ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

SEARCH FOR DIRECT SUPERSYMMETRIC TOP QUARK WITH DEEP LEARNING TECHNIQUES AT LHC AND HL-LHC

M.Sc. THESIS

Ceren GÜZELGÜN

Department of Physics Engineering

Physics Engineering Programme

SEPTEMBER 2018

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

SEARCH FOR DIRECT SUPERSYMMETRIC TOP QUARK WITH DEEP LEARNING TECHNIQUES AT LHC AND HL-LHC

M.Sc. THESIS

Ceren GÜZELGÜN
(509161104)

Department of Physics Engineering

Physics Engineering Programme

Thesis Advisor: Assoc. Prof. Dr. Altan ÇAKIR

SEPTEMBER 2018

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

DERİN ÖĞRENME TEKNİKLERİ İİLE LHC VE HL-LHC'DE SÜPERSİMETRİK TOP KUARK ARAYIŞI

YÜKSEK LİSANS TEZİ

Ceren GÜZELGÜN
(509161104)

Fizik Mühendisliği Anabilim Dalı

Fizik Mühendisliği Programı

Tez Danışmanı: Assoc. Prof. Dr. Altan ÇAKIR

EYLÜL 2018

Ceren GÜZELGÜN, a M.Sc. student of ITU Graduate School of Science Engineering and Technology 509161104 successfully defended the thesis entitled "SEARCH FOR DIRECT SUPERSYMMETRIC TOP QUARK WITH DEEP LEARNING TECHNIQUES AT LHC AND HL-LHC", which he/she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Assoc. Prof. Dr. Altan ÇAKIR** Istanbul Technical University	.....
Jury Members :	**Prof. Dr. Cenap Ş. Özben** Istanbul Technical University	.....
	Assis. Prof. Dr. Bora Işıldak Ozyegin University	.....
		.....

Date of Submission : **10 August 2018**
Date of Defense : **14 September 2018**

To my family and friends,

FOREWORD

First and foremost I would like to thank my supervisor Assoc. Prof. M. Altan Cakir for putting up with me in my most hopeless moments, his patience, and his motivation. Whatever success I have today is thanks to him. I would also like to thank Assoc. Prof. Tolga Birkandan for pushing me towards particle physics when I first graduated.

I am thankful to my oldest friend Gizem Bayram for being there in every crisis since I was six years old. My gratitudes to Mustafa Can Uslu for challenging me constantly and for his neverending support in all the important moments. And to rest of my friends, thank you all for offering whatever help you can.

As comes to family, I am thankful to my mother Oya Yesildag, without a doubt the best mother in the world. To my aunt Banu and her husband Tayfun, for doing millions more than an aunt and uncle normally would do. To their son Keremcan, for being my brother. My grandparents Nuray and Ismet for their unbelievable love and support for me, and for each other. I owe many thanks to my father Tufan, and to his lovely wife Elvan for offering their love, support and a sweet escape in their wonderful home. Finally, I thank Tore Erel, who has been through all the midterms and finals, success and failures, and most importantly this thesis, with me. You are an amazing human being.

My gratitude to these people and many more I may have forgot, the annoying belief you all seem to have for me is what brought me to this day. So long, and thanks for all the fish!

September 2018 Ceren GÜZELGÜN

TABLE OF CONTENTS

Page

ABBREVIATIONS

AdaBoost : Adaptive Boosting
ALICE : A Large Ion Collider Experiment
ANN : Artificial Neural Network
ATLAS : A Toroidal LHC Apparatus
BDT : Boosted Desicion Trees
BSM : Beyond Standard Model
CERN : European Organisation for Nuclear Research
CMS : Compact Muon Solenoid
ECAL : Electromagnetic Calorimeter
GUI : Graphical User Interface
HCAL : Hadron Calorimeter
HEP : High Energy Physics
HL-LHC : High-Luminosity LHC
LHC : Large Hadron Collider
LHCb : Large Hadron Collider beauty
MC : Monte Carlo
MSSM : Minimal Supersymmetric Standard Model
MVA : Multivariate Data Analysis
PDF : Probability Density Function
pMSSM : Phenomenological Minimal Supersymmetric Standard Model
QCD : Quantum Chromodynamics
QFT : Quantum Field Theory
SM : Standard Model
SUSY : Supersymmetry
TMVA : The Toolkit for Multivariate Data Analysis

SYMBOLS

$\mathbf{c}$	: Speed of Light
$\hbar$	: Planck's Constant
$\mathbf{p_T}$	: Transverse Momentum
$\mathbf{p_x}$, $\mathbf{p_y}$, $\mathbf{p_z}$	: Momentum components
$\mathbf{m_T}$	: Transverse Mass, Invariant Mass
$\mathbf{m_x}$, $\mathbf{m_y}$, $\mathbf{m_z}$	: Mass components
E_T^{miss}	: Missing Transverse Energy

LIST OF TABLES

Page

LIST OF FIGURES

Page

SEARCH FOR DIRECT SUPERSYMMETRIC TOP QUARK WITH DEEP LEARNING TECHNIQUES AT LHC AND HL-LHC

SUMMARY

The idea that certain particles can come together to form up matter has been around since 6 BC. Although these ideas started gaining physical credibility around the 19th century, the description of elementary particles have changed multiple times throughout the years. At first, elementary particles were thought to be indestructible, nevertheless we now know thanks to modern physics that they can decay, go under destructive collisions and either stop existing, or cause new particles to be created.

The search for the building blocks of matter gave way to smaller and smaller particles throughout years e.g molecules that are made up of atoms, atoms that are made up of subatomical particals. Particle physics take interest in these particles, their interactions, and their behaviour under certain energy scales.

For more than half a century, the most advanced research on this area have been made at European Organization for Nuclear Research (CERN) in Geneva, Switzerland. About 11000 scientist coming from over 100 countries around the world have been benefiting from the tools in this gigantic facility. Worlds largest and most complex equipments are being used at CERN to have a better understanding of the universe.

Large Hadron Collider (LHC) is a particle accelerator that is 25 km long in circumference and placed 175 meters underground. This complex structure is the largest experimental facility that has ever been built. First research with LHC was done in March, 2009 and continued until 2013. A world record has been broken with this research when a 4 TeV of energy per beam was reached. LHC had to go under maintenance for two years after this experiment, and started running again in early 2015.

LHC focuses on testing various particle physics theories. It seeks the answers to the questions such as the properties of the Higgs boson or observing particles that are anticipated by beyond the standard model theories such as supersymmetry and exotic models.

Seven detectors are placed at four points of the collider. Although LHC primarily collides protons, if necessary it can also work with heavy ions. Proton-lead collisions in 2013 and 2016, or lead-lead collisions in 2010, 2011, and 2013 are given examples.

Another record breaking quality of LHC is the computing-grid. Data from collisions are produced at an extraordinary rate, and processed at 170 computing centers located in 42 countries around the world.

The afforementioned detectors built on LHC serve to 4 main experiments named ATLAS, CMS, ALICE and LHCb. CMS and ATLAS experiments seek the answers to the same question with different experimental techniques. The difference emerges from the use two individually built, general purpose detectors that bear the same name

with their corresponding experiments. ALICE experiment takes interest in quark-gluon plasmas, whereas LHCb dwells on the aftermath of antimatter.

Monte Carlo (MC) simulation is a vital tool in experimental particle physics. MC is used in the high energy physics (HEP) area for designing or optimizing detectors, simulating particle interactions within the detector, and physics analysis. It helps simulate predicted physics such as the SM and BSM particles, is an effective tool in background estimation and event selection. In this manner this thesis uses a large portion of Monte Carlo samples to understand unknown interactions at LHC.

SUSY samples used in this thesis are produced with Madgraph and Pythia, whereas the standard model (SM) background comes from Snowmass effort. Various deep learning methods were studied to separate signal from the background. Detector response was simulated with Delphes software, letting us observe different pile-up scenarios. No pileup, 50 pileup, and 140 pileup are chosen as different cases. Signals at no pileup and 50 pileup are simulated at LHC conditions, with $300 fb^{-1}$, whereas 140 pileup case is simulated at the High Luminosity LHC (HL-LHC) $3000 fb^{-1}$.

Tree format is consisted of branches and leaves that contain physical information. At this step the files are large enough to cause disadvantages for analysis. After deciding on the necessary information and variables, lighter ntupled trees are generated.

Certain cuts rooting from physical properties were applied in order to separate signal from background. Two steps of cuts are followed, namely; baseline and signal cuts. In baseline cuts good events are selected and made sure the resulting events define the signature we want.

In signal cuts the first step is to decide on good kinematical variables that supress the SM background the best. In this thesis, these variables are H_T, E_T^{miss}, M_T and M_{T2}^W. A new varaible is proposed that is defined as $\alpha_{M_{T2}^W}$. Cutflow tables were then built for left signal and background events.

Toolkit Multivariate Data Analysis (TMVA) software was used to analyse kinematical variables. Boosted desicion trees (BDT) and neural networks (MLP) methods were chosen for analysis. By doing the trainings at LHC and HL-LHC, the effect of pileup is carefully observed. BDT and MLP responses of the same signals at different pileup scenarios differed from each other. The aim in using different variables at the same conditions is to have a better understanding of their performance.

The thesis is structured as follows: an introduction to the standard model and supersymmetry is given in Chapter 1. Technical details of the Large Hadron Collider and CMS detector is discussed in Chapter 2. Chapter 3 is devoted to the analysis strategy, selections, kinematic variables and multivariate analysis techniques. Results are interpreted and discussed in detail in Chapter 4. Finally, in Chapter 5 the work is concluded with final remarks, and improvements were discussed for future works.

DERİN ÖĞRENME TEKNİKLERİ İİLE
LHC VE HL-LHC'DE SÜPERSİMETRİK TOP KUARK ARAYIŞI

ÖZET

Belli tiplerdeki parçacıkların bir araya gelerek maddeyi oluşturması, milattan önce altıncı yüzyılda bile filozofların akıllarında yer eden bir düşünceydi. Bu fikirler 19. yüzyılın başlangıcında fiziksel olarak da güvenilirlik kazanmaya başlasalar da, günümüzdeki haliyle bildiğimiz "temel parçacık" konsepti seneler içerisinde anlamını bir çok kere değiştirdi. Örneğin modern fizik; uzun yıllar boyu "dayanıklı" kabul edilen temel parçacıkların bozunabileceğini ya da yıkıcı çarpışmalara girebileceğini gösterdi. Temel parçacıklar bu çarpışmaların sonucunda var olmayı durdurabilirler, veya yeni parçacıkların oluşmasına sebep olabilirler.

Maddenin yapıtaşı arayışı tarih boyunca daha küçük yapıda parçacıkların bulunmasına yol açtı. Molekülleri oluşturan atomlar, atomları oluşturan atom altı parçacıklar gibi. Parçacık fiziği, evreni oluşturan bu parçacıkları, etkileşimlerini ve çeşitli enerji seviyelerindeki yapılarını inceler.

Bu alanda en gelişmiş çalışmalar yarım yüzyılı aşkın bir süredir İsviçre'nin Cenevre kentinde konuşlanmış olan Avrupa Nükleer Araştırmalar Merkezi (CERN)'de yürütülmektedir. 21 üye ülke başta olmak üzere, dünyanın 100 farklı ülkesinden gelen (veya uzaktan çalışan) 11000 bilim insanı, CERN'ün imkanlarından faydalanmaktadır. Dünyanın en büyük ve kompleks enstrümanları temel parçacıkları daha iyi anlayabilmek için bu merkezde kullanılır.

Large Hadron Collider (LHC), yerin 175 metre altında ve 25 kilometre çapındaki dairesel bir tünelin içerisinde yer alan bir parçacık hızlandırıcısıdır. Bu komplex yapı, inşa edilmiş en büyük ve kompleks deneysel fasilitedir. Hızlandırıcıdaki ilk araştırma, 2010 yılının mart ayında başlamış ve 2013'e kadar sürmüştür. Bu araştırma çerçevesinde ışın başına 4 teraelektronvolt mertebesinde enerjilere ulaşılmış, ve bir hızlandırıcı için o zamanın dünya rekoru dörde katlanmıştır. İki yıllık bir bakım sürecine giren LHC, 2015'in başlarında yeniden çalıştırılmış ve günümüzde de çalışmaya devam etmektedir.

LHC'nin odağı çeşitli parçacık fiziği teorilerini test etmektir. Higgs parçacığının özellikleri, süpersimetri tarafından öngörülen yeni parçacıkları gözlemleyebilmek gibi, henüz çözülememiş fizik sorularına yanıtlar aramaktadır.

Hızlandırıcının dört noktasında, farklı amaçlara hizmet eden yedi dedektör bulunur. LHC, genel olarak proton demetlerini çarpıştırsa da, gerektiğinde ağır iyonlarla da deneyler yapılabilir. 2013 ve 2016 yıllarında yapılan proton-kurşun çarpışmaları, 2010, 2011 ve 2013 yıllarında tekrarlanan kurşun-kurşun çarpışmaları buna örnek verilebilir.

LHC'de bir başka rekortmen, hesaplama ağıdır. Çarpışmala verisi, eşi benzeri görülmemiş bir hızda üretilir, ve dünya üzerinde 42 ülkede bulunan 170 merkezden oluşan bir ağda işlenir.

Önceki paragraflarda bahsi geçen LHC üzerindeki dedektörler, 4 ana deneye hizmet eder. Bu deneyler ATLAS, CMS, ALICE ve LHCb olarak isimlendirilmiştir. CMS ve ATLAS deneyleri, aynı fiziksel sorunun cevabına farklı tekniklerle ulaşmaya çalışmaktadır. Deneylerde farklı yapılarda tasarlanmış, genel kullanıma yönelik dedektörler kullanılır. ALICE deneyi kuark-gluon plazmayı incelerken, LHCb madde ve antimadde miktarı arasındaki eşitsizliği araştırır. Tezin içeriğini oluşturan simülasyon ve incelemeler CMS dedektörü bünyesinde gerçekleştirildiğinden, özette sadece bu deneyden detaylıca bahsedilmesi uygun görülmüştür.

Türkçe anlamı "kompakt müon solenoit" olan CMS, LHC'de yer alan iki ana dedektörden biridir. Bu dedektörle yapılan deneyler kara maddeyi oluşturabileceği düşünülen parçacıkları araştırma, ekstra boyutlar ve Higgs bozonunun özellikleri gibi konulara yöneliktir.

CMS dedektörü 21.6 metre uzunluğunda, 15 m çapında ve 14000 ton ağırlığındadır. Fransa'nın Cessy kentinde, yerin altında bulunur. 43 ülke ve 199 enstitüden gelen 3800 bilim insanı, dedektörü kuran ve işleyişini sağlayan CMS işbirliğini oluşturur. Deney üyeleri, ATLAS deneyi üyeleri ile birlikte 2012 yılının temmuz ayında, Higgs bozonunun keşfini açıklamışlardır. Bu, fizik tarihinde çığır açan bir gözlem olmakla birlikte, fizikte hala cevaplanamamış bir çok önemli soru olduğu unutulmamalıdır. Bu soruların başında, özetin devamında değinilecek olan Standart Model'in eksiklikleri, kara madde, ve kara enerji gelir.

Elektromanyetik, güçlü nükleer, zayıf nükleer kuvvetler ile yerçekimi, evrendeki dört temel kuvvetlerdir. Evreni oluşturan parçacıkları ve bu kuvvetleri tanımlayan teoriye ise, Standart Model adı verilir. Bu teorinin oluşturulmasındaki ilk adım, 1961 yılında Sheldon Glashow'un elektromanyetik ve zayıf etkileşimleri birleştirebilmesiyle atılmıştır.

Standart modele göre 1/2 spinli fermiyonlar ile 1 spinli bozonlar (ve bu fermiyon ile bozonların antiparçacıkları) temel parçacıkları oluşturur. Model 3 farklı lepton ve kuark jenerasyonu içerir, her jenerasyon 2 üye barındırır. Modeldeki 6 kuark üst, alt, tılsım, garip, yukarı ve aşağı olarak isimlendirilmiştir. 6 adet lepton; elektron, müon, tau ve bunların nötrinolarından oluşur. Kuvvet etkileşimlerini sağlayan 8 gluon, 1 foton, Z^0, 2 adet $W^{\pm}$ ve 2012 yılında keşfi ile standart model parçacıklarını tamamlayan Higgs bozonu olmak üzere toplam 13 adet bozon bulunur.

Kendi içinde tutarlı olan ve fizik tarihinin en sağlam teorilerinden birisi olan standart model, yine de bazı soruları cevaplamakta güçlük çekmektedir. Madde-antimadde arasındaki dengesizlik, evrenin neredeyse %95'ini oluşturan kara madde ve kara enerji gibi konular henüz bu modelin yardımıyla açıklanamaktadır.

Bu nedenle fizikçiler, cevabı standard modelin ötesindeki teorik modellerde aramaya başlamışlardır. Deneysel anlamda bir kanıt sağlanamamış olsa da, teorik yönden otuz yılı aşkın bir süredir en başarılı olan kuram, Süpersimetridir (SUSY). Standart model tarafından tanımlanan her parçacığın bir partneri olduğunu savunan SUSY, fermion ve bozonlar arasında bir simetri kurar. Hesapsal olarak fermiyonları bozonlara, bozonları da fermiyonlara dönüştürebilir. Teorik alanda vaat ettiği başarıların da yardımıyla SUSY günümüzde hala bir çok deney ve teori grubu tarafından üstünde çalışılan bir alandır.

SUSY araştırmalarında aranan parçacıklar sinyal olarak adlandırılırken kalan bütün standart model kaynaklı fiziksel olaylar ardalan olarak bahsedilir. Sinyali ardalandan

ayırt edebilmek için bir çok teknik kullanılmakta olup, derin öğrenme ve yapay zeka sık başvurulan yöntemlerdendir.

Tezde SUSY sinyalleri Monte Carlo simülasyon teknikleriyle üretilmiş olup, stardart modelden kaynaklanan ardalan Madgraph ve Pythia isimli bilimsel yazılımlar yardımıyla sağlanmıştır. Sinyal ve ardalanı birbirinden ayırabilmek adına çeşitli derin öğrenme yöntemleri üzerinde çalışılmıştır. Dedektör etkisi sinyal ve ardalan verilerine Delphes adındaki bir başka bilimsel yazılım yardımıyla eklenmiş, ve veriler ROOT veri analizi programı tarafından okunabilen ağaç (tree) formatına çevrilmiştir. Farklı yığılım sonuçları, Delphes programındaki kartlar yardımı ile simüle edilmiş, yığılımsız ve 50 yığılımlı senaryolar $300 fb^{-1}$ luminositede incelenirken 140 yığılımlı senaryoda $3000 fb^{-1}$ kullanılmıştır.

Bu ağaç formatında fiziksel bilgiler barındıran veri dalları bulunur ve ilk aşamada Delphes'ten çıkan bu dosyalar dezavantaj sağlayacak büyüklüklere sahiplerdir. Analizin yararına olacak parametreler belirlendikten sonra özelleştirilmiş yazılım yardımıyla bu bilgileri barındıracak ntuple adı verilen yeni, ve daha küçük boyuttaki ağaçları yaratılmıştır.

Sinyali ardalandan ayırmak amacıyla, bilinen fiziksel gerçeklerden yola çıkılarak veride bir takım sınırlamalar - veya kesintiler - uygulanmıştır. Örneğin, sadece belirlenen bir momentum sınırından yüksek veya alçak olayları kabul etmek gibi. İki aşamalı bir kesme sistemi uygulanmakta olup, ilk aşamada genel parametrelere dikkat edilirken ikinci aşamada sinyal bölgesini ele alarak hareket edilmiştir.

Yıllardır bu alanda yapılan çalışmalar ve yayınlardan yararlanarak sinyali ardalandan ayırmada en etkili değişkenler H_T, E_T^{miss}, M_T ve M_{T2}^W olarak seçilmiş olup, $\alpha_{M_{T2}^W}$ isimli yeni bir değişken de tez sahibi ve danışman hocanın çalışmaları sonucunda teklif edilmektedir. Bu değişkenlere uygulanan işlemler sonucu kalan sinyal ve ardalan miktarı sayılır.

