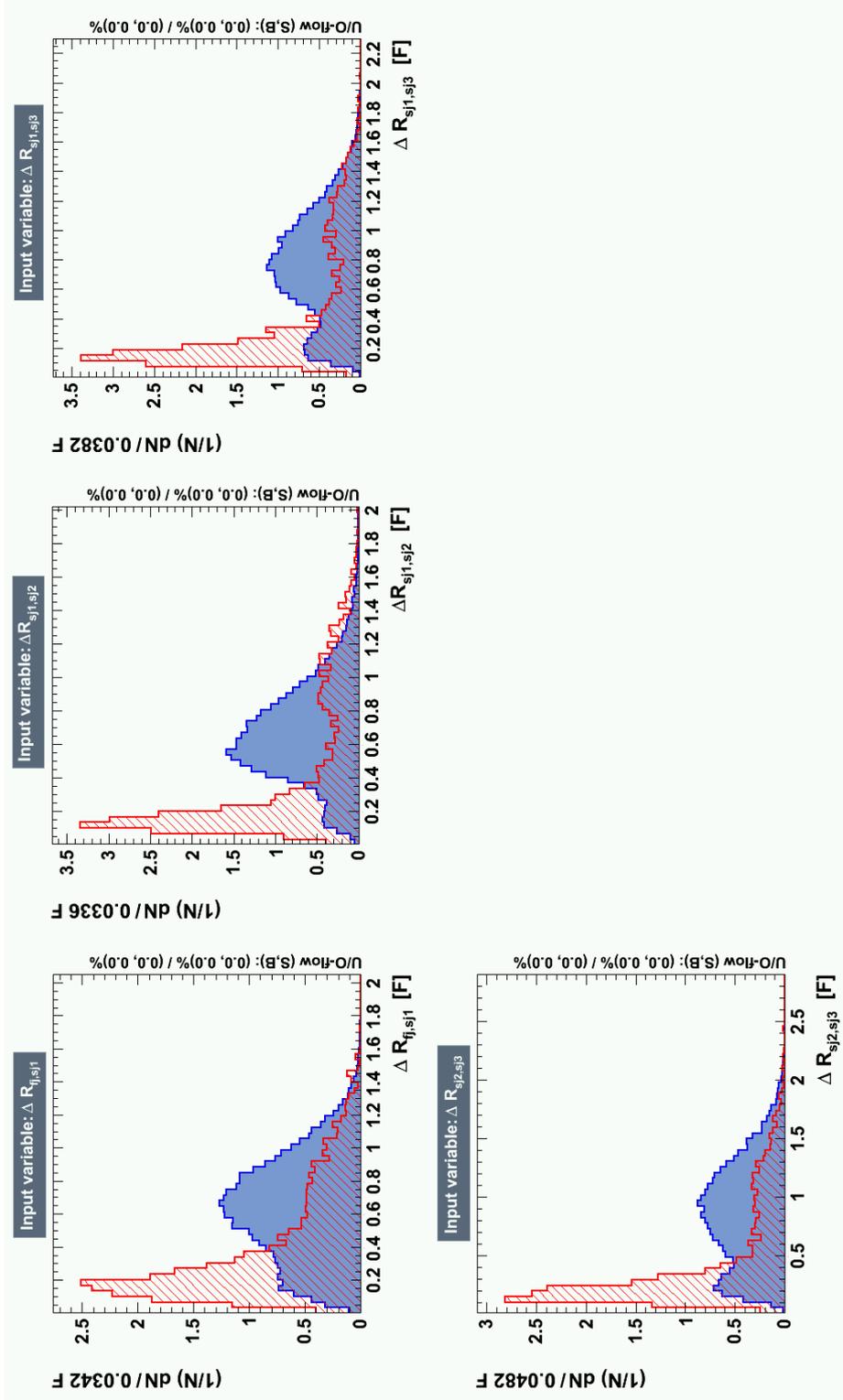


Figure 9.8: Possible input variables for the MV analysis as processed by the TMVA package (2).