Son adım, değişkenler üzerinden analiz yapan TMVA (Toolkit Multivariate Data Analysis) programına seçilen değişkenlerden oluşan ağaçların beslenmesidir. TMVA bünyesinde ötelenmiş karar ağacı (BDT) ile, çok katmanlı sinir ağları (MLP) metodları incelenmiştir. Aynı sinyallere farklı luminosite ve yığılım senaryolarında bakarak, yığılımın öğrenimleri nasıl etkilediği anlanmaya çalışılmıştır. Öğrenim sonuçlarındaki farklılıklar, yığılımın kendi başına bir etken olabileceğine işaret etmektedir. Aynı koşullarda farklı değişkenler kullanılarak yapılan öğrenimlerde ise, bu değişkenlerin performansları daha iyi anlanmaya çalışılmıştır.

Tezin planlanması şu şekilde yapılmıştır: Bölüm 1'de Standart Model ve Süpersimetri teorilerine kısa bir giriş yapılmıştır. Bölüm 2, LHC ile CMS'in teknik yapısını tartışır. Bölüm 3, analiz stratejisini, yapılan seçilimleri, kullanılan değişkenleri ve derin öğrenim metodlarını tanıtır. Elde edilen sonuçlar, Bölüm 4'te ayrıntılı bir şekilde tartışılır. Bölüm 5, analizi ve elde edilen sonuçlardan dersler çıkararak gelecek işlerde yapılabilecek iyileştirmeler için öneriler teklif eder.

1. INTRODUCTION TO PARTICLE PHYSICS

1.1 The Standard Model

Standard Model (SM) is a mathematical description of the elementary particles of matter and three of the four known fundamental forces; namely, electromagnetic, weak and strong interactions. These elementary particles can be categorised according to their certain properties i.e. mass, energy, momentum and electrical charge. Another cruically important property is called *spin*, describing an elementary particle's intrinsic angular momentum. As an intrinsic property it has a fixed value that can be associated with certain types of particles. Due to quantum mechanical rules that apply to small particles the value of spin can only be an integer or an half-integer [10]. Particles with half-integer spins are called *fermions*, whereas the integer spinned ones *bosons*. Fermions of SM are summarized in Table 1.1. For every particle shown in 1.1, there is an antiparticle with all the parameters exact but charge. An antiparticle's charge is the reversed sign value of a particle.

Table 1.1 : Summary of three generaions of fermions in the SM with their electrical charge.

Generation	Quarks		Leptons	
1	u up	d down	e electron	ν_e electron neutrino
2	c charm	s strange	μ muon	ν_μ muon neutrino
3	t top	b bottom	τ tau lepton	ν_τ tau neutrino
El. charge (e)	$+2/3$	$-1/3$	-1	0

As mentioned in the first paragraph, the SM only explains three out of four fundamental forces. Gravity is not included in the theory, due to difficulties in combining general relativity with quantum mechanics. Gravity is known to be 10^{40} times weaker than the electromagnetic force and thus its effect is expected to be negligible. The SM forces are mediated via spin-1 boson exchange. Table 1.2 summarizes the SM bosons and force carrying particles.

Table 1.2 : Summary of the SM bosons with their mass, electrical charge and spin.

Boson	Mass [GeV]	Electrical Charge [e]	Spin
γ	0	0	1
g	0	0	1
W^+, $W^\pm$	80.385 ± 0.015	± 1	1
Z	91.188 ± 0.002	0	1
H	125.09 ± 0.24	0	0

The electromagnetic force is a force that is mediated by photons and it acts on charged particles. Almost all of the interactions we see in our environment is caused by the electromagnetic force. Along with chemical reactions, it covers electrical and magnetic interactions.

The strong force is mediated via gluons and it has a range around $\sim 10^{-15}$m. Unlike photon, gluon is charged under the influence of the strong force, hence it can couple to itself. Only color-charged particles are affected by the strong nuclear force. There are three kinds of color charge called red, green, and blue; also three anticolors associated with these. Quarks are charged with one color, and antiquarks with anticolor. Color charge is conserved, therefore gluons carry a color and an anticolor.

Color confinement results in only color-neutral particles being observed in nature. A particle is color-neutral in one or two ways. If a color get cancelled by its anticolor, namely; the combination of two such quarks. Such combination makes up *mesons.* Another way is when all color charges combine and lead to color-neutralness. Latter combination makes up *baryons*. As a consequence of aforementioned phenomenon, gluons or quarks can not be observed in the nature in bare form. These composite combinations are often mentioned as *hadrons*. Protons and neutrons are typical examples to hadrons, with the former containing two up quarks and and a down, whereas latter contains one up quark and two downs. Nevertheless it must be taken into account that a gluon can make up a quark-antiquark pair, and these quarks and antiquarks can recombine to form a gluon.

The weak nuclear force is in control of the decay of the particles. Unlike electromagnetic, or strong forces, the weak force acts on the entirety of SM fermions. This force has a range that is even shorter than the strong force. Depending on which

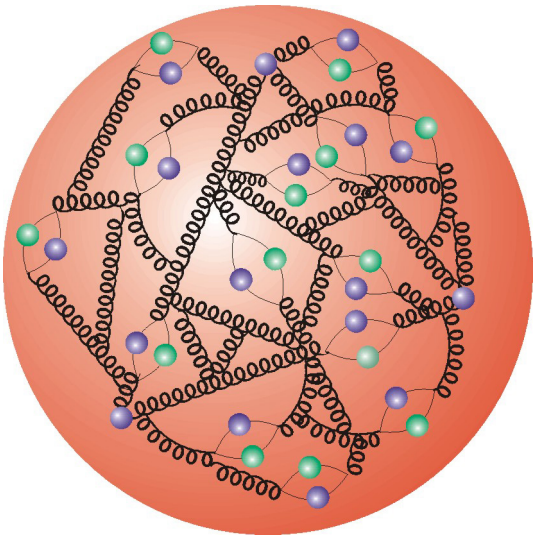


Figure 1.1 : Quarks are represented with green, while antiquarks orange. The image describes the interaction of quarks & antiquarks by exchanging gluons, which are represented with springs.

mediator is in charge ($W^{\pm 1}$ or Z), weak force can be responsible from neutral or charged current interactions. Due to charge conservation, W bosons will always link two particles with a charge difference of 1 e, i. e a lepton and a neutrino, an up and a down quark..

Higgs boson, (H) is the last particle of SM that is discovered in 2012. It is responsible for the mass of fermions, the W and Z bosons, and itself.

A summary of all the interactions within the SM is presented in Figure 1.2.

Quantum field theory (QFT) governs the behavior of fast moving small particles. It describes the fermions with a Dirac field $\psi(x)$. Lagrangian density $\mathscr{L}_\psi$, a function of the fields and their derivatives in the framework, explains the dynamics of a fermion and its antifermion.

$$\mathscr{L}_\psi = \mathscr{L}(\psi(x), \partial_\mu \psi(x)) = \mathscr{L}(\psi(x), \partial_\mu \psi(x), \psi^\dagger(x), \partial_\mu \psi^\dagger(x)), \tag{1.1}$$

where $\partial_\mu = \partial/\partial x^\mu$. An antifermion's field corresponds to the hermitian conjugate of a fermion's, clearly, $\psi^\dagger(x)$. When there is no interaction between a fermion and other particles, this density reduced to a kinetic, and a mass term,

$$\mathscr{L}_\psi = i\overline{\psi}\gamma^\mu \partial_\mu \psi - m\overline{\psi}\psi, \tag{1.2}$$

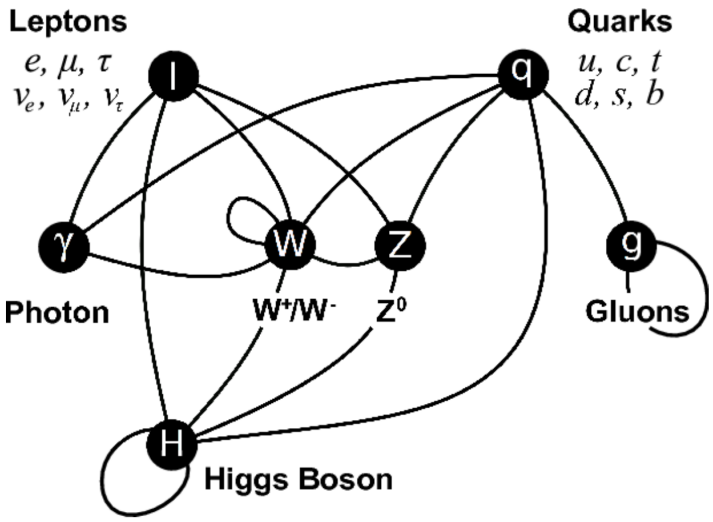


Figure 1.2 : A simple diagram that outlines the interactions between SM particles.

where m is the fermion mass, $\overline{\psi} = \psi^{\dagger}\gamma^{0}$ is the adjoint fermion field, and γ^{μ} are the Dirac matrices that are defined by $\{\gamma^{\mu}, \gamma^{\nu}\} = \gamma^{\mu}\gamma^{\nu} + \gamma^{\nu}\gamma^{\mu} = 2g^{\mu\nu}$.

Particle interactions emerge as a result of the symmetries that are enforced on the Lagrangian. If $\psi(x)$ transforms locally,

$$\psi(x) \rightarrow \psi'(x) = e^{i\alpha(x)}, \tag{1.3}$$

an extra term is added to the Lagrangian as a result of the derivation in Eq. (1.2),

$$\mathscr{L}_{\psi'} = i\overline{\psi}\gamma^{\mu}\partial_{\mu}\psi - \overline{\psi}\gamma^{\mu}(\partial_{\mu}\alpha)\psi - m\,\overline{\psi}\psi, \tag{1.4}$$

where $\alpha(x)$ in Eq. (1.3) is an arbitrary function of the local coordinates x. If the transformation governed by $\alpha(x)$ can be composed as a function of the generators of a Lie group, $\alpha(x) = \varepsilon(x)\cdot\vec{T} = \varepsilon^{i}T_{i}$, by including a vector boson field A^{i}_{μ} for every generator T_i, the Lagrangian can be made invariant under the transformation in Eq. (1.3).

The electroweak theory is a combination of the interactions of the electromagnetic and weak nuclear force, and $SU(2)\times U(1)$ characterize the symmetry of this theory. $SU(3)$ is the term that describes the symmetry of strong nuclear force, a force that is governed by quantum chromodynamics (QCD).

$$D_\mu = \partial_\mu + ig_s\frac{\lambda_a}{2}G^a_\mu + ig\frac{\sigma_i}{2}W^i_\mu + ig'\frac{Y}{2}B_\mu, \tag{1.5}$$

where g, g' and g_s are the coupling strengths of the vector fields in each term, is the covariant derivative that leaves the SM Lagrangian invariant under a G_{SM} transformation. It is established with the second term of Eq. (1.5) that the SM lagrangian is invariant under $SU(3)$ transformations. A set of eight linearly independent, 3×3 Hermitian matrices that are called the Gell-Mann matrices [10], are generators of the $SU(3)$ group and they describe the symmetry of QCD. From the eight matrices also arises eight gluon fields G^α_μ, with massless gluons as excitations.

The local phase transformation under $SU(2) \times U(1)$ governs the electroweak interactions. Some experiments [11] have observed electroweak parity violation, and in order to take this into account only left handed particles are allowed to couple to $SU(2)$ [12]. Pauli matrices and the hypercharge Y are the generators of the electroweak symmetry group. This results in three W^i_μ fields under $SU(2)$ transformation, and one B_μ field for $U(1)$. These fields can not be associated with any bosons in Table 1.2.

The massless vector fields contradict with observations, therefore mass terms need to be added to the SM Lagrangian in order to give mass to the W and Z bosons. This results in symmetry breaking in the gauge invariance of the Lagrangian. However as the Brout–Englert–Higgs mechanism [13, 14] demonstrates, electroweak symmetry can be broken spontaneously under specific conditions. The Lagrangian can be made invariant again by introducing a scalar field

$$\Phi = \begin{pmatrix} \phi^a_1 + i\phi^a_2 \\ \phi^b_1 + i\phi^b_2 \end{pmatrix}, \tag{1.6}$$

that breaks the symmetry of the vacuum state. This field has a potential of

$$V(\Phi) = \mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2, \tag{1.7}$$

where μ^2 is a mass parameter. The shape of the potential can be seen in Fig 1.3 in the case of $\mu^2 < 0$. Further details on the Higgs mechanism can be found in the following reference. [14].

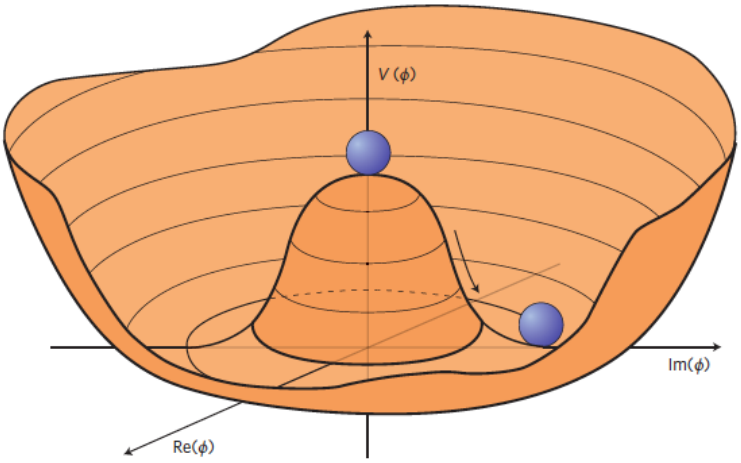


Figure 1.3 : An illustration that demonstrates the "Mexican Hat" potential that is responsible of the "spontaneous" symmetry breaking [2].

The SM is far from being an abstract theory. It is remarkably simple and very powerful. Mathematical framework allows theorists to make extraordinarily precise predictions. Over the past five decades, many quantities are being measured in particle physics laboratories all around the world, and almost all of these quantities fall on the mathematically predicted values by the SM.

If so, a question comes to mind: then what is missing in the SM? Essentially, the model lacks stability at higher energies.

1.2 Beyond the Standard Model

The SM is one of the most robust and succesfull theories in physics. It was able to predict the existence of the Higgs boson, 50 year before it experimentally observed. However, despite its capabilities, it also leaves some open questions in theoretical and experimental areas that are well worth addressing.

Most obvious question would be, what about gravity? Although there are many studies upon the issue [15] [16], gravity is not implemented into the SM. Even though gravity is negligible at the energy scales that we currently look into, the great role it plays at higher energy scales e.g the Planck scale, Λ_p 10^{19} GeV is undeniable. According to the SM, graviton must be the mediator for gravity, yet, at the time of writing this thesis it has not been observed. It is observed for the past 20 years that the expansion of the universe is accelerating. Measuring the magnitude of a supernovae granted the first clue for this [17].

This phenomenon could either be explained by accepting the cosmological constant Λ as zero, or that some exotic fields to the SM were needed. Our solar system, the stars, the sun, the galaxies; in other words the visible universe is all made of protons, neutrons and electrons. It is a mind blowing discovery that this baryonic matter is only less than 5 percent of the total mass of the universe. The rest is estimated to be 26 % of dark matter, and 69 % of dark energy that is hypothesized to cause the acceleration. None of the SM particles are suitable candidates for these observations.

Originally, the SM formulated the neutrinos to be massless. Nevertheless, the SNO experiment [18] could observe one third of the expected neutrinos, where the expected amount was determined from the sun's nuclear activity. The detector used in the experiment was designed to count the electron neutrinos arising from nuclear fusion reactions. A type of neutrino transforming into another, namely, neutrino oscillations are a possible explanation to this insufficiency. However, for this to occur neutrinos must have a different *mass*. The SM can not explain how do these neutrinos acquire their mass, but several models outside the SM have been proposed [19].

Theoretical developments have been made for these and other questions that SM fails to answer. These developments are generally referred to as physics beyond the standard model (BSM). Big portions of BSM studies mainly include supersymmetry (SUSY) searches, which is the main topic that governs this thesis.

1.2.1 Supersymmetry

SUSY [1] [20] [21] [22] [23] is a well motivated BSM theory that originates from a 1974 paper written by Wess and Zumino [24]. For the past three decades, the theory has been of great importance in theoretical physics developments, and a major influence on experimental physics.

Generally, SUSY proposes a symmetry between fermions and bosons by blending them together via assocating each particle with partners. A SM spin-1/2 fermion would be paired with a boson, and vice verca. The SM quarks would be paired with spin-0 *squarks*, a name made up for supersymmetric quarks. Likewise, the SM bosons have fermion superpartners.

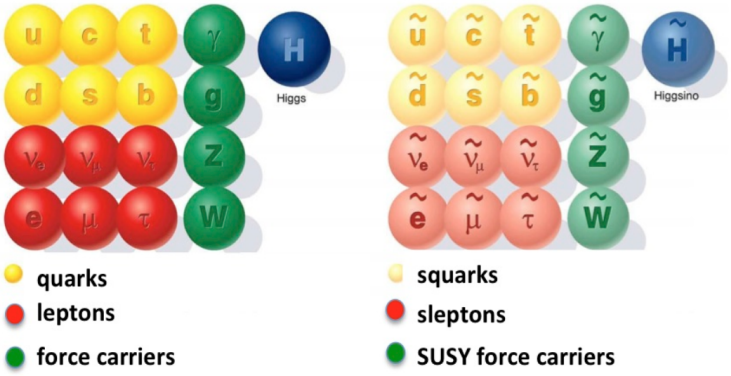


Figure 1.4 : The SM particles and their SUSY superpartners shown side by side. Mixing of force carrying superpartners are not depicted [3]

.

Exact invariance for this theory would describe a scenario if particles and sparticles have the same mass and quantum numbers. Only their spins would differ by 1/2. Such thing has not been observed in nature; e. g. a spin-0 particle with the same mass as an electron. But we already know that we don't live in a world where there is exact invariance under symmetry: if we were, all gauge bosons would have no mass. But it is well established that unlike photon, $W^{\pm}$ and Z have mass. The SM overcame this issue by with the spontaneous symmetry breaking. Therefore, it is only fair to claim there is another mechanism that would make SUSY a broken symmetry. There is an abundance of possible theories for SUSY.

Table 1.3 : Particle spectrum of the MSSM [1]. Brown and blue blocks describe sleptons and leptons, green and purple blocks describe squarks and quarks, red and orange blocks describe Higgs and Higgsino, respectively.

Superfield	**Bosons**	**Fermions**	$SU(1)_C$	$SU(2)_L$	$U(1)_Y$
Gauge					
G^a	g^a	$\tilde{g}^a$	8	0	0
V^k	$W^\mu(W^\pm, Z_0)$	$\tilde{W}^\mu(\tilde{W}^\pm, \tilde{Z}_0)$	1	3	0
V'	B/γ	$\tilde{B}/\gamma$	1	1	0
Matter					
L_i	$\tilde{L}_i = (\tilde{\nu}, \tilde{e})$	$L_i = (\nu, e)$	1	2	-1
E_i	$\tilde{E}_i = \tilde{e}_R$	$E_i = e_R$	1	1	2
Q_i	$\tilde{Q}_i = (\tilde{u}, \tilde{d})_L$	$Q_i = (u, d)_L$	3	2	1/3
U_i	$\tilde{U}_i = \tilde{u}_R$	$U_i = u_R$	3	1	-4/3
D_i	$\tilde{D}_i = \tilde{d}_R$	$D_i = d_R$	3	1	2/3
Higgs					
H_d	H_d	$\tilde{H}_d$	1	2	-1
H_u	H_u	$\tilde{H}_u$	1	2	1

In this thesis Minimal Supersymmetric Model (MSSM) is considered. The word *minimal* here emphasizes the fact that the model contains the smallest number of new particle states and interactions that are consistent with the phenomenology [25]. The particle spectrum of MSSM is shown in Table 1.3.

In the SM, the Higgs boson mass is not stable under quantum corrections. According to the theory the weak scale must be weaker compared to what has been observed.

2. THE LARGE HADRON COLLIDER EXPERIMENT

The Large Hadron Collider (LHC) experiment is the largest and compelling accelerator around the world. It is a 27-kilometer ring that is consisted of superconducting magnets and accelerating structures for boosting the particles. This ring is contained in a depth ranging from 50 to 175 meters underground, and placed in the region between the Geneva International Airport and Jura mountains. It got started up in 10th of September, 2008 and since then remains the largest addition to CERN's accelerator complex [26].