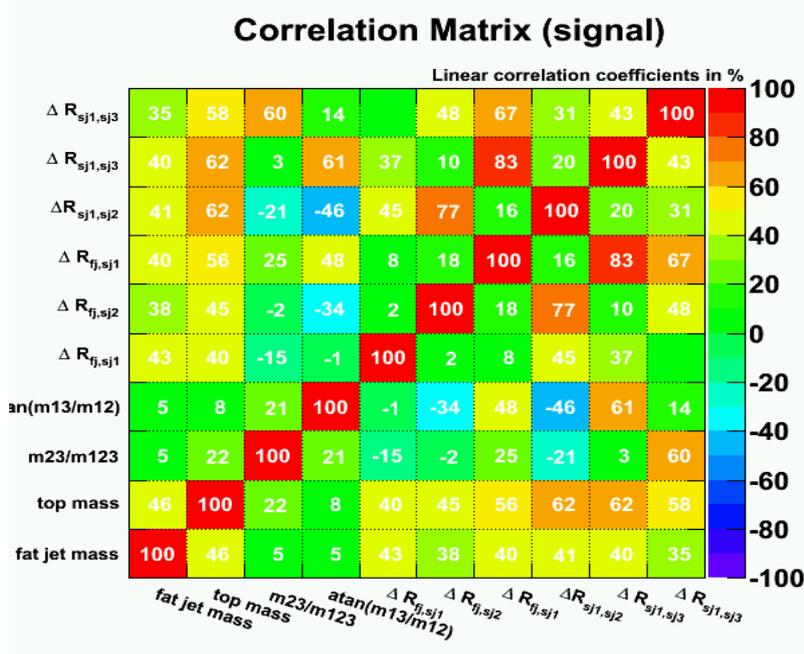


Figure 9.9: Correlations matrix of all possible discriminant variables in signal ($Z' \rightarrow t\bar{t}$).

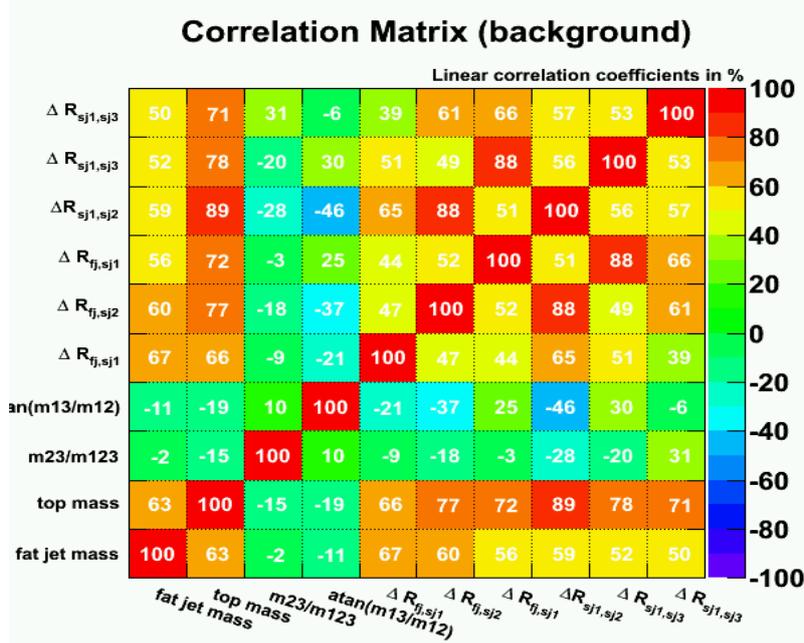


Figure 9.10: Correlations matrix of all possible discriminant variables in background ($W+\text{jets}$).

9.4.2 Multivariate Optimization

The MV methods used are as explained: simulated annealing optimized cuts (CutsSA), artificial neural network (MLP) and boosted decision trees (BDT). Within the TMVA framework, each of these MV methods contains tunable parameter to improve performance. For the CutsSA method, these are the minimum and initial temperatures as explained in subsection 9.3.1. These parameters were modified and compared within the same MV methods. The ones returning the best performance in the correspondent p_T bin were chosen to be presented here. For the MLP, the process was similar. Different number of hidden layers and cycles was tested, and in the end the best performing setting were used. Many MLP's with different settings were tested but none of them could return a performance better than the BDT. For the BDT no parameters were modified since it always returned a good performance and it is known to work very well out-of-the-box.

The MV analysis was carried out incrementally, starting with 4 variables, then moving to 5 and 7 variables. Using more than 7 variables did not yield any increase in performance. The background rejection vs. signal efficiency figures are shown for 4,5 and 7 variables in the fat jet p_T ranges of 200 – 300 GeV and 300 – 450 GeV, respectively in Figures 8.11-8.16. The color points in each figure correspond to the efficiencies calculated in Table 9.1 and Table 9.2. They allow for a visual measure on how much it is possible to improve performance using a MV analysis.