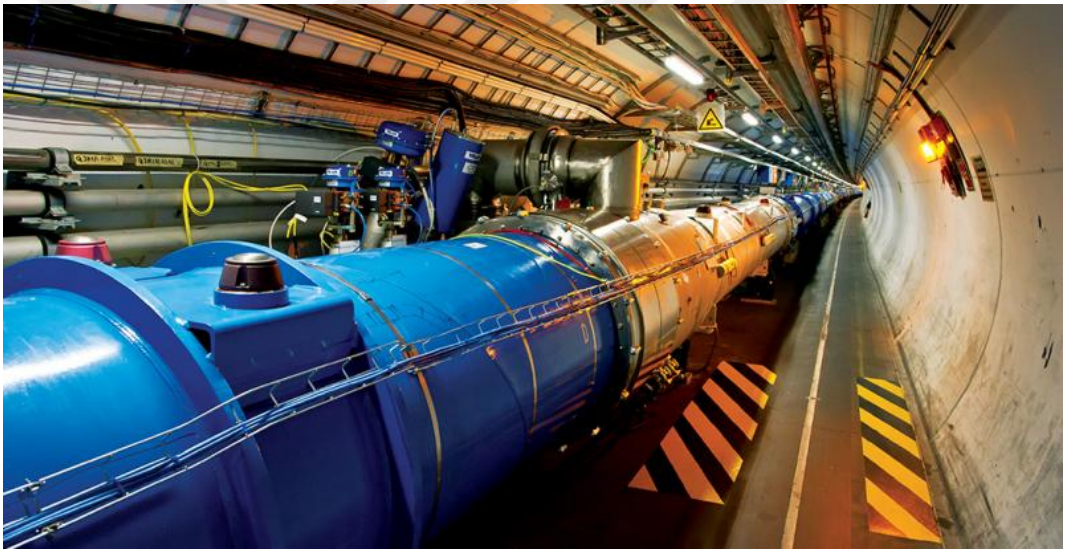


Figure 2.1 : A section of the LHC, located underground in CERN, Geneva [4].

Two high-energy beams travel inside the collider in the opposite directions before they collide with a speed that is close to that of light. These beams are consisted of ultra-relativistic particles and they move with 0.999997828 c at the time of injection, and reach a maximum of 0.999999991 c at top energy [27]. The beams are kept in ultra vacuum, by cryogenic pumping of 9000 cubic metres of gas, and the environment they are in is extremely cold. The magnets around are cooled down to $-271.3\,^{\circ}$C, a temperature that is even colder than the outer space. This requires an enormous supply of liquid helium.

The beams acquire direction with the help of 1232 dipole magnets that are in charge of bending, and 392 quadrupole magnets that provides focusing. Just before collision, different type of magnets "squeeze" the beams to increase the probability of colliding. An analogy described by CERN itself associates the task of making these beams collide with firing needles 10 kilometers apart with a precision so high that they meet on air, half-way [4].

The entire accelerator is controlled and maintained at CERN Control Centre. The centre aids the beams to collide in four different accelerators located on the ring. These accelerators (and corresponding experiments) are named A Toroidal LHC Apparatus (ATLAS), The Compact Muon Solenoid (CMS), A Large Ion Collider Experiment (ALICE), and Large Hadron Collider beauty (LHCb). Among these, only the CMS experiment will be discussed, for this thesis focus on simulations within the CMS detector.

2.1 The CMS Experiment

With a broad physics agenda ranging from inspecting Standard Model, to looking for particles that might associated with dark matter, CMS is one of the LHC detectors. The experiment's main purpose is to reach the same scientific goals with the ATLAS by using different technique and approaches. A grand example of this would be the announcement made on 4 July 2012, referring to the new particle observed both by ATLAS and CMS experiments in the mass region 126 GeV [28], [29].

The collaboration consists of scientists from all around the world working together towards advancing human knowledge on the universe. Over 4000 particle physicists, engineers, computer scientists, technicians and students coming from more than 40 different countries participate in the CMS experiment.

CMS detector weighs 14000-tonne, even two times heavier than the iron giant Eiffel Tower. True to its name, CMS is quite compact; being 15 meters high and 21 meters long. The solenoid magnet that the CMS is built around is the most powerful that is ever made.

The detector resembles a well advanced camera that records the collisions from all angles up to 40 million times each second [30]. Onion like layers make up

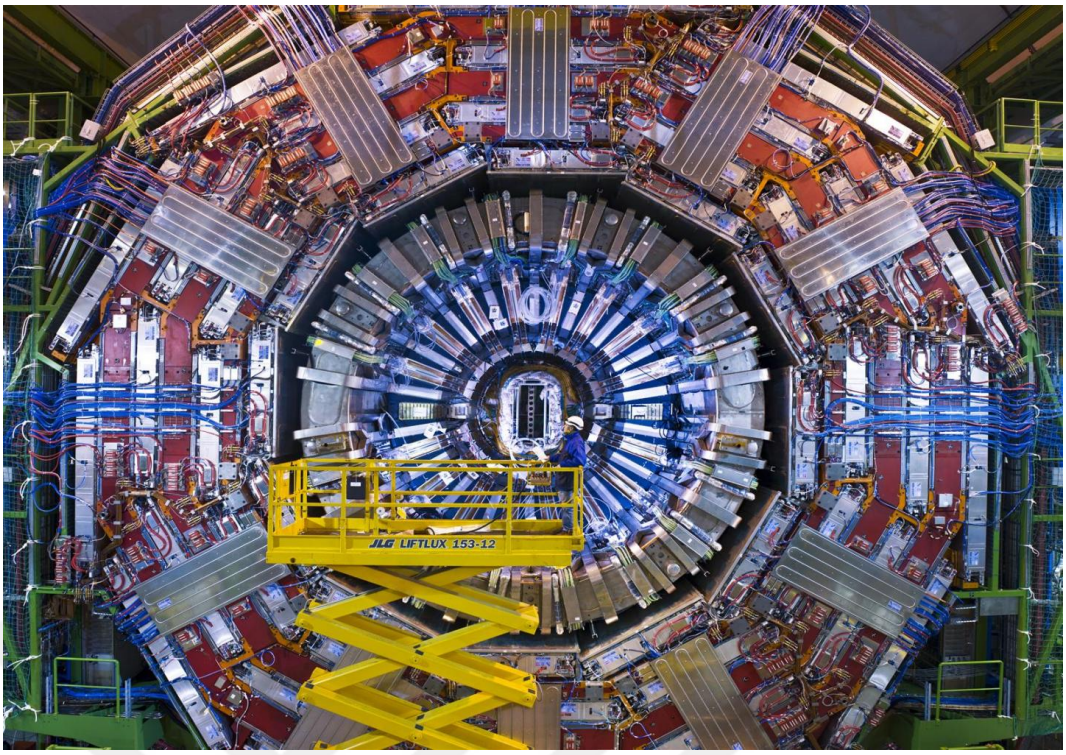


Figure 2.2 : The CMS detector that is located in one of the four collision points withing the LHC [5].

the detector's structure, each having a separate duty. These duties include bending particles, identifying the tracks, measuring energy, and detecting muons.

CMS's heaviest component is by 12500 tonnes, the magnet coils. This powerful magnet aids bending the particles as they scatter out from the collision point. Bending is necessary for a couple of reasons. First of all, identifying the particles by looking at the tracks of positively and negatively charges particles; for they bend in the opposite directions under the same magnetic field. Second purpose is to measure the momenta. Particles with high momentum bend less compared to low momentum particles.

The solenoid magnet is made up of a cylindrical coil of fibres that are superconducting. 18500 amps of electricity flow through and circulate within these coils, and superconductivity reassures that there is no resistance. Hence generated a magnetic field of 4 Tesla. Tracker and calorimeters are placed in this giant magnet.

After the bending, CMS must also correctly identify these particles from their tracks. This duty belongs to the silicon Tracker. It is the innermost layer. The tracker is able to reconstruct the hadron (composite particles made of quarks) and some lepton tracks:

muons and electrons. Tracks emerging from the decaying of short lived particles are also reconstructed.

Tracks must be reconstructed as precisely as possible, yet the tracker needs to be lightweight so it does not disarray the particles. This can be achieved by measuring the positions very accurately, so the reconstruction is reliable from just a few points. The accuracy of each measurement is to $10\mu m$, even thinner than a strand of human hair. Being at the core of the detector, the tracker must be made from materials that resist radiation damage as best they can.

After particle identification their energies in each collision must be calculated. This information is vital in understanding the physics happened at the vertex. Calorimeters of CMS assembles the information on energy. Two kinds of calorimeters are placed in the detector, Electromagnetic (ECAL) and Hadron (HCAL) calorimeters.

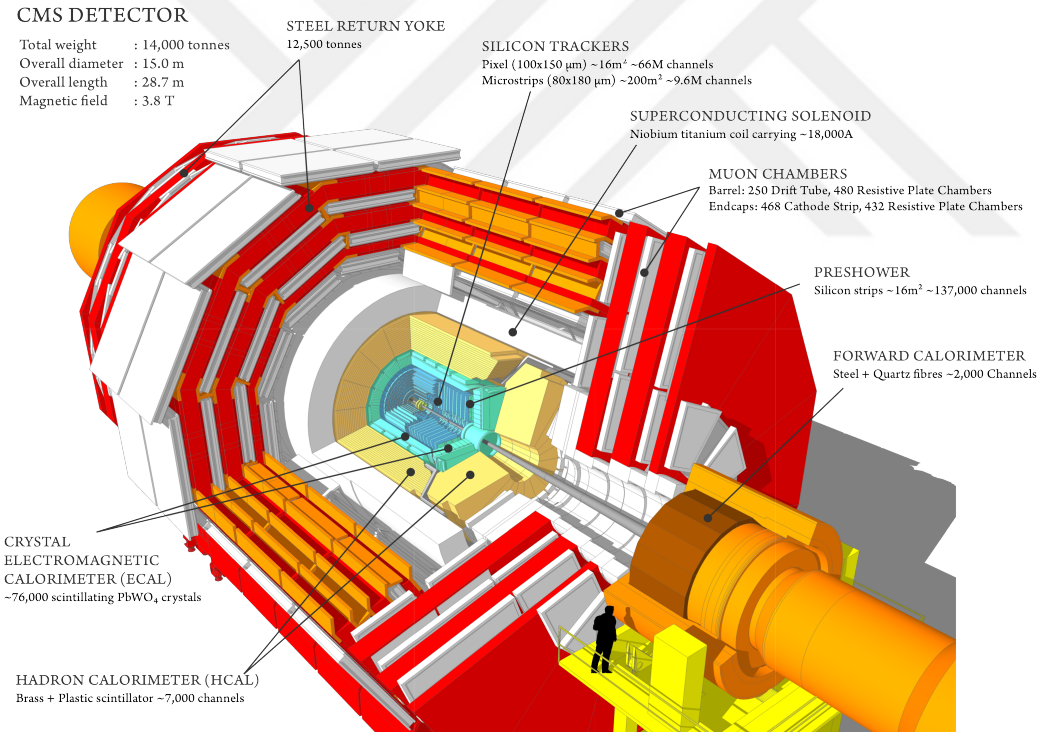


Figure 2.3 : The CMS detector that is located in one of the four collision points withing the LHC [6].

ECAL is the inside layer of the two calorimeters, and primarily it measures the energies of photons and electrons. These particles have been of great assistance in the quest of finding the Higgs boson, and they are suspected to be vital in pursue of new physics. The material that the calorimeter is made of is mostly see through, and it scintillates

whenever a photon or an electron gets past. This scintillation is essentially the light production that is proportional to the passing particle.

The calorimeter has two components called the "barrel" and the "endcaps". The barrel has a cylindrical structure and consists of 61200 crystals. Unlike the barrel, endcaps have a flat design with more than 15000 crystals.

Hadron calorimeter measures the energies of hadrons, and grants indirect measurements of neutral, non-interacting particles. These particles are of importance for they give information on the possible formation on the supersymmetric particles. Hadron decay may include daughter particles that are invisible to the CMS detector. In order to detect these particles, HCAL must capture every particle possible. If particles tend towards one specific side of the detector and not the others, an imbalance in the transverse plane emerges in momentum and energy. This hints at the production of "invisible" particles.

A scintillating system is also present in HCAL, whenever a particle passes through there is a light pulse. Fibers of optic collects this light, and photodetectors amplify the signal. The calorimeter consists of many layers, and summing up this light pulses in all the layers make up a "tower"; the measured energy of the particle. These layers of HCAL are spread in several regions, called the barrel (HB, HO), endcap (HE), and forward (HF).

One might deduct from the name Compact Muon Solenoid, that detecting muons is a vital duty in this experiment. The outer layer of the onion structure that makes up the CMS detector is devoted to detecting muons. The long distance muons can cover without interacting with iron makes them unstoppable by neither type of CMS calorimeters. Therefore, sub-detectors that are specifically designed for muon identification are placed within the CMS. The identification is mostly executed in the outer edges, for muons are the only type of particles that are likely to register a signal in that region.

3. ANALYSIS FRAMEWORK

3.1 Analysis Flow Diagram

Monte Carlo (MC) simulation is a vital tool in experimental particle physics. The method relies on a set of rules and random probabilistic events to mimic the atomic motion, defining the system by probability density functions (PDF). They describe partons coming into the process.

MC is used in the high energy physics (HEP) area for designing or optimizing detectors, simulating particle interactions within the detector, and physics analysis. It helps simulate predicted physics such as the SM and BSM particles, is an effective tool in background estimation and event selection.

Physics analyses at the Monte Carlo level within HEP follow a six step procedure that is summarized in Fig 3.2. The first two steps govern the generation of matrix element samples. Afther the Monte Carlo generation, we use DELPHES framework for simulating the detector response to analyze data.

Events are generated with Monte Carlo event generators such as PYTHIA, HERWIG, andMADGRAPH. Generation helps making predictions for collider experiments, and solving physics problems. Several steps make up the generation of an event [31], as illustrated in Figure 3.1.

- Hard Process
- Parton Shower
- Hadronization
- Underlying Event
- Unstable Particle Decay

In a collision the majority of the events lie along the beam axis. What is of interest is events containing high momentum-transfer, and only a fraction of events fit this category. Therefore, it is not necessary to simulate all possible proton-proton collisions. Instead, the simulation must be focused on the necessary hard process. This concept is achievable by starting to simulate at the center of the collision, then calculating the probability distribution of a hard scattering by using perturbation theory. With the help of PDFs and perturbation theory combined, simulating hard scattering is a straightforward process.

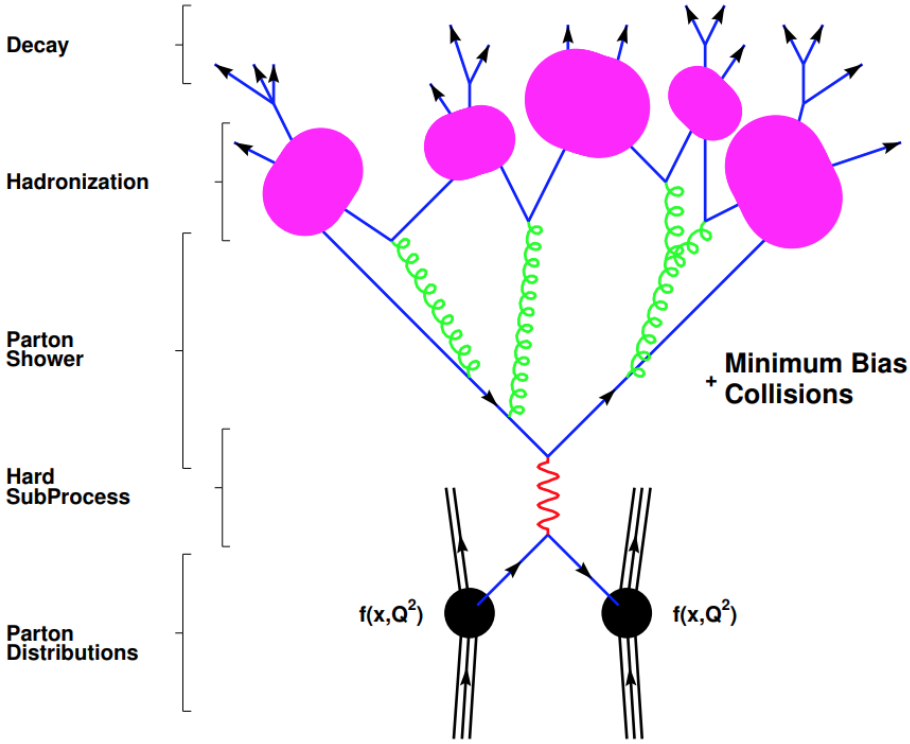


Figure 3.1 : Showering and hadronization structure in event generators. [7].

Parton shower step governs the aftermath of the incoming and outgoing partons in a hard collision. These partons are colored particles: quarks and gluons. The confinement of a system of partons into hadrons is taken into account in the hadronization step. Secondary interactions between proton remnants make up the underlying events. This causes a contamination in the hard process simulation, for underlying events produce soft hadrons everywhere. These hadrons are unstable particles with heavy resonances that decay afterwards, making up the final state of

event generation chain. Many improvements have been made to better model these secondary decays.

3.2 Event Samples

Many searches proceeding at the LHC for new physics beyond the SM are tailored to search for specific scenarios, or being kept as inclusive as possible. In this analysis, we focus on the direct top-squark pair production. Previous searches performed by CMS collaboration focus on simplified SUSY models [32].

Detector response is simulated both for signal and background events using DELPHES 3.0.10 fast simulation program [33]. This version of DELPHES is able to include pileup from minimum bias collisions. These collisions are selected randomly, from a file containing inelastic proton-proton interactions produced with PYTHIA8 [34]. It is a two step procedure of

- mixing of the pileup events and main interaction
- substracting the pileup with fast jet area technique.

Three separate cards are used to simulate no pileup, 50 pileup, and 140 pileup scenarios. PileUpMerger package is used to complete the mixing procedure. There are parameters that can be managed by the user, namely PileUpFile, MeanPileUp, PileUpDistribution, ZVertexWidth and ZVertexResolution.

Pileup file is a binary file that contains the pileup events. Binary format is substantially faster when reaching these random events compared to root trees. Four momentum, vertex positon and Particle ID information is stored in the file. In this thesis, the pileup file is prepared with PYTHIA8, and converted to binary.

MeanPileUp is the parameter that contains average pileup events per bunch-crossing. N Pileup events are randomly chosen from the PileUpFile for every hard scattering. PileUpDistribution decides on the random number N. 0 means Poisson distribution, whereas 1 is uniform.

Hard scattering and pileup events are randomly distributed in time, and z coordinate. This distribution is controlled with user specifications, with ZVertexSpread.

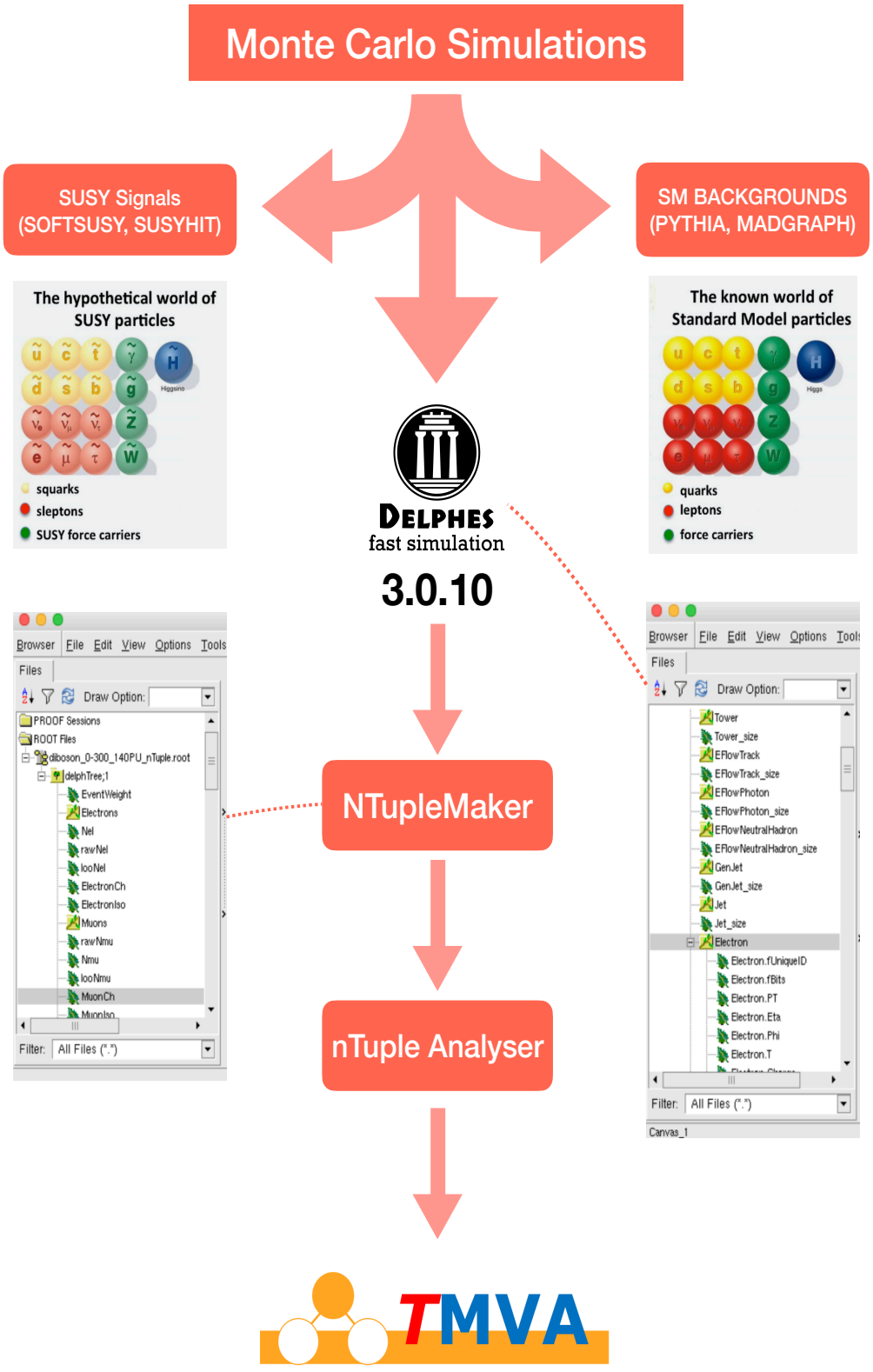


Figure 3.2 : An illustrative summary of the analysis workflow, starting from Monte Carlo signal and background simulations, adding the detector response, ntuplizing the resulting data, finally doing the analysis by employing various techniques.