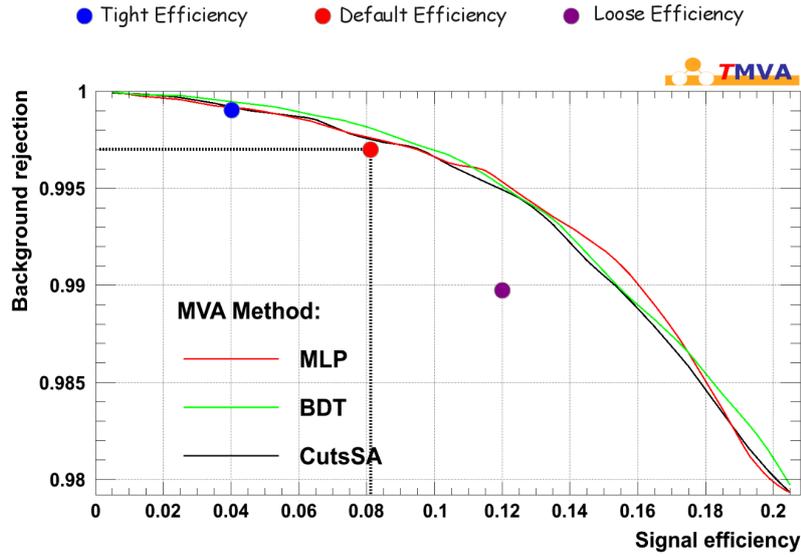


Figure 9.11: Background rejection vs. signal efficiency in the range **200-300 GeV** in fat jet p_T using **4 variables**: top candidate mass, m_{23}/m_{123} , $\text{atan}(m_{13}/m_{12})$ and $\Delta R_{sj1,sj2}$.

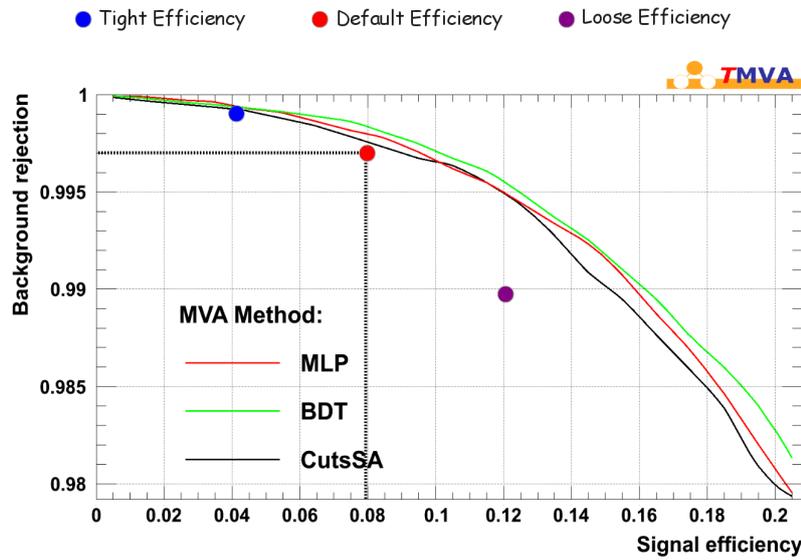


Figure 9.12: Background rejection vs. signal efficiency in the range **200-300 GeV** in fat jet p_T using **5 variables**: top candidate mass, m_{23}/m_{123} , $\text{atan}(m_{13}/m_{12})$, $\Delta R_{sj1,sj2}$ and the fat jet mass.

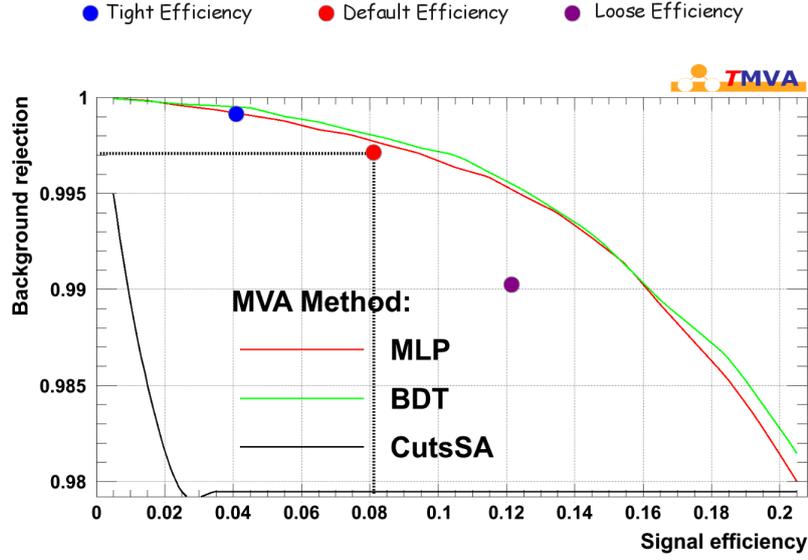


Figure 9.13: Background rejection vs. signal efficiency in the range **200-300 GeV** in fat jet p_T using **7 variables**: top candidate mass, m_{23}/m_{123} , $\text{atan}(m_{13}/m_{12})$, fat jet mass, $\Delta R_{sj1,sj2}$, $\Delta R_{sj1,sj3}$, and $\Delta R_{fj,sj3}$