The SM background to this SUSY search arises mainly from the following processes: $t\tilde{t}$ + jets events, t + jets events, W (V)+jets and Z (V) + jet events, WW (VV)+jets, WZ (VV)+jets and ZZ (VV)+jets events. Between 10-100 million signal events per background were produced for the Snowmass effort [35] with MADGRAPH5 [36]. The SM background has a cross section that is normalised to (NLO), and discussed in great detail in the Snowmass studies. [35], [37], [38]. A maximum of four partons that arise from initial and final radiation are matched to PYTHIA for hadronisation and fragmentation. These signal samples that are used in the analysis are summarized in Table 3.1. ML* corresponds to the given name to each signal, whereas $m_{\tilde{t}}$ is the mass of stop, the SUSY partner of the top quark.

Table 3.1 : Signal sample properties for the analysis. First column implies the names given to signal samples, second and third column describes the stop and LSP masses, final column adresses the cross section.

Signal	$m_{\tilde{t}}$ [GeV]	m_{LSP} [GeV]	σ [pb] (NLO)
ML1	483	300	0.8101815
ML2	566	375	0.32587
ML3	800	450	0.0381673
ML4	1100	450	0.00439055
ML5	1125	25	0.00373237

These samples will be used within the scope of different scenarios. The scenarios will vary with respect to pile-up, and luminosity. Pile-up events occur when a single bunch crossing produces a few seperate events. In this thesis, 50PU and 140PU scenarios are governed. Events with 50PU are produced with 300 fb^{-1} luminosity, where 140PU has the luminosity 3000 fb^{-1}, which corresponds to the High-Luminosity LHC. (HL-LHC)

HL-LHC project focuses on increasing the LHC performance. The increase in performance is anticipated to result in an increased probability for new discoveries after 2025. The main goal is to increase LHC's luminosity by a factor of 10, which is outside the limits of the detector's design value.

One of the most important properties that effects the performance in an accelerator is luminosity. The number of collisions that are likely to occur in a time interval is directly proportional to this value. The chance of observing rare processes increase by

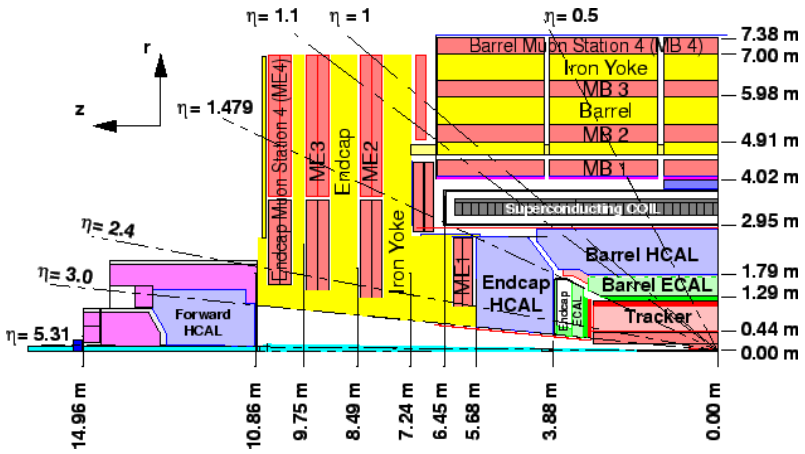


Figure 3.3 : Longitudinal view of one quadrant of CMS [8].

gathering more data; and in order to accomplish that higer luminosity values need to be reached.

HL-LHC is planned to be in working order starting from 2026 and observations in greater detail is an anticipated outcome of this project. Compared with the LHC's 2017 run where 3 million Higgs bosons per year were produced, HL-LHC is expected to produce 15 million.

3.3 Baseline Selection

For a successful analysis, excessive amount of signal samples are produced. Detector response then added to these samples, then they are converted to tree form with the help of DELPHES. This form is consists of independed branches that can be detected and analysed with the ROOT [39] software. In every branch, physical information of an object is stored.

Baseline selections is of service in suppressing the calorimeter noise. The most effective approach in this step is to put certain limits on the transverse momentum and pseudorapidity. Transverse momentum, p_T is the component of momentum perpendicular to the beam line. Momentum along the beamline may include remnants

of the beam particles, however, p_T is always associated with the physics at the vertex. The pseudorapidity η represents the angles with respect to the axis of the colliding beams. It is valued 0 for rays which are perpendicular to the beam, and and large positive or negative values for rays at small angles to the beam [40]. η regions corresponding to the CMS detector can be examined in Fig. 3.3. Both variables are represented in the mathematical framework as

$$p_T = \sqrt{p_x^2 + p_y^2}, \qquad \eta = \ln \tan \frac{\theta}{2}. \tag{3.1}$$

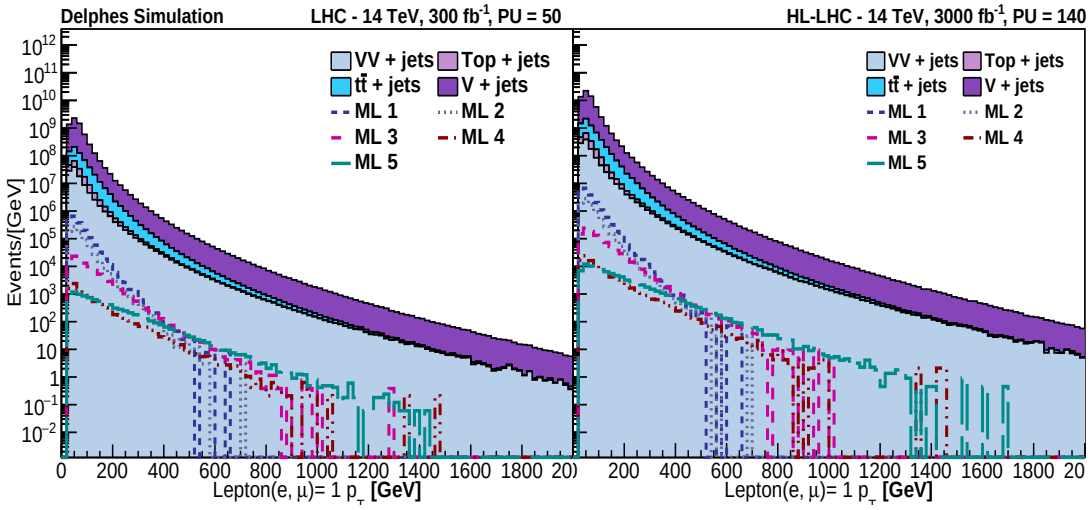
This analysis follows the baseline selections of requiring at least four jets and at least one b-tagged jet with $p_T > 40GeV$ and $|\eta| < 2.4$. Events with more than one lepton are vetoed.

The effect of baseline requirements on signal and background samples can be examined in Figures 3.4, 3.5, 3.6 and 3.7. In every figure, background samples are represented with color filled areas, whereas dashed lines with different colors represent signal samples. In HEP, the convention is to represent samples with hypothetical particles with dashed lines. Y-axis of the plots describe the number of events of the shown samples, while x-axis illustrate the value of the distribution shown.

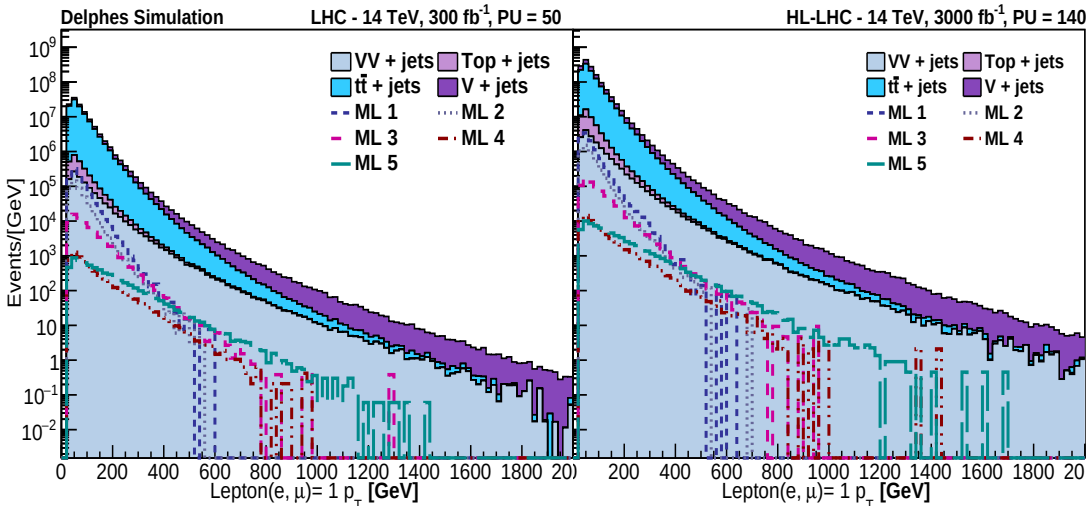
A count of the weighted signal and background events left was made, after applying the baseline requirements. As a result of these requirements, it can be seen that V+jets events are suppressed the most compared to other types of backgrounds, whereas t+jets events are effected the least. Details can be studied from Table 3.2. What is anticipated from the baseline selections is rigorous supression in the background while keeping the signal events as much as possible. A comparison for the SM background versus all signal samples can be seen from Table 3.3

3.4 Signal Selection

Events that pass the baseline requirements are subjected to signal selection. Majority of background suppression takes place in this step. A variety of special variables that are deducted from event kinematics are used in controlling the background, such as



(a)



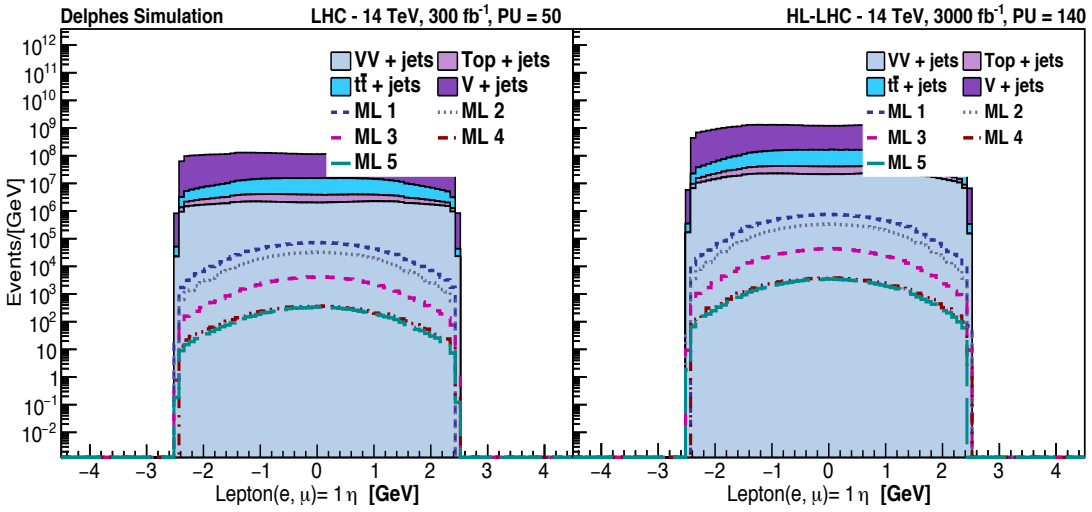
(b)

Figure 3.4 : p_T of the leptons before (a) and after (b) baseline requirements. 50PU events with 300 fb^{-1} are displayed on the left, 140PU with 3000 fb^{-1} is on the right. The full histograms describing the backgrounds are stacked, and the five inclusive signal models are shown as dotted lines, not stacked.

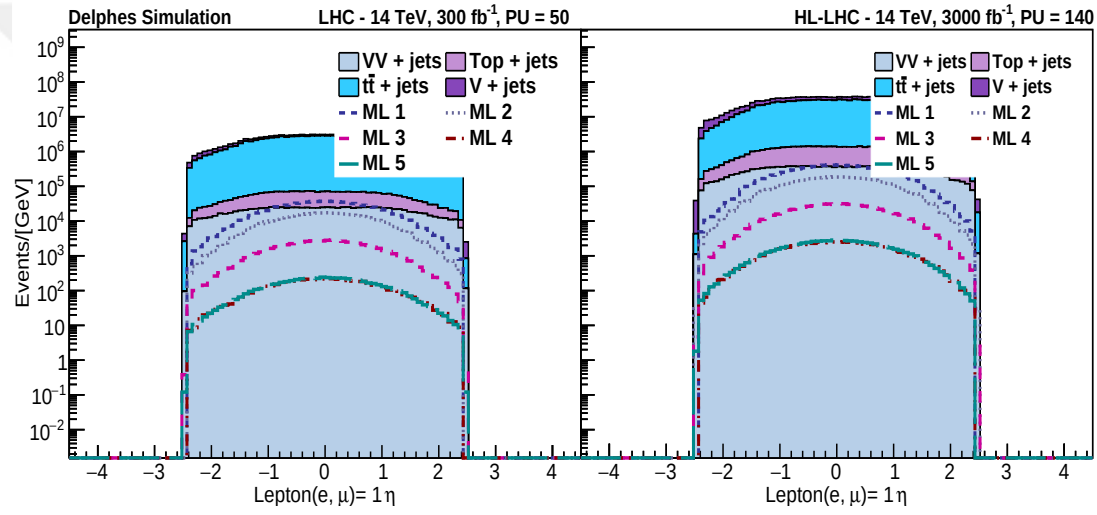
the scalar sum of the jets H_T, missing transverse energy E_T, transverse mass M_T are among the most commonly used.

3.4.1 Missing transverse energy (E_T^{miss})

Missing transverse energy, E_T^{miss} is the magnitude of all invisible transverse momentum arising from invisible particles within detector that do not interact with matter strongly enough to be detected. If the vectoral sum of the transverse momenta emerging from the observed particles in the event is not zero, from conservation of transverse



(a)



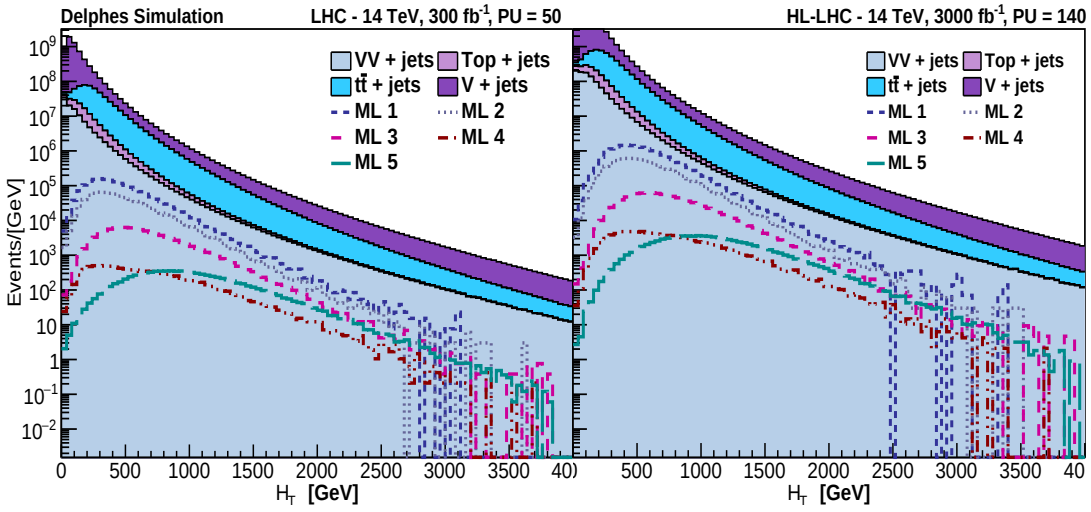
(b)

Figure 3.5 : η of the leptons before (a) and after (b) baseline requirements. 50PU events with 300 fb^{-1} are displayed on the left, 140PU with 3000 fb^{-1} is on the right. The full histograms describing the backgrounds are stacked, and the five inclusive signal models are shown as dotted lines, not stacked.

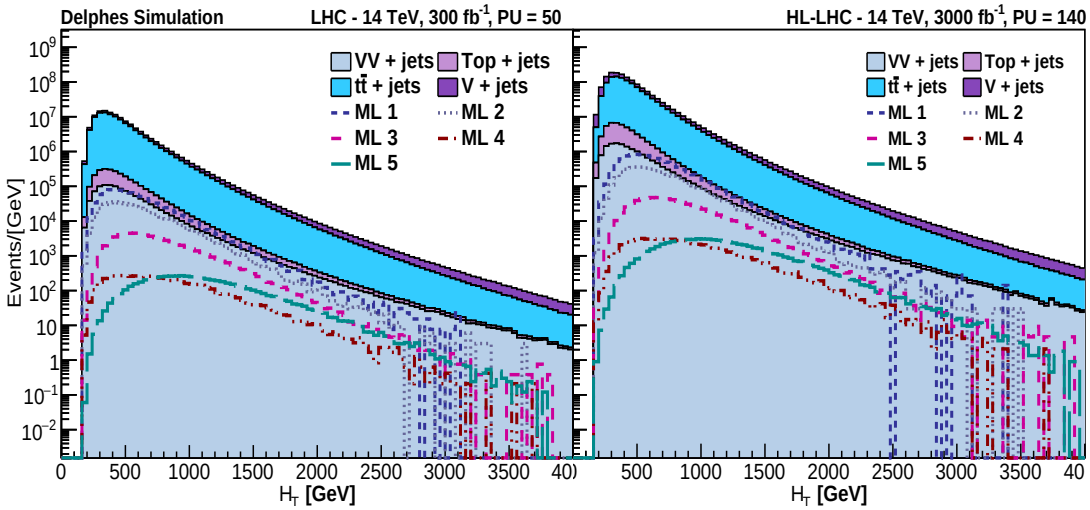
momentum that implies "invisible" particles with "invisible" momenta to cancel this remnant. It can be calculated as $E_T^{miss} = -|\sum \vec{p}_T|$.

3.4.2 Transverse mass M_T

The transverse mass component that is perpendicular to the beam direction is called m_T. It is a variable that is invariant under Lorentz boost, in the z-axis that is along



(a)



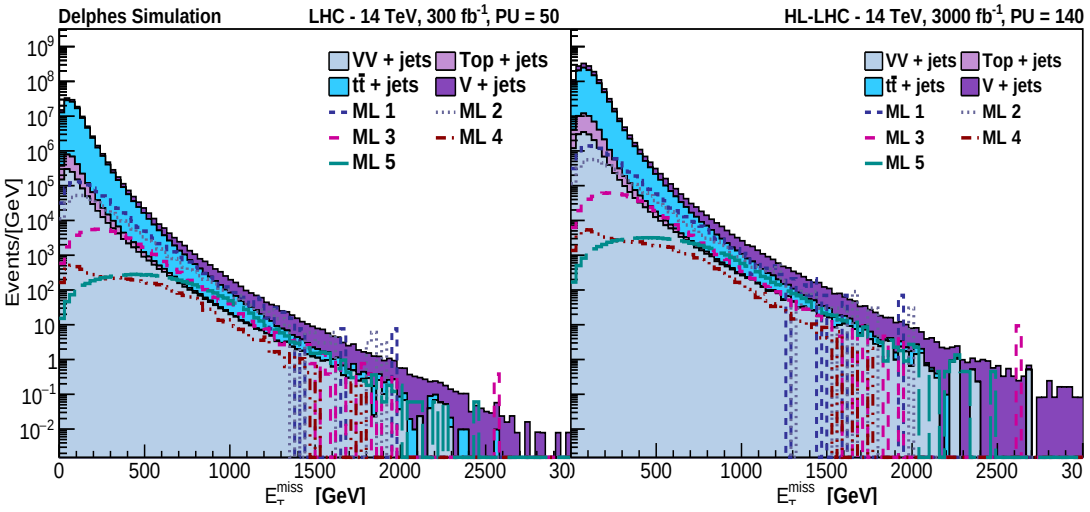
(b)

Figure 3.6 : Scalar sum of jets H_T before (a) and after (b) baseline requirements. 50PU events with 300 fb^{-1} are displayed on the left, 140PU with 3000 fb^{-1} is on the right. The full histograms describing the backgrounds are stacked, and the five inclusive signal models are shown as dotted lines, not stacked.

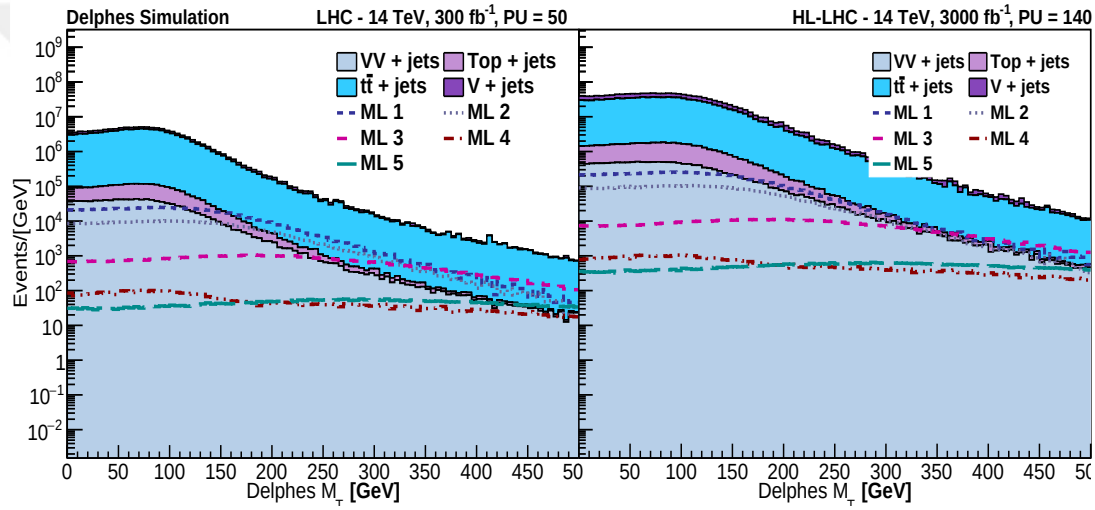
the beam pipe. M_T helps in the leptonically decaying W boson events. m_T is mathematically described in natural units ($\hbar = c = 1$) as

$$m_T^2 = m^2 + p_x^2 + p_y^2 + p_z^2 \tag{3.2}$$

where m is the invariant mass, p_x and p_y are the transverse components of momenta, that are perpendicular to the beam pipe. In this work, M_T is represented as



(a)



(b)

Figure 3.7 : (a) E_T^{miss} and (b) M_T after baseline requirements. 50PU events with 300 fb^{-1} are displayed on the left, 140PU with 3000 fb^{-1} is on the right. The full histograms describing the backgrounds are stacked, and the five inclusive signal models are shown as dotted lines, not stacked.

$$M_T = \sqrt{2p_T^{lep}E_T^{miss} - 2\vec{p_T^{lep}}\vec{E_T^{miss}}} \tag{3.3}$$

where p_T^{lep} is the momentum of leptons, and E_T^{miss} is the missing transverse energy.

3.4.3 M_{T2}^W

M_{T2}^W is a variable introduced by Yang Bai et al in 2012 [41]. It is a kinematic variable defined specifically for single lepton search channel, which this thesis mainly focuses on. The main background in top squark direct production emerges from the SM

Table 3.2 : Summary of the individual background and resulting total SM background events after the baseline requirements, for 50PU at 300 fb^{-1}.

Cutflow	VV+jets	$t\bar{t}$+jets	V+jets	Top+jets	**Total SM**
#Events	11082100000	21569500000	1684020000000	6206650000	1722878250000
Lepton(e,μ)	1042020000	4600470000	55007300000	771728000	61421518000
nJets ≥ 4	51045400	1191630000	1441450000	45384500	2729509900
bjets ≥ 2	15617400	1036150000	336086000	36346400	1424199800

Table 3.3 : Summary of individual SUSY signal events left after the baseline requirements, for 50PU at 300 fb^{-1}.. The total SM background is added for comparison.

Cutflow	Total SM	ML1	ML2	ML3	ML4	ML5
#Events	1722878250000	24305400	9776100	1145020	131716	111971
Lepton(e,μ)	61421518000	5680390	2301950	290616	31899	27442
nJets ≥ 4	2729509900	4248170	1753300	248609	25672	24364
bjets ≥ 2	1424199800	3032490	1271970	209584	21128	22281

$t\bar{t}$ events, and this variable substantially supresses the background by working with dileptonic and semileptonic topologies.

M_{T2}^{W} is defined as the as the minimum 'mother' particle mass compatible with all the transverse momenta and mass-shell constraints.

It is mathematicaly defined as:

$$M_{T2}^{W} = \min \left\{ m_y \text{ consistent w/: } \left[\begin{array}{r} \vec{p}_1^{\,T} + \vec{p}_2^{\,T} = \vec{E}_T^{miss}, p_1^2 = 0, (p_1 + p_l)^2 = p_2^2 = M_W^2, \\ (p_1 + p_l + p_{b_1})^2 = (p_2 + p_{b_2})^2 = m_y^2 \end{array} \right] \right\}. \tag{3.4}$$

3.4.4 Topness

As an alternative to M_T^{2W}, *topness* variable can also be employed in order to suppress the top background, as defined in [42]. The suitable selection for topness variable is determined after multiple studies. The most optimal value at the rough estimate was to accept events that has $Topness < 7$. Similar studies were used as a guide in this decision. However, the future works will involve further optimizations by employing algorithmical checks.

3.4.5 $\alpha_{M_{T2}^{W}}$

In this thesis, we have also defined another alternative kinematic variable called $\alpha_{M_{T2}^{W}}$,

$$\alpha_{M_{T2}^W} = \frac{\sqrt{H_T + E_T^{miss}}}{M_{T2}^W}. \tag{3.5}$$

Every fundamental variable used in analyses give information about a physics object. The combination of M_T, H_T, and E_T^{miss} that are explained in previous subsections are being used to derive new SUSY variables. Distributions that are comparable to each other, where the signal and the background show similar behavior are in general preferrable. The reason behind this is the way the software TMVA works smoother with this type of inclusive distributions.

In $\alpha_{M_{T2}^W}$, the numerator $\sqrt{H_T + E_T^{miss}}$ and the denominator M_{T2}^W provide a comparable distribution of this type, where signal and background have similar values. Other possible combinations of H_T, E_T^{miss} and M_{T2}^W that do not behave accordingly are rejected during the studies to derive $\alpha_{M_{T2}^W}$. The aim was to find a combination that behaves anological with M_{T2}^W at the denominator.

$\alpha_{M_{T2}^W}$ at different scenarios is shown at Figure 3.8, where the afforementioned inclusive distribution is clearly observed between signal and background events.

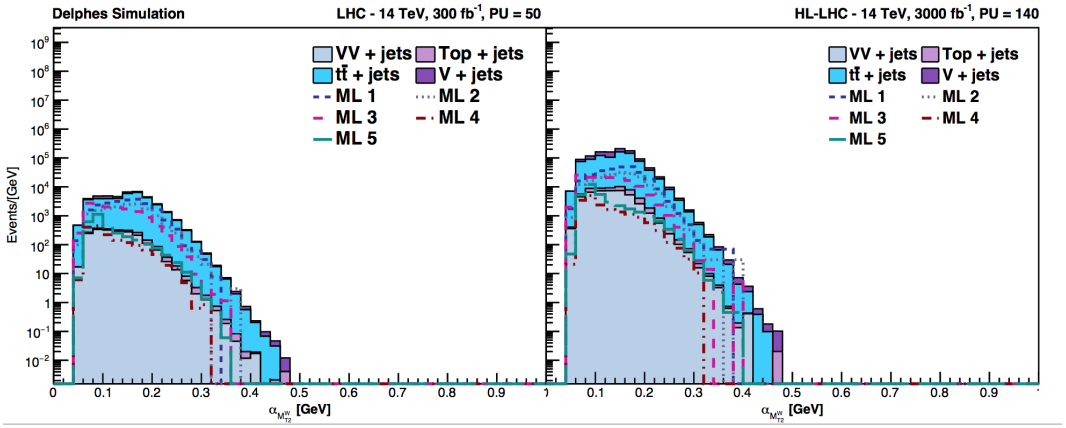


Figure 3.8 : $\alpha_{M_{T2}^W}$ for 50PU, 300 fb^{-1} on the left and 140PU, 3000 fb^{-1} on the right.

Certain values that are determined over these kinematic variables make up the signal selection. Events with $E_T^{miss} > 120$ GeV, $M_T > 120$ GeV and $H_T > 500$ GeV were accepted. For M_{T2}^w, topness, and $\alpha_{M_{T2}^w}$ variables, requirements are applied interchangeably, by picking one variable at a time.

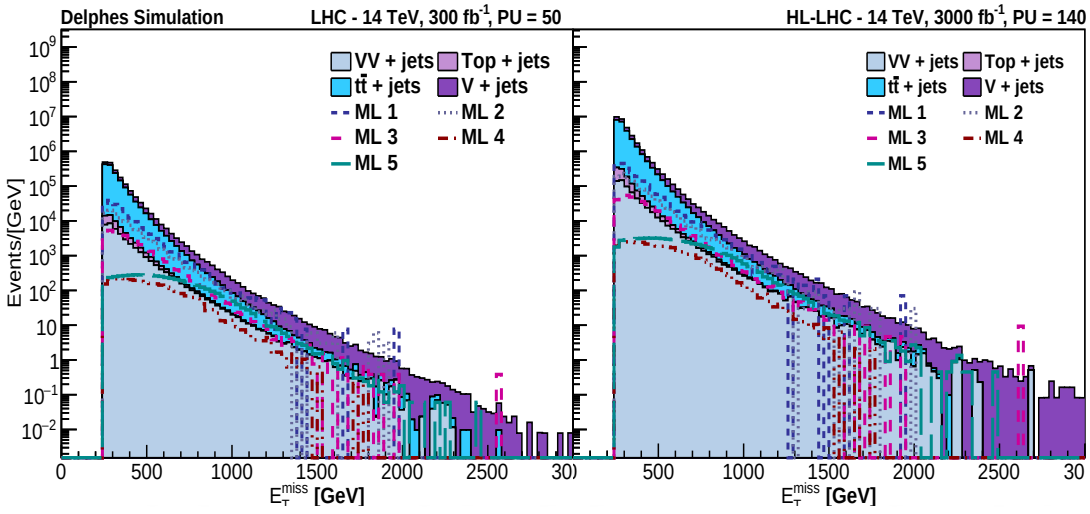
Resulting event numbers after each step can be examined in detail in the results section, in Table 5 for 50PU and Table 7 for 140PU.

3.4.6 Variable comparison

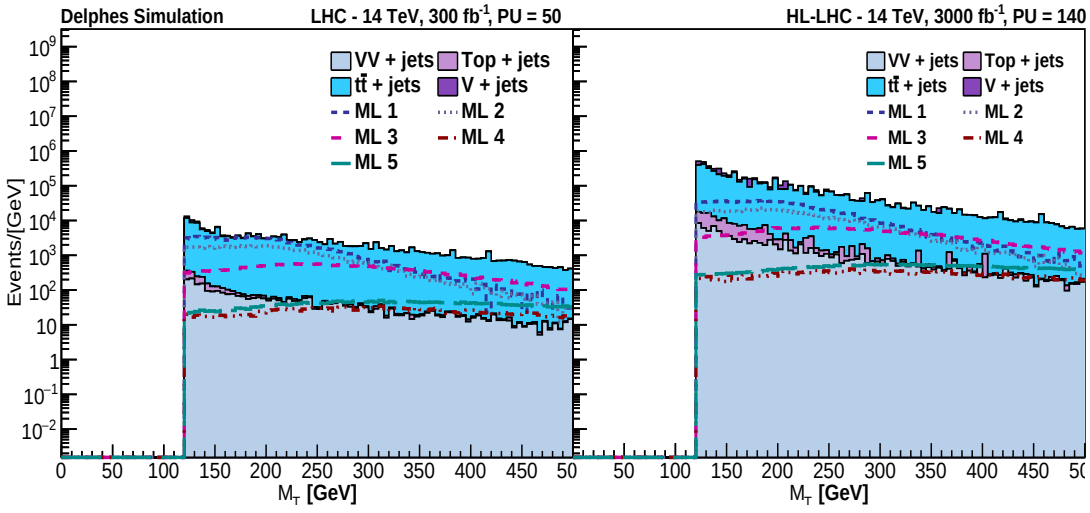
The cuts applied on the interchangeable variables M_{T2}^{W}, $\alpha_{M_{T2}^{W}}$ and topness effect rest of the training variables H_T, M_T and E_T^{miss} differently.

In Figures 3.11 and 3.12, these distributions are shown for ML1 and ML5 for different interchangeable cuts.

For ML1, all variables are the most effected by $\alpha_{M_{T2}^{W}}$ and M_{T2}^{W} similarly, in both pileup scenarios. It is on the contrary for ML5, where topness supresses the signal more.

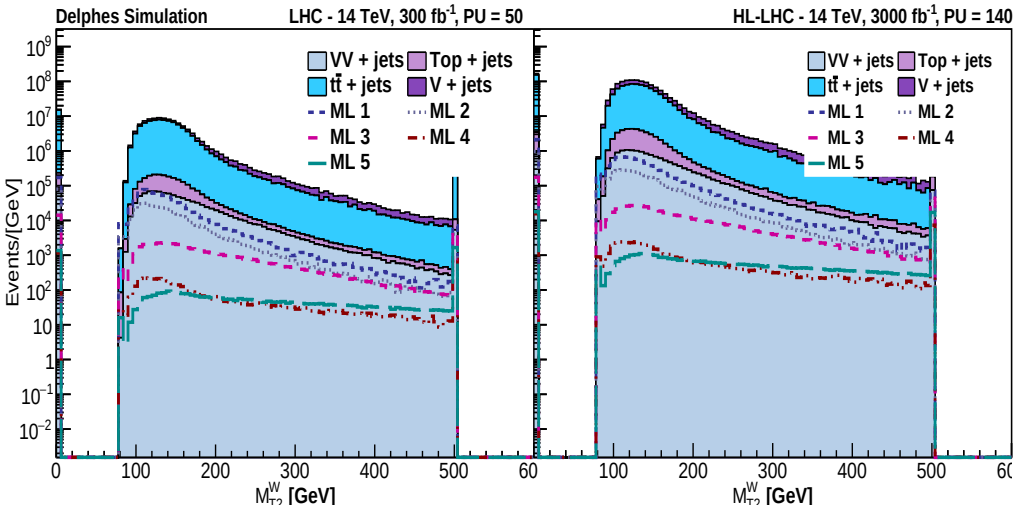


(a)

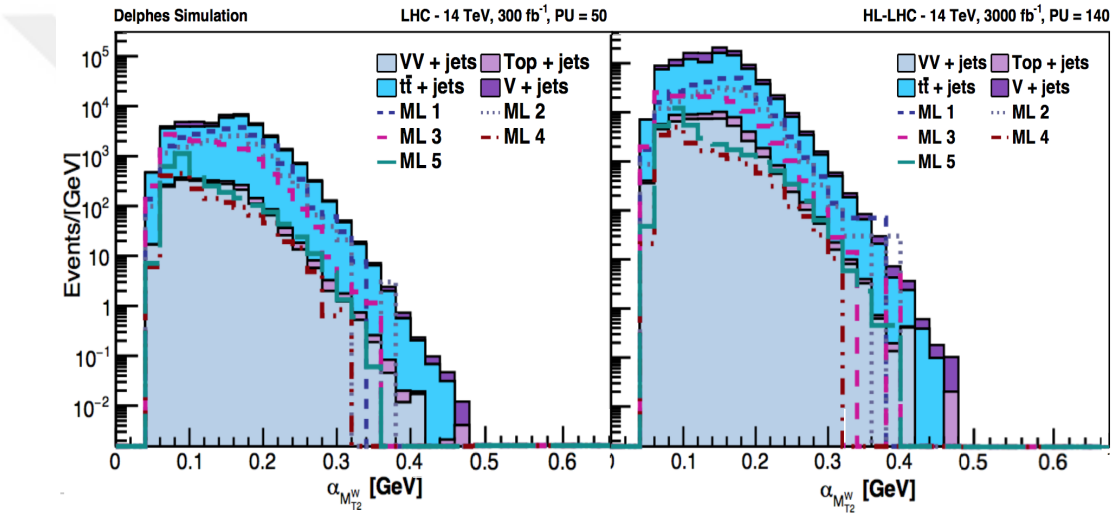


(b)

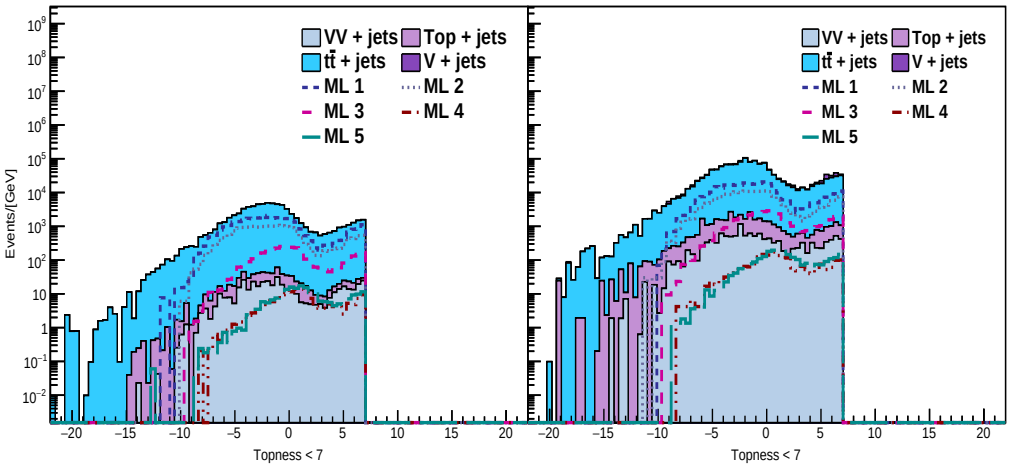
Figure 3.9 : Missing transverse energy E_T^{miss} (a), and transverse mass m_T (b) distributions after all cuts. 50PU, 300fb^{-1} events are displayed on the left, whereas on the right is events with 140PU, 3000fb^{-1}. The full histograms describing the backgrounds are stacked, and the five inclusive signal models are shown as dotted lines, not stacked.