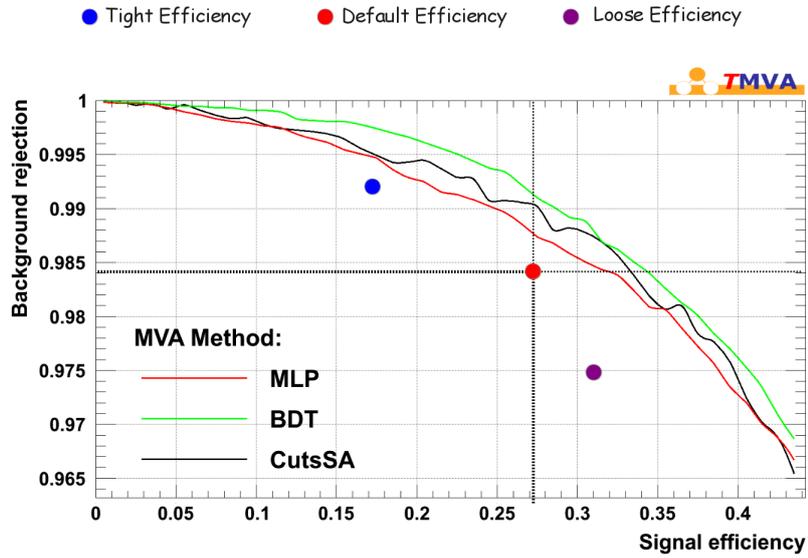


Figure 9.14: Background rejection vs. signal efficiency in the range **300-450 GeV** in fat jet p_T using **4 variables**: top candidate mass, m_{23}/m_{123} , $\text{atan}(m_{13}/m_{12})$ and $\Delta R_{sj1,sj2}$.

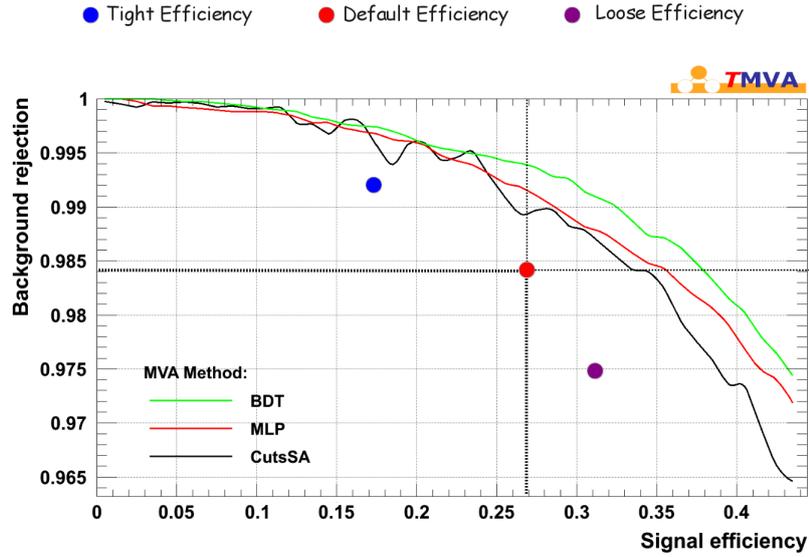


Figure 9.15: Background rejection vs. signal efficiency the range **300-450 GeV** in fat jet p_T using **5 variables**: top candidate mass, m_{23}/m_{123} , $\text{atan}(m_{13}/m_{12})$, $\Delta R_{sj1,sj2}$ and fat jet mass.

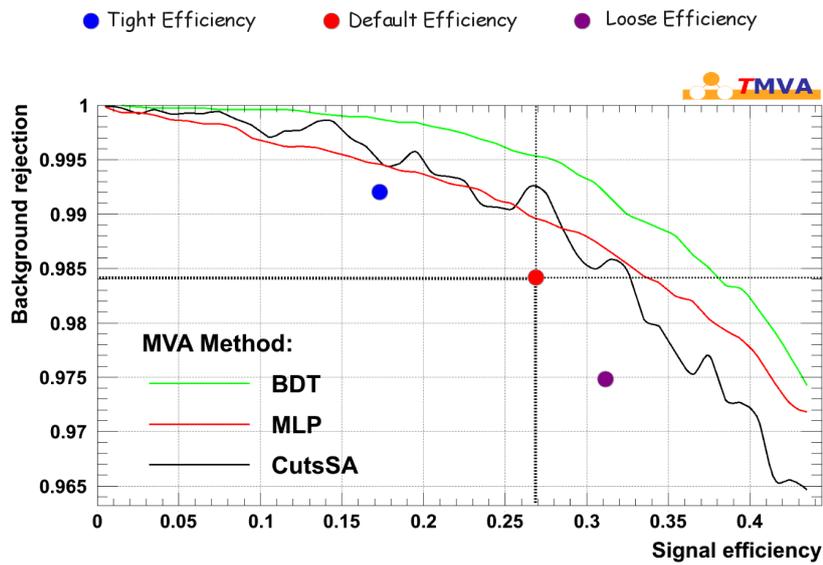


Figure 9.16: Background rejection vs. signal efficiency in the range **300-450 GeV** in fat jet p_T using **7 variables**: top candidate mass, m_{23}/m_{123} , $\text{atan}(m_{13}/m_{12})$, $\Delta R_{sj \text{ all-pairwise}}$ and fat jet mass.