(a) M^w_{T2},

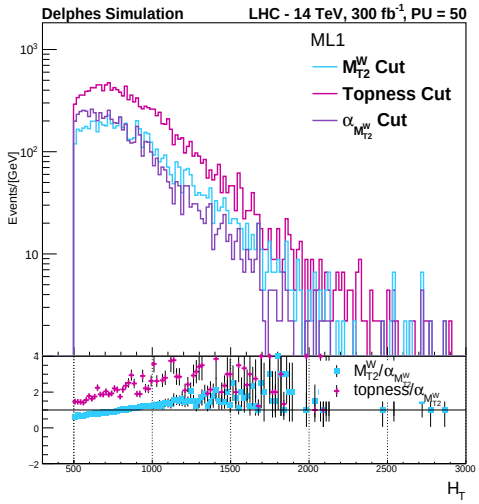


(b) $\alpha_{M^W_{T2}}$

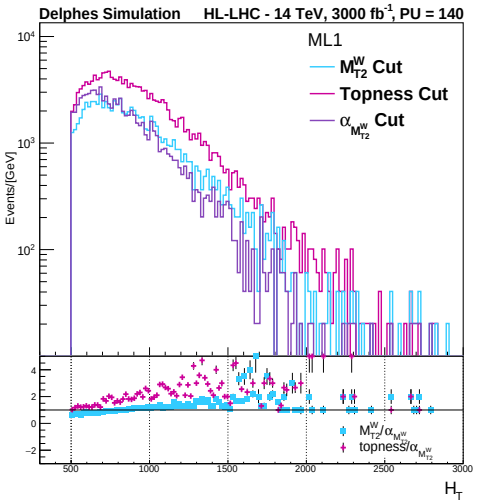


(c) *Topness*

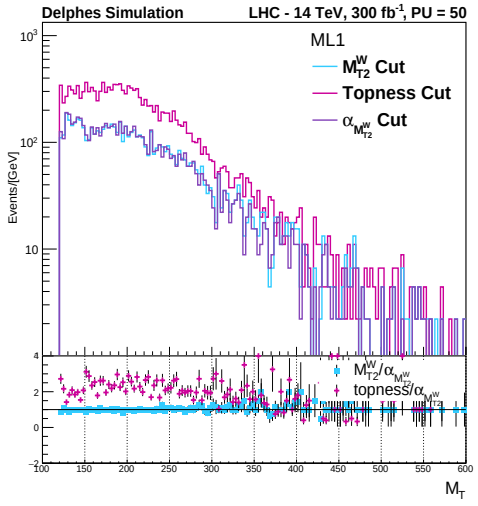
Figure 3.10 : MVA variables. 50PU, $300 fb^{-1}$ events are displayed on the left, whereas on the right is events with 140PU, $3000 fb^{-1}$. The full histograms describing the backgrounds are stacked, and the five inclusive signal models are shown as dotted lines, not stacked.



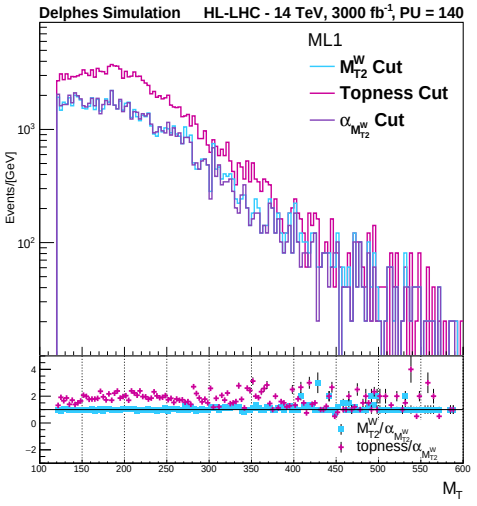
(a) H_T, LHC



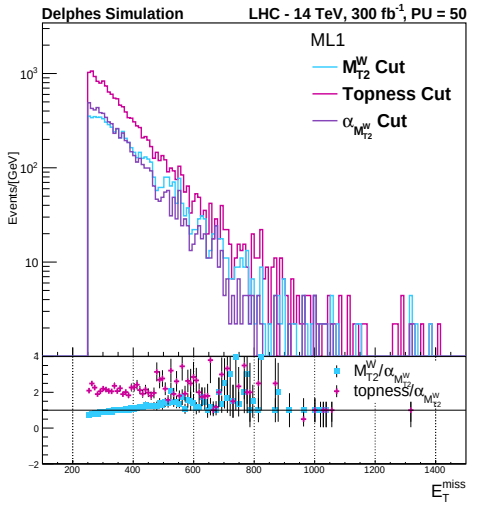
(b) H_T, HL-LHC



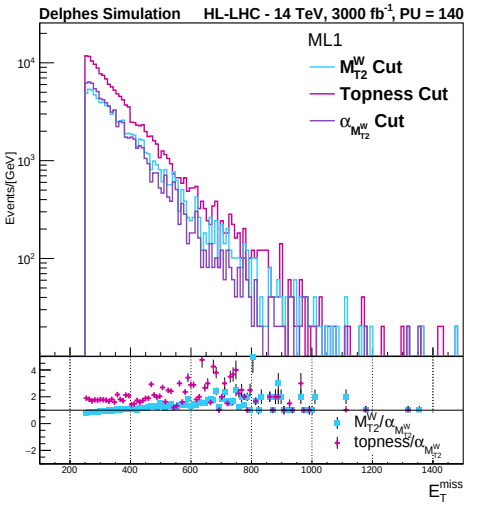
(c) M_T, LHC



(d) M_T, HL-LHC

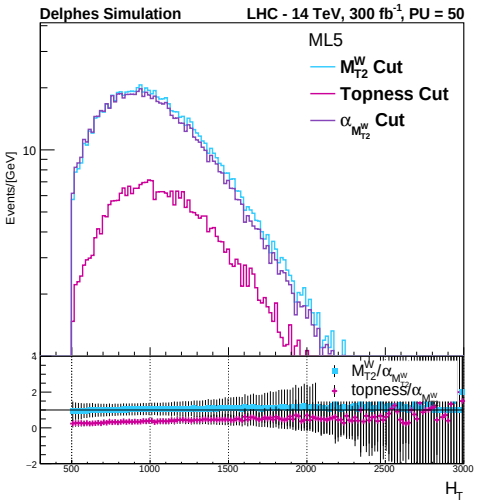


(e) E_T^{miss}, LHC

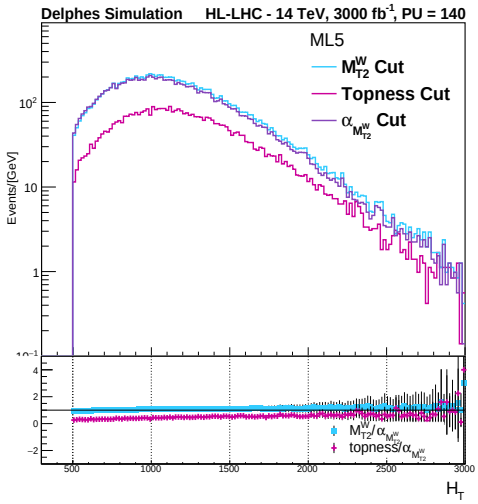


(f) E_T^{miss}, HL-LHC

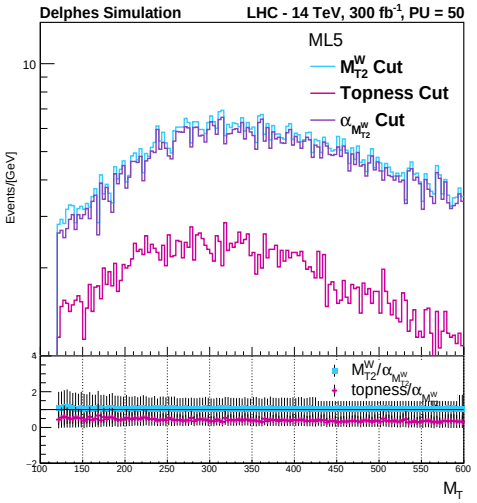
Figure 3.11 : The effect of the interchangeable variable selections on training variables, for signal ML1 at 50PU, 300 fb^{-1} and 140PU, 3000 fb^{-1}.



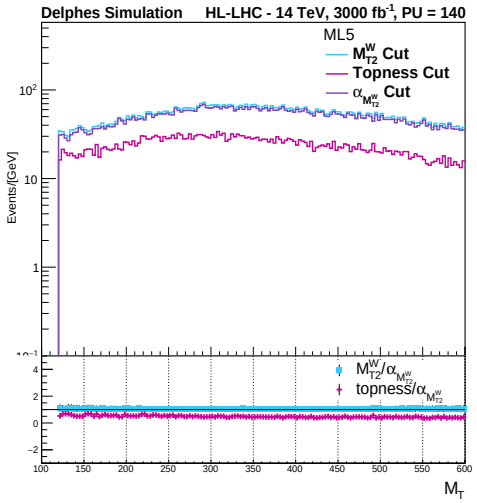
(a) H_T, LHC



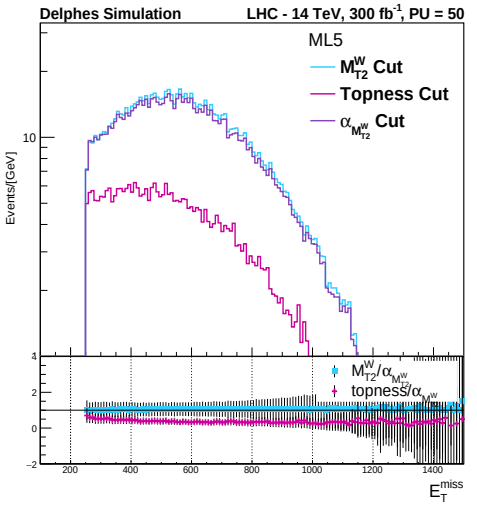
(b) H_T, HL-LHC



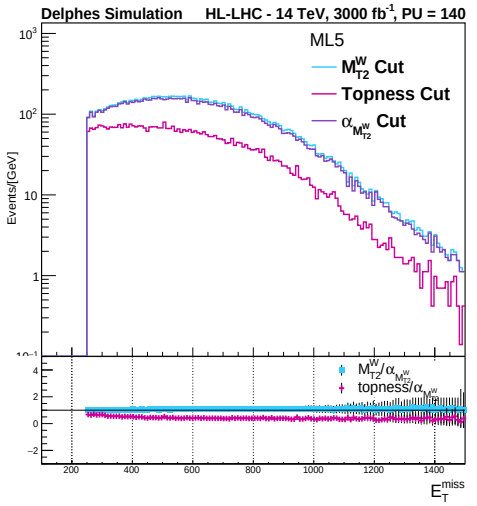
(c) M_T, LHC



(d) M_T, HL-LHC



(e) E_T^{miss}, LHC



(f) E_T^{miss}, HL-LHC

Figure 3.12 : The effect of the interchangeable variable selections on training variables, for signal ML2 at 50PU, 300 fb^{-1} and 140PU, 3000 fb^{-1}.

What is expected from a variable in the ideal case is to have it supress the background by leaving the signal events untouched as much as possible. It can be deduced from Figure 3.11 that M_{T2}^{W} and $\alpha_{M_{T2}^{W}}$ cuts result in different behaviors on the H_T and E_T^{miss} distributions both at LHC and HL-LHC for ML1. M_{T2}^{W} cut starts gaining on $\alpha_{M_{T2}^{W}}$ starting from the beginning. Whereas the M_T give a rather flat behavior at the $M_{T2}^{W}/\alpha_{M_{T2}^{W}}$ cuts ratio, except after 400-450 GeV where statistics are lower.

This observations are not likewise for ML5, shown in 3.12. M_{T2}^{W} and $\alpha_{M_{T2}^{W}}$ cuts affect all distributions both in LHC and HL-LHC similarly.

$t\bar{t} + Jets$ background is suppressed the least with topness cuts, such can be seen in Figure 21 in all three distributions. $\alpha_{M_{T2}^{W}}$ and M_{T2}^{W} show a similar efficiency, where the first works better at H_T distribution and the latter at M_T. Again, likewise signal samples, $\alpha_{M_{T2}^{W}}$ and M_{T2}^{W} shows similar effect on M_T. However in H_T and M_{T2}^{W}, there is a difference in higher energies where $\alpha_{M_{T2}^{W}}$ performs better than M_{T2}^{W}.

Results of samples ML2, ML3 and ML4 is provided in Figures 15, 16 and 17, rest of the background samples are provided in Figures 18, 19 and 20 in the Appendix A.3.

3.5 TMVA

Considering the ROOT-integrated environment it provides, the Toolkit for Multivariate Analysis (TMVA) [9] has been a very useful tool in this analysis. It uses supervised learning teqhniques, in which the multivariate method is trained over a sample where the result is known. (e.g. Monte-Carlo simulation of signal and background).

TMVA offers various multivariate methods that is a set of statistical analysis methods that simultaneously analyze multiple measurements on the object studied. Commonly used methods are as Likelihood, Fisher Method, NN, BDT, SVM... These methods vary in level of complexity, background rejection performance, training speed, robustness against number of input variables and against overtraining. Training and test samples are used in every MVA method in order to make sure the training assembles a reasonable solution that does not depend on the statistical fluctuations.

Neural Networks is amongst the methods that are widely used. It is made up of neurons that takes the linear combination of previous neuron outputs, as the input. During training, the weights of each neurons are found by minimizing the error on NN output.

NN takes advantage of correlations, and is robust against overtraining and increased number of variables.

Another method is the desicion trees, which essentially is a binary tree that composed of cuts that paves the phase space of input variables. For every event, YES/NO desicions are repeated before reaching a stop criterion. A colossal drawback of this method is the extreme sensitivity of training samples, likely to yield to overtraining. Different techniques can be employed for performance stabilization, such as boosting, bagging, or random forests.

In this analysis we have used BDT with the following settings: boosting with AdaBoost, 850 trees in the forest, minimum node size 5% for classification, maximum depth of decision trees as three, 20 grid points in variable range used in finding optimal cut in node splitting and separating with GiniIndex.

3.5.1 Boosted desicion trees

A decision tree is a classifier that is structured in binary. It takes a stop criterion for the process to be completed, and yes/no decisions are applied repeatedly till the criterion is reached. These decisions are made on one variable at a time, and the number of decisions is controlled with an internal parameter called *node*. Signal and background are classified from the regions of the splitted phase space by looking at the majority of the events ending up at the final node. Example of a simple decision tree can be examined in Figure 3.13.

In the Figure 3.13, the binary decisions that are applied from the nodes can be observed, starting from the root node. The decisions are applied via discriminating variables such as x_i in the example. Each desicion is made so the best separation among signal and background is established. Not all variables are used necessarily, some of them may be used repeatedly and some of them not at all, depending the efficiency they provide.

Decision trees provide straightforward interpretation with the way they are visualized in a 2D tree format. With this method, the phase space can be splitted into larger number of hypercubes compared to other methods. These hypercubes are labeled as signal or background-like. Each route travelled from the root node to the final leaf represent a sequence of cuts applied for separation.

Without boosting, decision trees are not stable enough. Instabilities arise from the statistical fluctuations of the training sample. A fluctuation may impact only one of the two variables in a two input variable scenario, resulting in a significantly different classifier response. This issue is solved with various boosting methods.

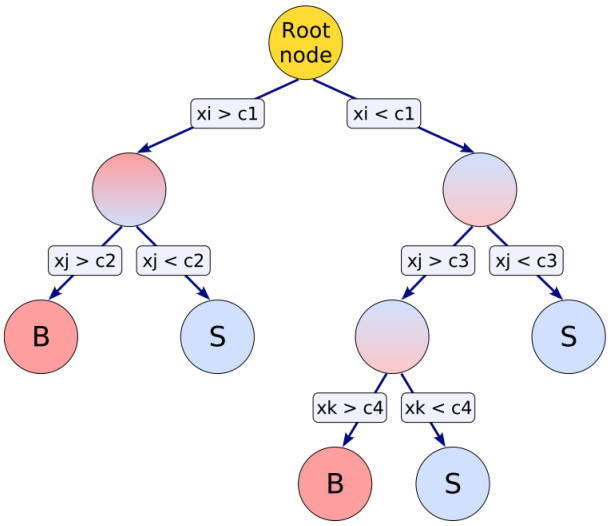


Figure 3.13 : Schematic description of a decision tree. x_i, x_j and x_k are the discriminators, whereas B symbolizes background and S symbolizes signal.

A forest can be constructed from many trees that are derived from the same training sample. Final classification is done by considering the majority vote coming from each separate classification in the trees of the forest. This so called process of boosting modifies the weights in the samples. With each step, future trees have better learning. In this thesis Adaptive Boosting (AdaBoost) technique will be used, where the final weight is the sum of all classiffiers that are weighted by their errors.

AdaBoost is the most popular boosting algorithm. The algorithm reweights the misclassified and classified events, while a higher weight was given to the misclassified events. The first tree is trained with the original event weights, whereas the next trees use modified event samples with previously misclassified events' weight is multiplied by a *boost weight* denoted as α. This weight is determined from the misclassification rate of the tree before it,

$$\alpha = \frac{1 - err}{err}, \tag{3.6}$$

where err is the so called misclassification rate. In order to keep the sum of weights constant, the weights of the event sample are renormalised afterwards.

The result of each individual classifier is defined as $h(\mathbf{x})$ where x indicated the input variables. The classifier for signal is described as $h(\mathbf{x}) = +1$ and for background it is $h(\mathbf{x}) = -1$. $y_{Boost}(\mathbf{x})$ defines the boosted event classifier and is calculated as

$$y_{Boost}(\mathbf{x}) = \frac{1}{N_{collection}} \cdot \sum_{i}^{N_{collection}} \ln(\alpha_i) \cdot h_i(\mathbf{x}). \tag{3.7}$$

When $y_{Boost}(\mathbf{x})$ has a small value, that implies a background-like event. Larger values then imply a signal-like event. Eq. (3.7) shows the standard boosting algorithm.

A weak classifier could be defined as small decision trees with a depth of 2 or 3, with little discrimination power. AdaBoost works very well with these weak classifiers. The performance could be improved by imposing *slow learning* by having more steps of boosting. β parameter inhibits the learning rate, and it is given as an exponent to the boost weight α as $\alpha \to \alpha^{\beta}$. It can be modified within TMVA.

The results are visualized and interpreted via TMVA's graphical user interface (GUI). The GUI interface can be examined in Figure 3.16.

3.5.2 Multilayer perceptron (MLP)

Artificial neural networks (ANN) differ from each other via neuran layouts, and connection weights. Technically, a NN that consists of n neurons is able to have n^2 connections. However, this system can be simplified with a new layout; in which the neurons are aligned in layers, and connected directly with the previous and following layers only. Such networks are called the multilayer perceptrons, as illustrated in Figure 3.14.

First and last layers of the perceptron are named input and output respectively. Every other layer in between is called hidden layers. For every variable in a classification process, there is a neuron in the input layers. Clearly, n_{var} neurons for n_{var} variables. These neurons contain the the input values $x_1, ...x_{n_{var}}$. The single neuron at the final layer that contains the output variable is called y_{ANN}, the NN estimator. In this thesis training are done with a similar structure. 4 input variables H_T, M_T, E_T^{miss} + 1 interchangeable variable make up the input layer.

There is 1 hidden layer, and 1 output layer, as illustrated in Fig. 3.15

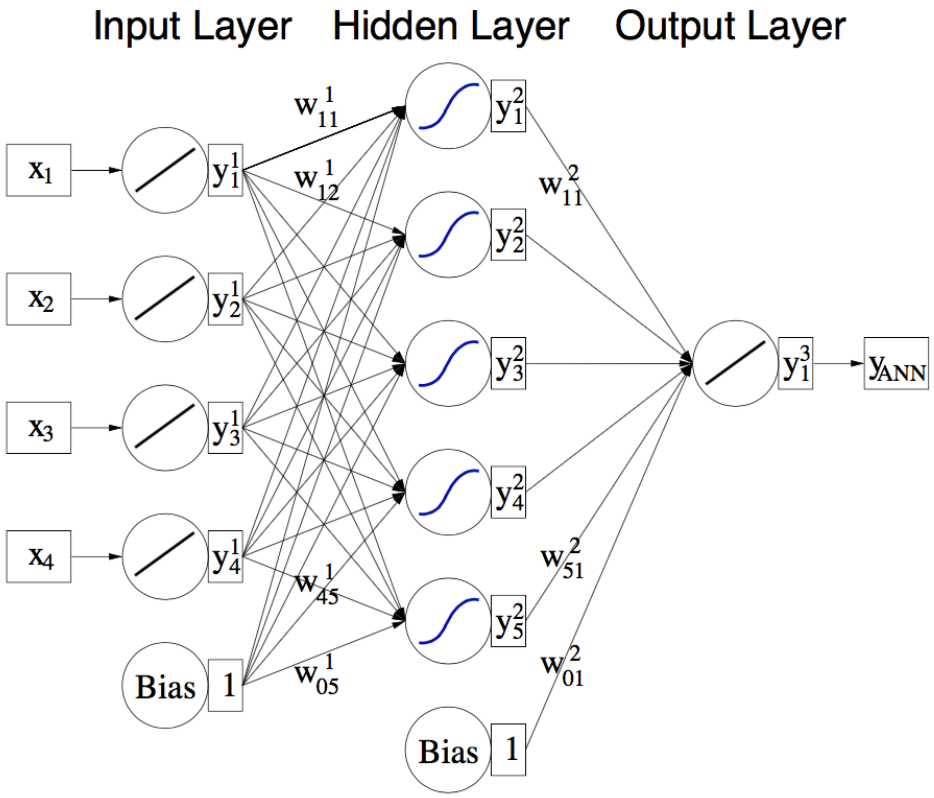


Figure 3.14 : Multilayer perceptron structure [9].