CutsSA Performance: The CutsSA method offers a set of selection criteria on the variables given, for a determined signal efficiency or background rejection. For the range 200-300 GeV CutsSA, (as for the other methods) does not yield an improvement. The Cuts SA curves show "wigliness" in 300-450 GeV. This is because for this range the number of background events is limited. Additionally, because the efficiency and background are calculated using one hundred bins, the background-signal efficiency may not reflect exact efficiencies. Thus, values for the efficiency and the corresponding selection criteria are extracted from an output TMVA file. The CutsSA performance does not increase when the number of variables is increased from 5 to 7 and it hovers mostly around $\epsilon_s = 0.32$ for the same r_{bkg} . The efficiencies obtained with the CutsSA method are shown on Table 9.3. An improvement of 5% is with the optimized cut with the same background rejection. Or one can reduce the background ~ 40 staying with the same signal efficiency.

	Same ϵ_s	Same r_{bkg}
Default Settings	$\epsilon_s = 0.267, r_{bkg} = 0.984$	$\epsilon_s = 0.267, r_{bkg} = 0.984$
CutsSA	$\epsilon_s = 0.265, r_{bkg} = 0.990$	$\epsilon_s = 0.324, r_{bkg} = 0.984$

Table 9.3: Comparison of efficiencies from default settings and CutsSA for same ϵ_s and same r_{bkg} .

The selection criteria as obtained by CutsSA are shown in Table 9.4.

	CutsSA ($\epsilon_s = 0.267, r_{bkg} = 0.990$)	CutsSA ($r_{bkg} = 0.984, \epsilon_s = 0.324$)
top cand mass	124-205 GeV	118-187 GeV
m23/m123	> 0.26	> 0.39
atan(m23/m123)	> 0.52	> 0.33
$\Delta R_{sj_1, sj_2}$	< 0.86	No cut

Table 9.4: CutsSA selection criteria for different values of background rejection and signal efficiency.

The selection criteria for variables m23/m123 and atan(m23/m123) for default set-

tings, cannot be compared to the CutsSA selection criteria. The default settings of the HEPTopTagger have selection criteria that are "curved" on the m_{23}/m_{123} - $\arctan(m_{23}/m_{123})$ space as seen on Figure 9.17. However, only with the inclusion of one extra angular variable, CutsSA is able to stay competitive with the default settings.

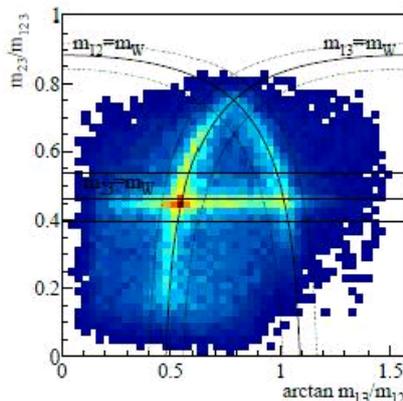


Figure 9.17: A triangular-shape cut is applied to the substructure variables obtained with the HEPTopTagger.

Figure 9.18 shows the selection criteria as obtained by CutsSA overlaid with their corresponding distributions. The black dots represent a minimum cut while the dark green dots represent a maximum cut. The distributions of the variables for signal and background in which the selection criteria are overlaid to observe where exactly the CutsSA optimization algorithm is making the the selection criteria. When a series a dots are aligned, for example in b) m_{23}/m_{123} it means that there is no maximum or minimum cut.

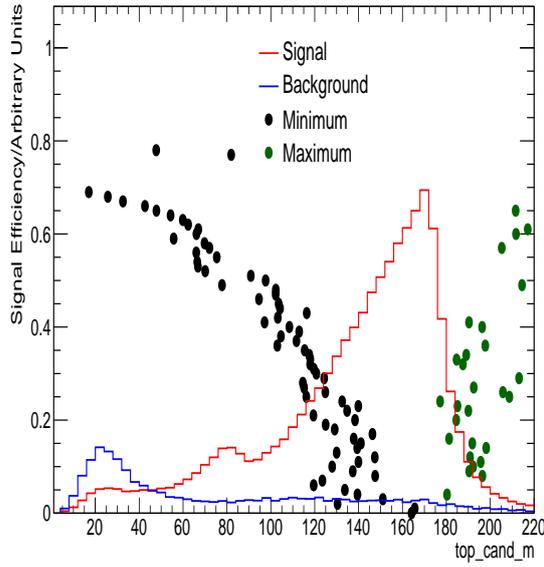
A summary of the best results found is given on Table 9.5 and Table 9.6 for the same signal efficiency and for the same background rejection as the default setting only for the 300-450 GeV range, respectively.

$\epsilon_s = 0.267$	Default	CutsSA	MLP	BDT
300-450 GeV	$r_{bkg} = 0.984$	0.990	0.998	0.995

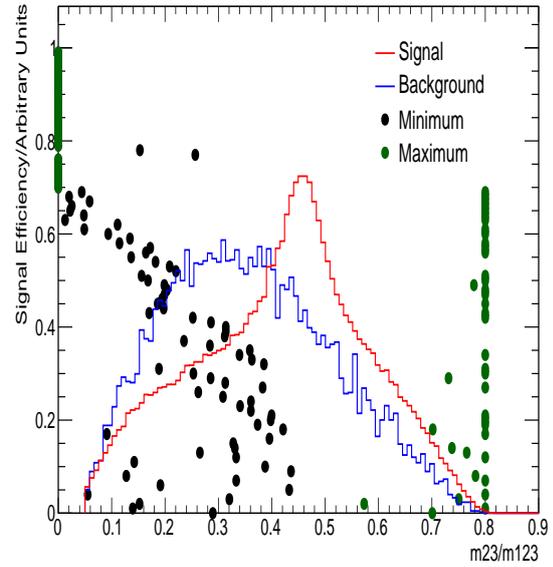
Table 9.5: TMVA optimization result for the same default signal efficiency (ϵ_s).

$r_{bkg} = 0.984$	Default	CutsSA	MLP	BDT
300-450 GeV	$\epsilon_s = 0.27$	0.32	0.35	0.38

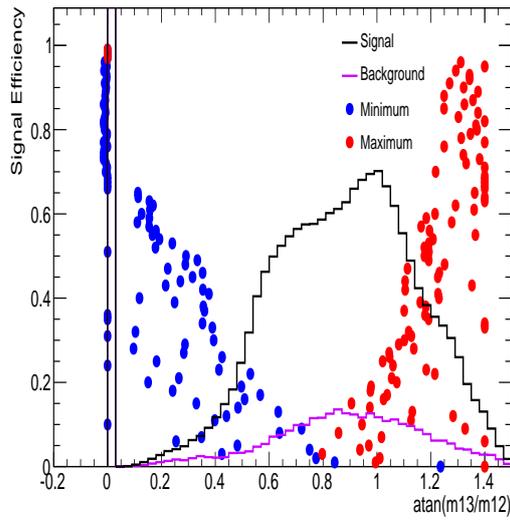
Table 9.6: TMVA optimization results for the same default background rejection (r_{bkg}).



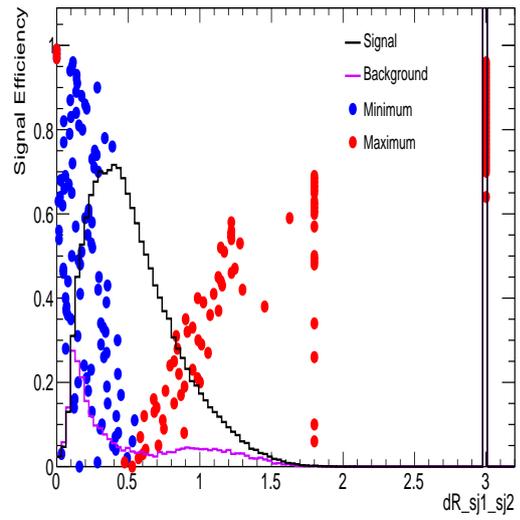
a) Top candidate mass



b) m_{23}/m_{123}



c) $\text{atan}(m_{13}/m_{12})$



d) $\Delta R_{s_{j1},s_{j2}}$

Figure 9.18: Selection criteria obtained by the CutsSA method for four variables. Black dots represent a minimum cut, while dark green dots represent a maximum cut for a given signal efficiency shown in the y-axis. The dots stop at 70% efficiency because this is where there are no more top quark to be found in the sample. The histograms are scaled to arbitrary units

10 Conclusion

For this thesis two set of results were presented: the first set of results dealt with the impact of different parameter settings on the performance of the HEPTopTagger. A parameter scan was performed for 300 possible HEPTopTagger parameter configurations. Figure 8.2 shows the relationship between signal and background. From this figure we learn that the figure of merit S/\sqrt{B} cannot be increased. However, there can be increase in S/B by varying parameters. Tight, medium and loose configurations are chosen in the different regions of the Figure 8.2. These configurations vary the value of S/B , with S/\sqrt{B} remaining constant. Also explored was the fat jet distance parameter, R_{fatjet} from 1.5 to 1.8 which showed that, as expected, signal efficiency is increased but there is a small drop on S/B . Therefore is not very useful to increase R_{fatjet} . All of these parameter variations of the HEPTopTagger performed with Monte Carlo simulations were validated with LHC data. This gives confidence that whatever the parameter configuration one decides to use for a particular analysis, the HEPTopTagger performance will be very well described by Monte Carlo simulations.