3.5.3 Deep neural networks (DNN)

Deep neural network (DNN) is a type of artificial neural network that is structured with more than one hidden layers, and quite a few number of neurons in these layers. Complex and non-linear relations can be observed with this network. It can work efficiently with training data that is very large.

The aim of its implementation is to provide a feed-forward MLP that can be trained on multi-core and GPU architectures. There are couple of backend methods. The standard backend that can work on any platform which containts ROOT software. CPU backend employs multithreading to do parallel training on multicore CPU structures. The GPU backend is used in training on CUDA-capable GPU structures.

To train the network, stochastic batch gradient descend method is used. The weight of each training step is W_{ij}^{k}, and bias term is θ_{i}^{k} where k is the layer number.

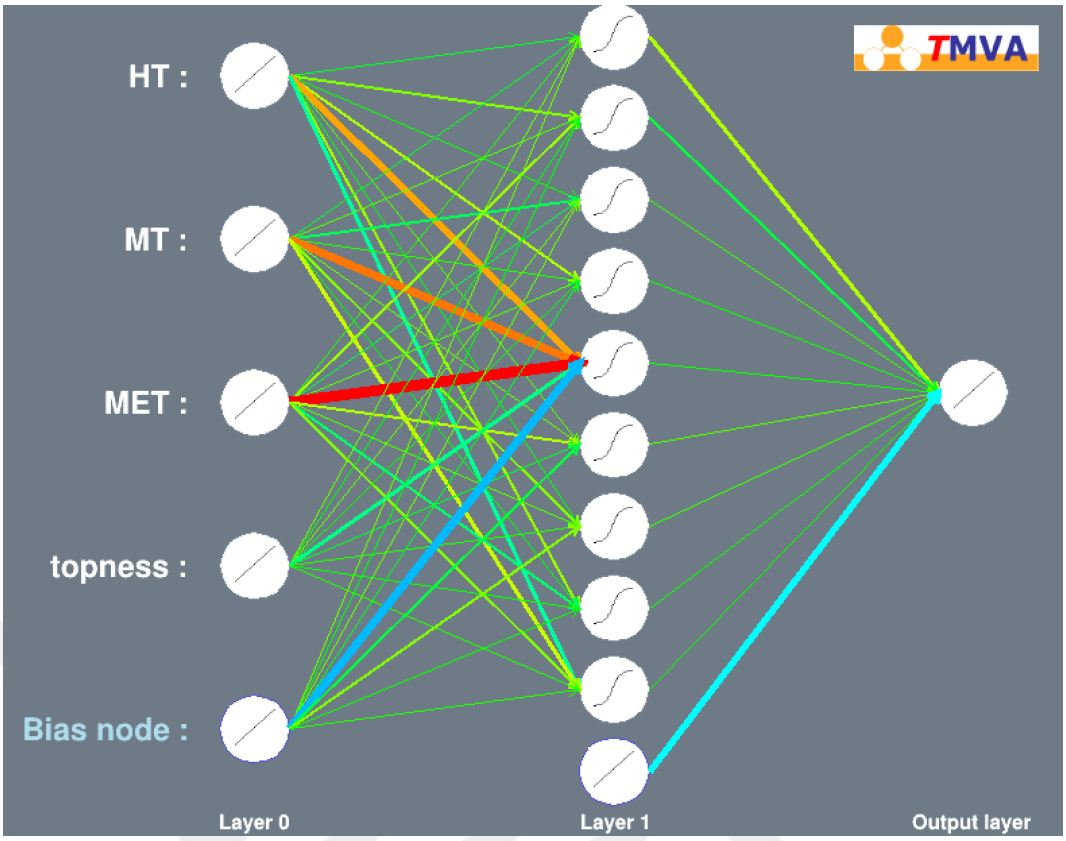


Figure 3.15 : NN structure for ML1 with topness variable at 50PU, $300 fb^{-1}$.

$$W_{ij}^k \to W_{ij}^k - \alpha \frac{\partial J(x_b, y_7 b)}{\partial W_{ij}^k} \tag{3.8}$$

$$\theta_i^k \to \theta_i^k - \alpha \frac{\partial J(x_b, y_b)}{\partial \theta_i^k}, \tag{3.9}$$

where x_b is the randomly chosen input batch, and y_b is the expected output. The function $J(x_b, y_b)$ is the corresponding value of the loss function. With this model, it is also possible to do trainings by using momentum.

$$\Delta W_{ij}^k \to p \Delta W_{ij}^k - (1.0 - p) \frac{\partial J(x_b, y_7 b)}{\partial W_{ij}^k} \tag{3.10}$$

$$W_{ij}^k \to W_{ij}^k - \alpha \Delta W_{ij}^k \tag{3.11}$$

$$\Delta \theta_i^k \to p \Delta \theta_i^k - (1.0 - p) \frac{\partial J(x_b, y_b)}{\partial \theta_i^k} \tag{3.12}$$

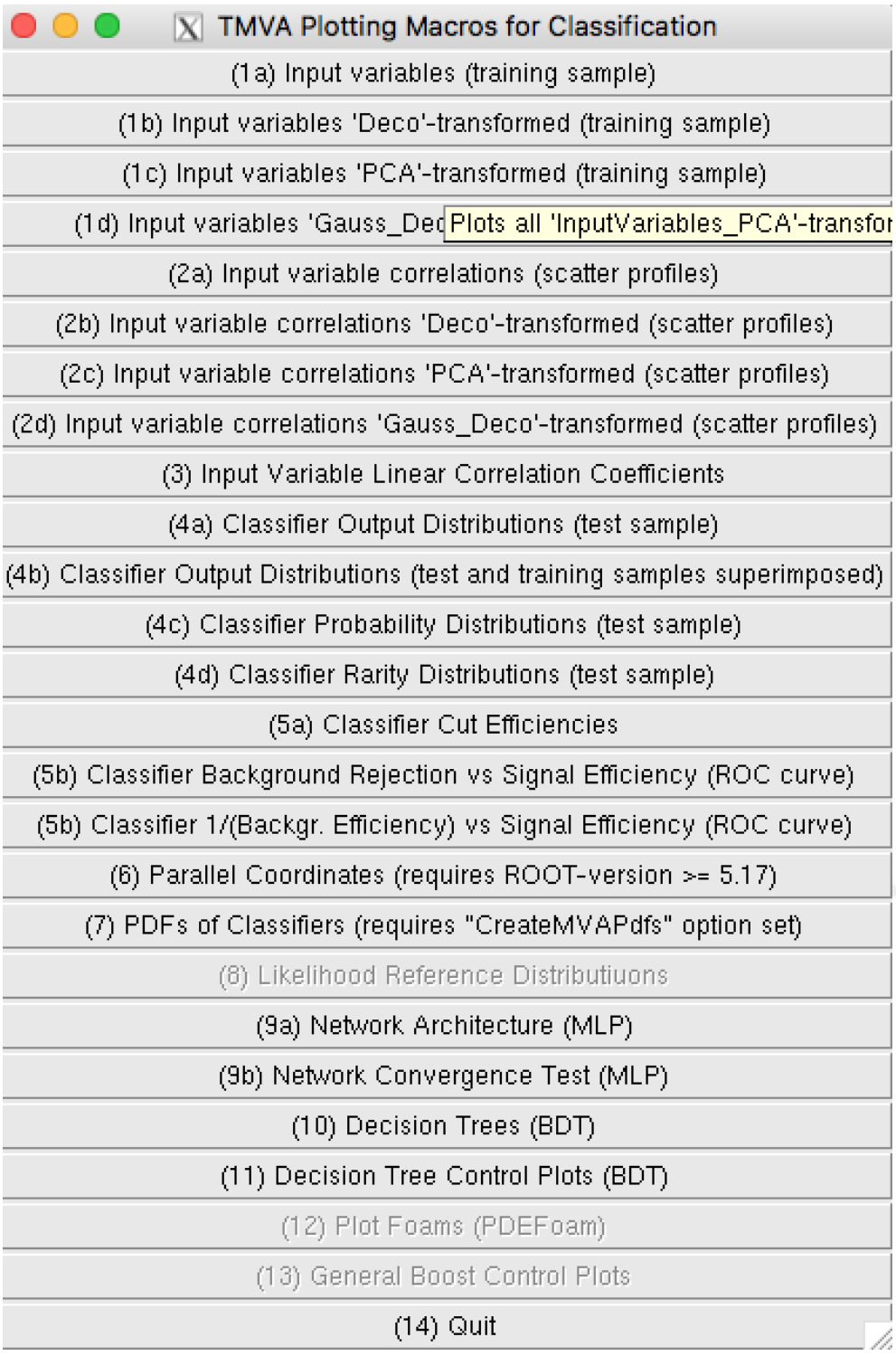


Figure 3.16 : MVA's graphical user interface.

$$\theta_i^k \rightarrow \theta_i^k - \alpha \Delta \theta_i^k \tag{3.13}$$

The standard stochastic batch gradient is derived by setting $p = 0$. Training batches are randomly chosen. Multiple steps are needed to complete the training set, and enough steps to traverse once is called the epoch.

4. RESULTS AND INTERPRETATION

SUSY samples and the SM backgrounds go through baseline and signal selections. Events left after each selection are counted, and they are filled in new trees to feed to TMVA. The selections are done for different scenarios with varying pileup and luminosity values. 50PU with 300 fb^{-1} and 140PU with 3000 fb^{-1} are used. Increasing the luminosity also increases the statistics. In total, there is 5 SUSY samples (ML1, ML2, ML3, ML4, ML5) and 4 SM backgrounds (V+jets, VV+jets, t+jets, $t\bar{t}$+jets).

Selecting SUSY signals among large SM backgrounds is a tedious job. TMVA methods are used to make the separation process rather simpler. Classification methods were used with TMVA's BDT, MLP, and DNN techniques. Selected variables for training are M_{T2}^{W}, $topness$, and $\alpha_{M_{T2}^{W}}$. The TMVA input trees are structured as a TTree with TreeS (signal tree) and TreeB (background tree).

Number of signal and background events left are counted after each selection. It is important to compare number of total SM background left with final SUSY signals events. Ideal scenario is to have as little SM background as possible, whereas having high number of SUSY signals. It is intended to have selections that suppress the background while not decreasing the signal count drastically.

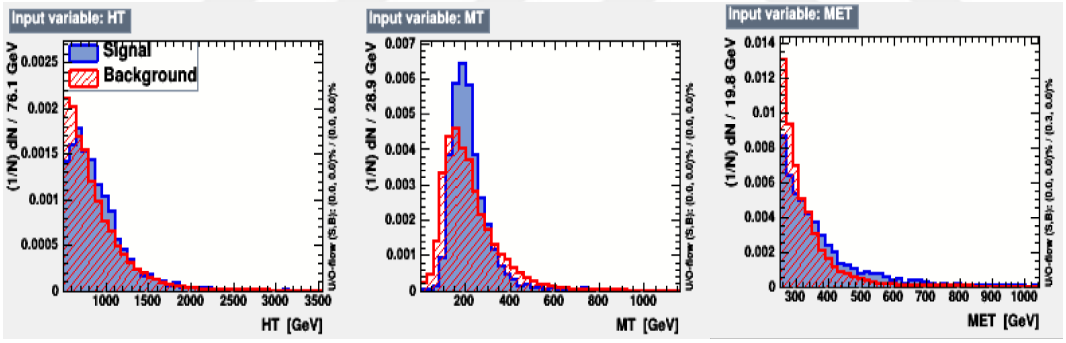
Results for baseline and signal selections for 50PU with $300 fb^{-1}$ is presented in Table 4. After these selections, top+jets events become the minority background with 821 events, whereas $t\bar{t}+jets$ events are the majority with 132101 events. ML1 is the signal sample with the maximum number of events (24758) after these selections, whereas the signal ML4 ends up with the least number of events, 744. Table 5 displays the results after applying chosen kinematic variables. It is important to notice that these variables are not applied consecutively, but interchangeably instead. Results for each row must be considered individually. When suppressing the dominant $t\bar{t}+Jets$ background, M_{T2}^{W} proves to be the most effective variable. Our variable $\alpha_{M_{T2}^{W}}$ is a close second, eliminating more than half of the background. Topness variable is seen to be the least

effective. Generally, $\alpha_{M_{T2}^W}$ and M_{T2}^W effect the signal samples similarly, whereas topness has the least influence on signal.

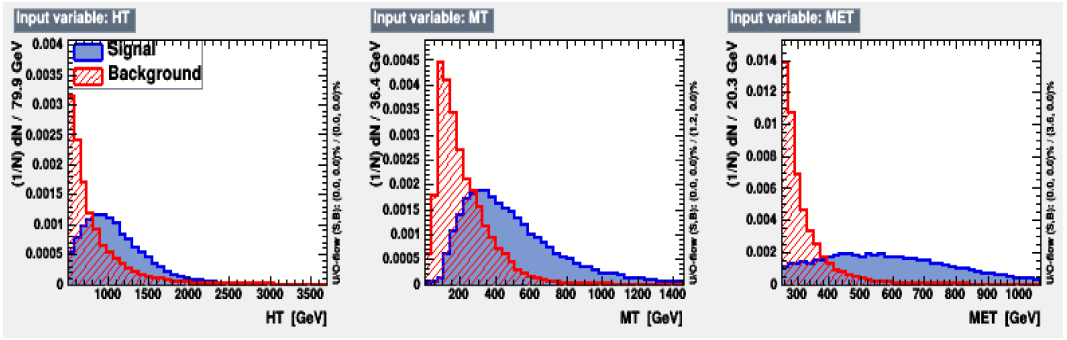
Results for baseline and signal selections for 140PU with $3000 fb^{-1}$ is presented in Table 4. 6. After these selections, top+jets events become the minority background with 54177 events, whereas $t\bar{t} + jets$ events dominate with 2748100 events. ML1 is the signal sample with the maximum number of events (290329) after these selections, whereas the signal ML4 ends up with the least number of events, 8794. Table 5 displays the results after applying chosen kinematic variables. Variable performances are consistent with 50PU $3000 fb^{-1}$ scenario.

4.1 Results of TMVA 50PU 300 fb^{-1}

Input variables fed for the training are shown in Figure 4.1. From the figure, it can be deduced that for ML1, the variables have similar signal and background distributions before training. For ML5, signal is already above the SM background in most areas. An easier separation for ML5 can be expected, from just looking at this distributions.



(a) ML1



(b) ML5

Figure 4.1 : Signal and background distributions of main input variables H_T, M_T and E_T^{miss} fed for each training.

Each training is done with four input parameters. Three out of four parameters are always chosen to be H_T, E_T^{miss} and M_T. In order to differentiate between M_{T2}^W, *topness*, $\alpha_{M_{T2}^W}$, only one of them are used as the fourth parameter in each turn. These interchangeable variables are shown in Figure 4.2. $\alpha_{M_{T2}^W}$ variable shows a dramatic change between ML1 and ML5. ML1 signal and the SM background distributions are quite similar to each other, which makes the separation harder. For ML5, signal and background already distributed in different regions. Rest of the variables exhibit similar behavior.

The results for the correlation between input variables for ML1 signal and background samples can be seen in Figure 4.3. Linear correlation coefficients is given in percentages in matrix form. Correlation varies from +100% to -100%. The red diagonal obviously implies that variables are fully correlated with themselves.

Discussing the signal values for the topness variable, highest correlation is with M_T by 15%, and lowest with E_T^{miss} by 6%. M_{T2}^W variable is most correlated with H_T with a 13%, and least correlated with E_T^{miss}. $\alpha_{M_{T2}^W}$ variable has a -11% correlation with E_T^{miss}, and -13% with H_T. Highest correlation is always between E_T^{miss} and H_T with 55+%.

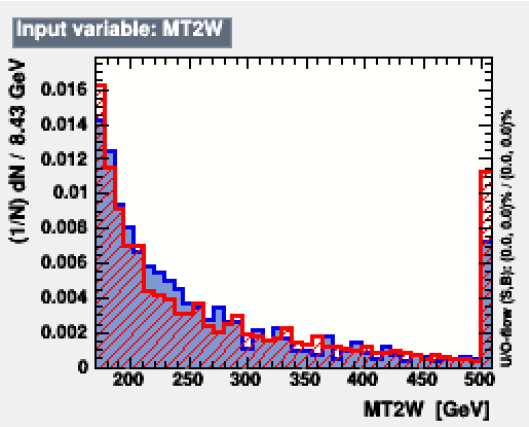
Discussing the background values for the topness variable, highest correlation is with M_T by 20%, and lowest with E_T^{miss} by 6%. M_{T2}^W variable is most correlated with H_T with a 12%, and least correlated with M_T, 3%. $\alpha_{M_{T2}^W}$ variable has a -9% correlation with M_T, and -13% with H_T. Highest correlation is always between E_T^{miss} and H_T with 20+%.

The results for the correlation between input variables for ML5 signal and background samples can be seen in Figure 4.4. Discussing the signal values for the topness variable, highest correlation is with M_T by 2%, and lowest with E_T^{miss} by -6%. M_{T2}^W variable is most correlated with E_T^{miss} with a 22%, and least correlated with H_T, with 14%. $\alpha_{M_{T2}^W}$ variable has a -13% correlation with M_T, and 12% with H_T. Highest correlation is always between E_T^{miss} and H_T with 30+%.

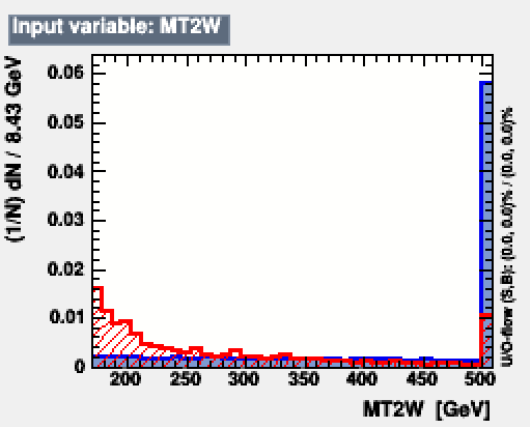
Discussing the background values for the topness variable, highest correlation is with M_T by 21%, and lowest with E_T^{miss} by 5%. M_{T2}^W variable is most correlated with E_T^{miss} with a 11%, and least correlated with M_T, 5%. $\alpha_{M_{T2}^W}$ variable has a 6% correlation with

M_T, and -12% with H_T and E_T^{miss}. Highest correlation is always between E_T^{miss} and H_T with 25+%.

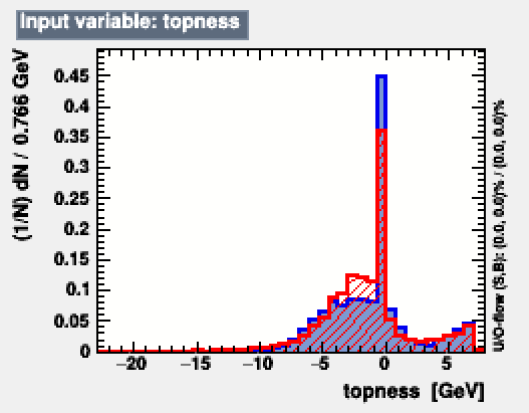
Classifier output distributions, and efficiency for the topness variable with BDT method can be observed from Figure 4.6 for ML1 and Figure 4.7 for ML5.