The second set of results presented in this thesis, was the performance optimization of the HEPTopTagger by means of a multivariate analysis. This is done by feeding possible discriminating variables to the TMVA package. With this analysis is found that it is possible to increase performance (i.e. signal efficiency, background rejection) over a traditional cut-based approach. With the optimized cuts (CutsSA) found by TMVA, it is possible to reduce the background by 50% while maintaining the same signal efficiency for default HEPTopTagger parameter settings. The boosted decision tree method of multivariate classification is able to increase the signal efficiency from 27% to 38% (absolute) while maintaining the same background rejection. With the multivariate analysis results we found there is still some room for improvement on the HEPTopTagger performance over a traditional cut-based analysis. As new analyses using the HEPTopTagger come out, the possibility to improve performance by means of a multivariate analysis on the HEPTopTagger variables is established.

References

- [1] Sven Olaf Moch. *Top quark production at the Tevatron and the LHC*. Fermilab Seminar. <http://theory.fnal.gov/seminars/slides/2009/SMoch.pdf> (2009)
- [2] Tilman Plehn and Michael Spannowsky. *Top Tagging*. [arXiv:1112.4441] [hep-ph] (2011)
- [3] Lyndon Evans and Philip Bryant. *LHC Machine*. JINST. 3(S08001) (2008)
- [4] Jean-Luc Caron. Original ref: E540. *LHC Project Illustrations / Civil engineering* LHC-PHO-1998-349 (1998)
- [5] K. Aamodt *et. al.*. *The ALICE Experiment at the CERN LHC*. JINST, 3(S08002) (2008)
- [6] ATLAS Collaboration. *The ATLAS Experiment at the CERN LHC*. JINST, 3(S08002) (2008)
- [7] R. Adolphi *et. al.*. *The CMS Experiment at the CERN LHC*. JINST, 3(S08002) (2008)
- [8] A. Alves *et. al.*. *The LHCb Detector at the CERN LHC*. JINST, 3(S08002) (2008)
- [9] CERN PhotoLab / Physics diagrams and charts. *The CERN accelerator complex* CERN-DI-0812015 Photograph: Christiane Lefèvre. <http://cdsweb.cern.ch/record/1260465> (2008)
- [10] ATLAS Collaboration. *Luminosity Determination in pp Collision at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC*. Eur. Phys. J. C71(1630) (2011)
- [11] ATLAS Collaboration. *Updated Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS detector*. (ATLAS-CONF-2011-011) (2011)

- [12] The ATLAS Collaboration. *Luminosity Determination in pp Collisions at $\sqrt{s}=7$ TeV Using the ATLAS Detector at the LHC* (arXiv:1101.2185) <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults> (2011)
- [13] Egede U. *The search for a standard model Higgs at the LHC and electron identification using transition radiation in the ATLAS tracker* (Doctoral Dissertation) (1997) Figure used in page page 9 retrieved from <http://www.hep.lu.se/atlas/thesis/egede/thesis-node39.html>.
- [14] Gabriel *et. al.* NIM A927:1-99
- [15] ATLAS Photos. Still image from the Episode 2 movie Photo 080302201. <http://atlas.ch/photos/events-general-detection.html>
- [16] D.H. Perkins. *Introduction To High Energy Physics*. Cambridge, University Press (2000)
- [17] F. Halzen and A.D. Martin. *Quarks and leptons: an introductory course in modern particle physics*. Wiley. (1984)
- [18] J. Beringer *et al.* *The Review of Particle Physics* (Particle Data Group), Phys. Rev. **D86**, 010001 (2012)
- [19] G. Gabrielse, D. Hanneke, T. Kinoshita, M. Nio, and B. Odom, *New Determination of the Fine Structure Constant from the Electron g Value and QED* Phys. Rev. Lett. **97**, 030802 (2006), Erratum, Phys. Rev. Lett. **99**, 039902 (2007)
- [20] Siegfried Bethke. *Experimental Tests of Asymptotic Freedom* arXiv:0606035 [hep-ex] (2006)
- [21] The ATLAS Collaboration. *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*. [arXiv:1207.7214] [hep-ex] (2012)

- [22] The CMS Collaboration. *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*. arXiv e-prints, [arXiv:1207.7235] [hep-ex] (2012)
- [23] K. Nakamura et al. *Review of Particle Physics: The CKM Quark-Mixing Matrix* Journal of Physics G **37** (75021): 150 (2010)
- [24] Stephen P. Martin. *A Super Symmetry primer*. [arXiv:9709356v6] (1997)
- [25] GianF. Giudice. *Beyond the standard model*. [arXiv:9605390v1] (1996)
- [26] R. M. Harris, C. T. Hill, and S. J. Parke. *Cross-section for topcolor Z0 decaying to tt: Version 2.6*, [arXiv:9911288] [hep-ph] (1999)
- [27] K. Agashe, A. Belyaev, T. Krupovnickas, G. Perez, and J. Virzi. *LHC Signals from Warped Extra Dimensions*. Phys. Rev. D **77** 015003, [arXiv:0612015] [hep-ph] (2008)
- [28] B. Lillie, L. Randall, and L.-T. Wang. *The Bulk RS KK-gluon at the LHC*. JHEP **0709** 074, [arXiv:0701166] [hep-ph] (2007)
- [29] CDF Collaboration and D0 Collaboration. *Combination of CDF and D0 results on the mass of the top quark using up to 5.8 fb⁻¹ of data*. [arXiv:1107.5255v3] (2011)
- [30] M. Baak et al. *Updated Status of the Global Electroweak Fit and Constraints on New Physics*. [arXiv:1107.0975] (2011)
- [31] R. Bonciani et al. *NLL resummation of the heavy-quark hadroproduction cross-section*. Nucl. Phys. B **529**, 424 [arXiv:9801375] [hep-ph] (1998)
- [32] M.Cacciari et al. *The t-tbar cross-section at 1.8 and 1.96 TeV: a study of the systematics due to parton densities and scale dependence*. High Energ.Phys. 0404, **068** [arXiv:0303085] [hep-ph] (2004)
- [33] T.Plehn et al. *Stop Reconstruction with Tagged Tops*. JHEP 1010:**078** [arXiv:1006.2833] [hep-ph] (2010)
- [34] T.Plehn, G.P. Salam and M.Spannowsky. *Fat Jets for a light Higgs*. Phys. Rev. Lett. **103** 241803, [arxiv:0910.5472] (2009)

- [35] J.M Butterworth *et al.* *Jet Substructure as a new Higgs search channel at the LHC.* Phys. Rev. Lett. **100** 242001, [arxiv:0802.2470] [hep-ph] (2008)
- [36] J.M Butterworth, B.E. Cox, and J.R. Forshaw. **WW scattering at the CERN LHC** Phys. Rev. Lett **D65** 0096014, [arxiv:0201.098] [hep-ph] (2002)
- [37] The ATLAS Collaboration. *ATLAS Experiment Online Luminosity Measurement.* <http://atlas.ch/proton-data-2012.html> (2012)
- [38] M. Dasgupta, L. Magnea, and G.P. Salam. *Non-perturbative QCD effects in jets at hadron colliders.* JHEP **02**. 055, [arxiv:0712.3014] [hep-ph] (2008)
- [39] M. H. Seymour. *Searches for new particles using cone and cluster algorithms: A Comparative study.* Z. Phys **C62** 127-138 (1994)
- [40] A. Hoecker *et al.* *Toolkit for Multivariate Analysis with ROOT.* TMVA version 4.0.3. CERN-OPEN-2007-007 [arXiv:0703039] (2009)
- [41] Tilman Plehn, Michael Spannowsky, Michihisa Takeuchi. *How to Improve Top Tagging.* [arXiv:1111.5034] (2011)
- [42] S. Catani. *Aspects of QCD, From the Tevatron to the LHC.* Proc. Workshop on Physics at TeV Colliders, Les Houches, France, Frontieres (Gif-sur-Yvette) [arxiv:0005233] [hep-ph] (1999)
- [43] R. Eusebi *et al.* *Combination of measurements of the relative fraction of ttbar events produced via gluon-gluon fusion.* CDF Note. <http://www-cdf.fnal.gov/physics/new/top/2007/topProp/ttProductionCombo/index.html>
- [44] Hadronic working group of the BOOST2010 workshop. *Boosted objects: a probe of beyond the standard model physics.* Eur.Phys.J **C71**:1661 (2011)
- [45] Y. Dokshitzer *et al.* *Better Jet Clustering Algorithms.* JHEP 9708 001, 835 [arXiv:9707323] [hep-ph] (1997)

- [46] S. Frixione and B.R. Webber. *Matching NLO QCD computations and parton shower simulations.* *JHEP* 0206 029 [arXiv:0204244] [hep-ph] (2002).
- [47] S. Frixione *et al.* Matching NLO QCD and parton showers in heavy flavour production. *JHEP* 0308 007, [arXiv:0305252] [hep-ph] (2003)
- [48] The ATLAS Collaboration. *Performance of large- R jets and jet substructure reconstruction with the ATLAS detector* ATLAS-CONF-2012-065
- [49] The ATLAS Collaboration. *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s}=7$ TeV.* [arXiv:1112.6426] [hep-ex] (2011)
- [50] M.L. Mangano *et al.* *ALPGEN, a generator for hard multiparton processes in hadronic collisions* [arXiv:0206293] [hep-ph] (2002)
- [51] John Allison. *Facilities and Methods: Geant4 A Simulation Toolkit.* Nuclear Physics, 17(2):2024, (2007)
- [52] The ATLAS Collaboration. *The ATLAS Simulation Infrastructure.* Eur.Phys.J., C(70):823 874, (2008)