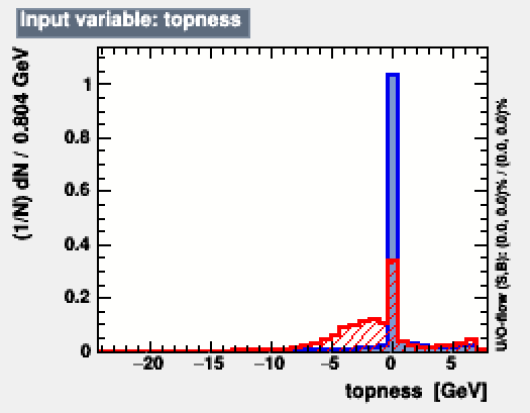
(a) M_{T2}^W - ML1



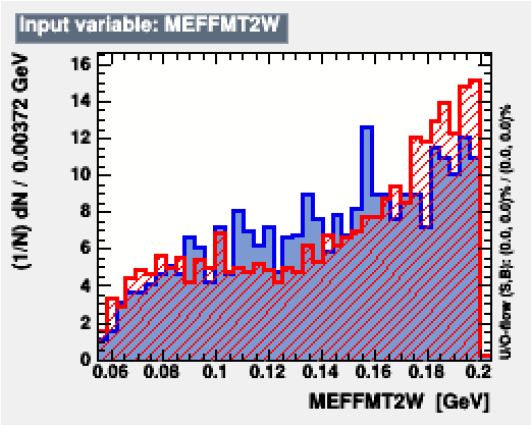
(b) M_{T2}^W - ML5



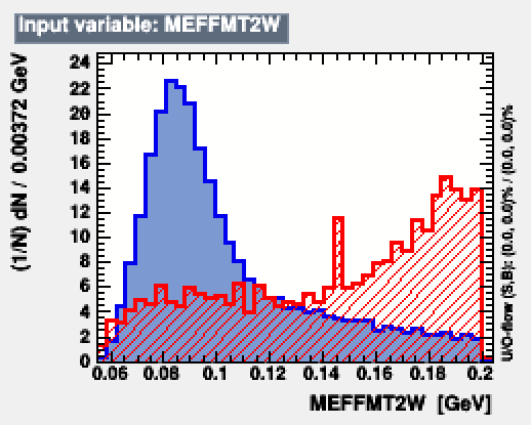
(c) Topness - ML1



(d) Topness - ML5

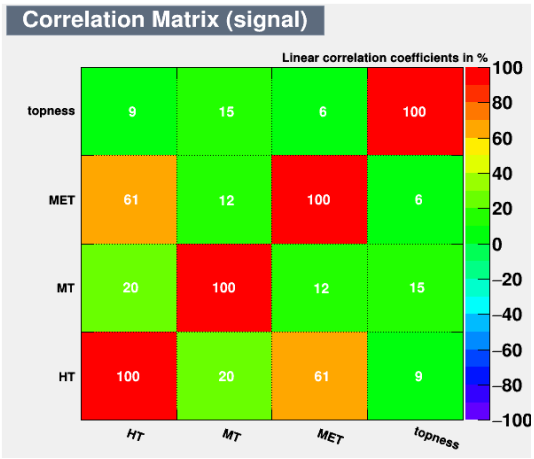


(e) $\alpha_{M_{T2}^W}$ - ML1

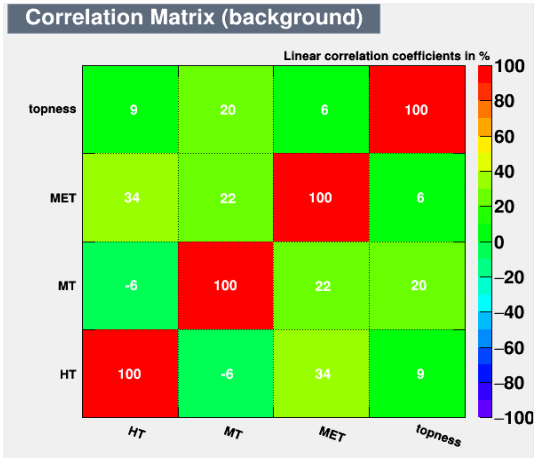


(f) $\alpha_{M_{T2}^W}$ - ML5

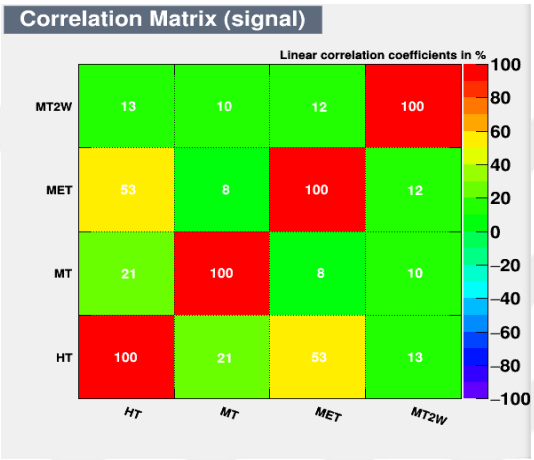
Figure 4.2 : Distribution of the input variable used for separate trainings, for 50PU and $300 fb^{-1}$. Variables and corresponding signals are addressed in the subcaptions.



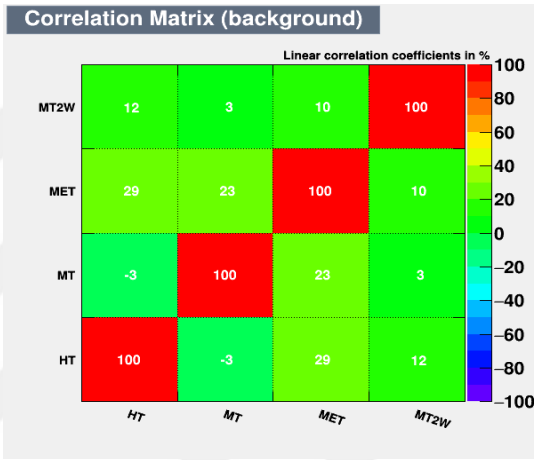
(a) Topness - ML1



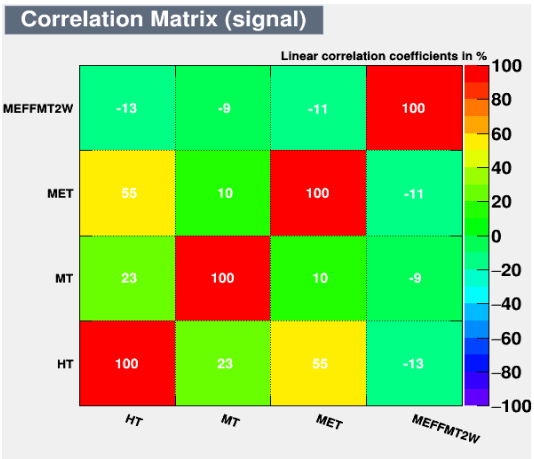
(b) Topness - ML1



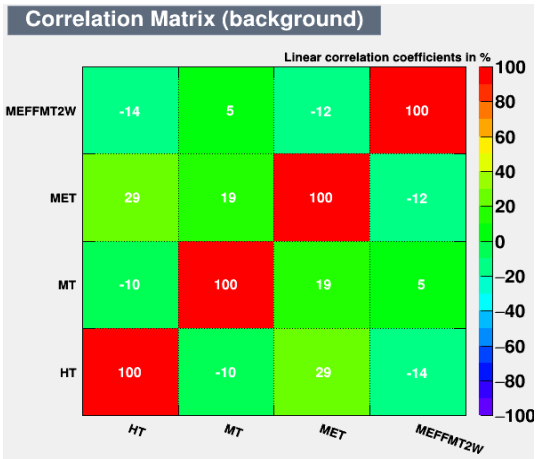
(c) M_{T2}^{W} - ML1



(d) M_{T2}^{W} - ML1



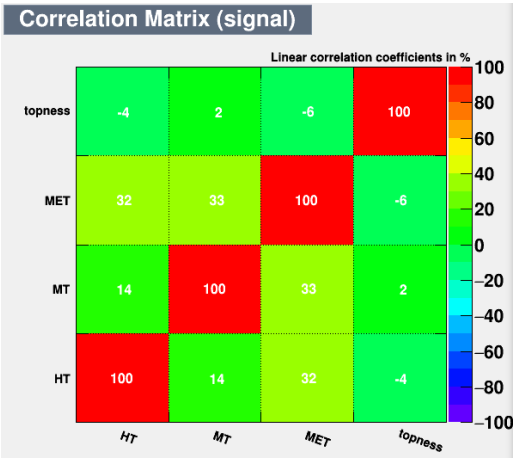
(e) $\alpha_{M_{T2}^{W}}$ - ML1



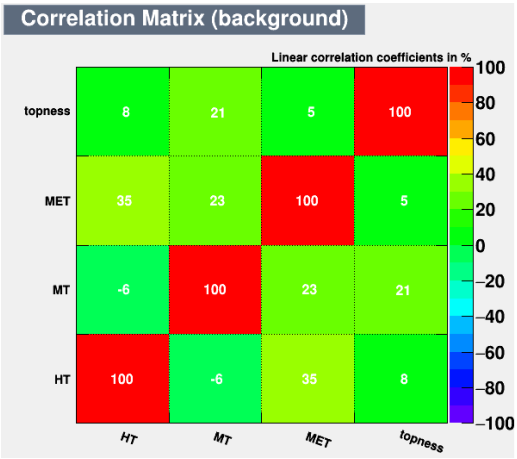
(f) $\alpha_{M_{T2}^{W}}$ - ML1

Figure 4.3 : Correlation between input variables illustrated in matrices for 50PU, 300 fb^{-1}, signal ML1.

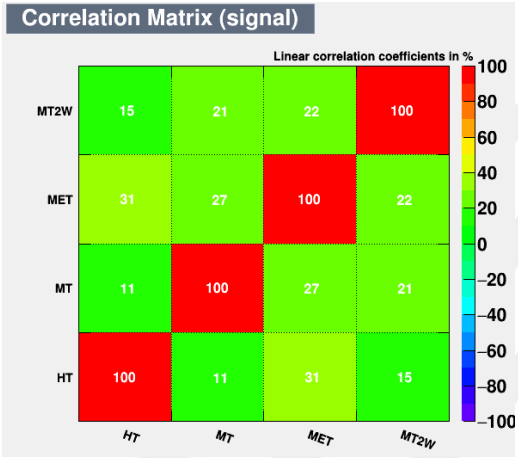
Left columns of this figures descibe the signal and background distributions for given variables. On the right column, cut efficiencies with optimal cut value can be observed.



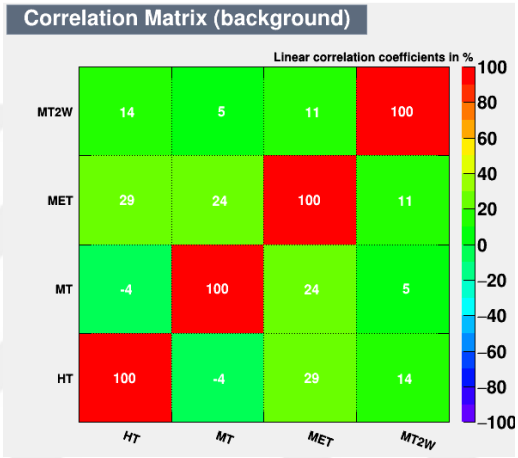
(a) Topness - ML5



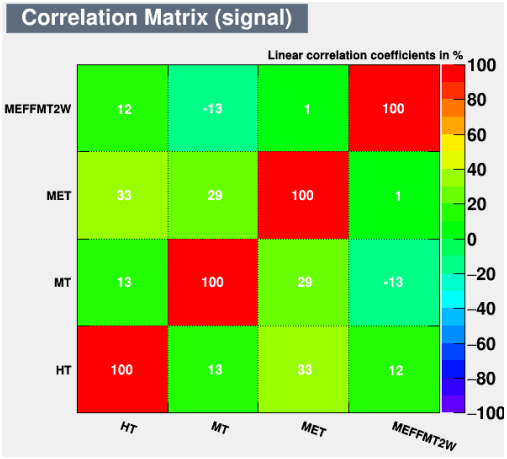
(b) Topness - ML5



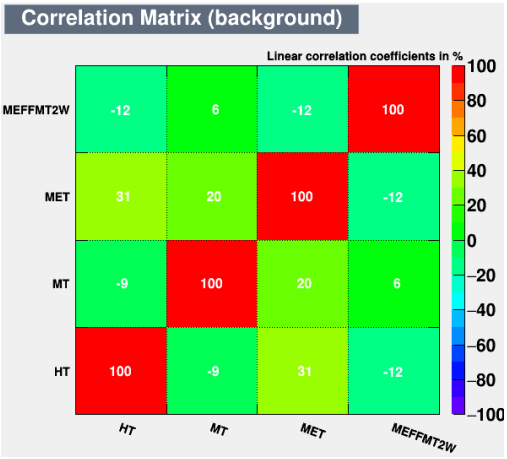
(c) M_{T2}^{W} - ML5



(d) M_{T2}^{W} - ML5



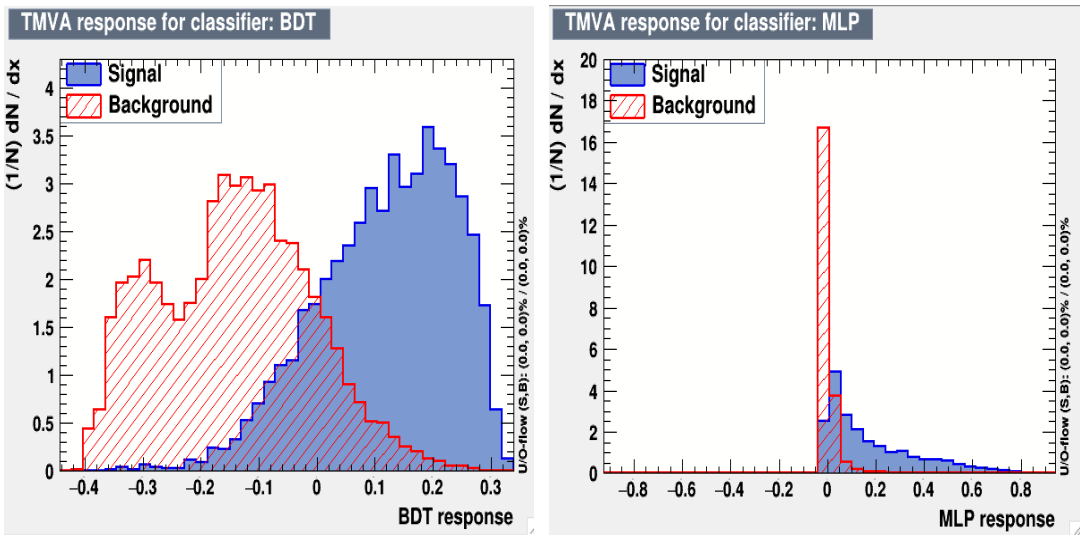
(e) $\alpha_{M_{T2}^{W}}$ - ML5



(f) $\alpha_{M_{T2}^{W}}$ - ML5

Figure 4.4 : Correlation between input variables illustrated in matrices for 50PU, 300 fb^{-1}, signal ML5. First column of plots (a, c, e) presents results for signal, second row (b, d, f) shows the background.

For discrimination, the discriminator values given with BDT is used. Discriminator is determined by examining the shape that gives $S/\sqrt{S+B} = 5$. More clearly,



(a) M_{T2}^{W} - ML4 (b) M_{T2}^{W} - ML4

Figure 4.5 : BDT and MLP responses of ML4, trained with M_{T2}^{W} at 50PU $300 fb^{-1}$.

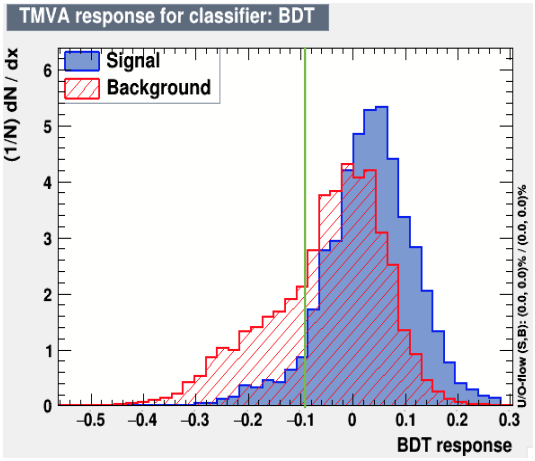
the region of interest should be greater than 5σ whereas 3σ is considered to be excess. When compared, trainings done with ML5 are dramatically more successful compared to ML1, with every variable granting distinguishable separation of signal and background.

Classifier output distributions, and efficiency for the topness variable with MLP method can be observed from Figure 4.8 for ML1. Left columns of this figures descibe the signal and background distributions for given variables. On the right column, cut efficiencies with optimal cut value can be observed. For discrimination, the discriminator values given with MLP is used. As compared in Figure 4.5, trainings done with BDT separated signal from the background better than the MLP method.

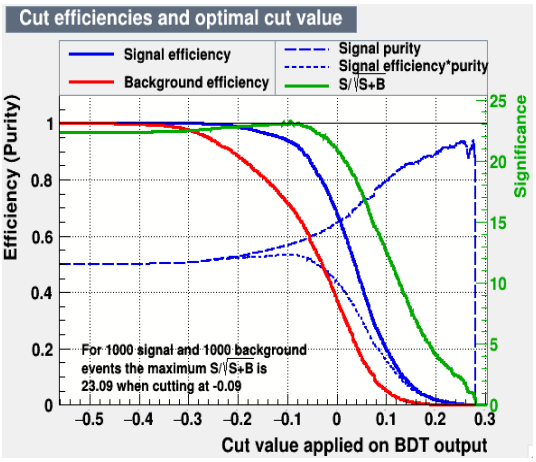
Figure 1 presents BDT results of different signal samples trained with $\alpha_{M_{T2}^{W}}$. In this scenario, ML1 proves to be the hardest signal to separate. Among all signals, with 483 GeV it has the lightest stop mass. As the stop gets heavier with ML3, ML4 and ML5, separations become more efficient.

4.2 Results of TMVA 140PU 3000 fb^{-1}

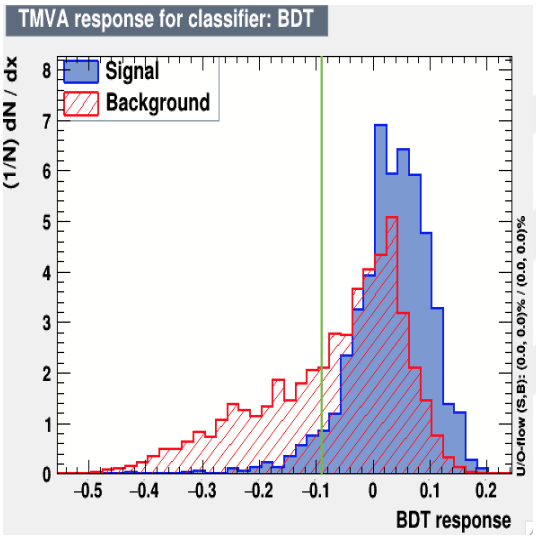
Input variables fed for the training are shown in Figure 4.10. For ML1, signal and background show similar distribution, which can indicate to harder separation. For ML5, signal is already more dominant in lower energy regions.



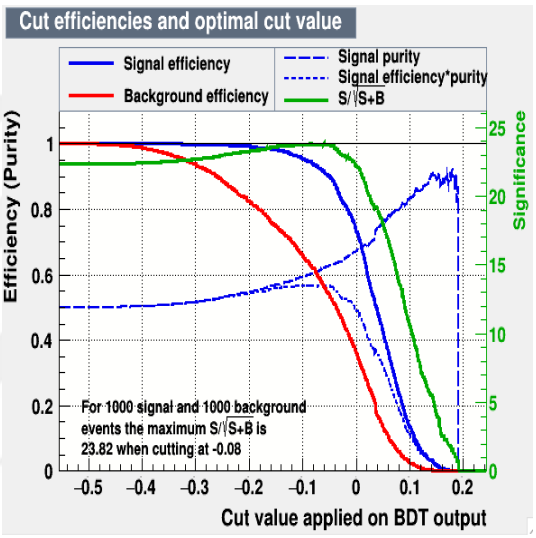
(a) Topness - ML1



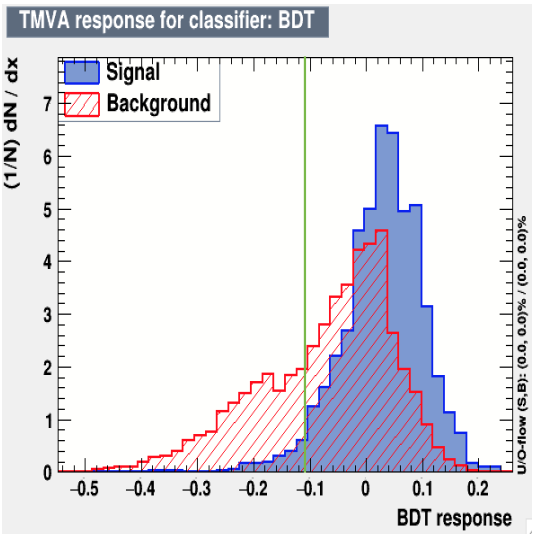
(b) Topness - ML1



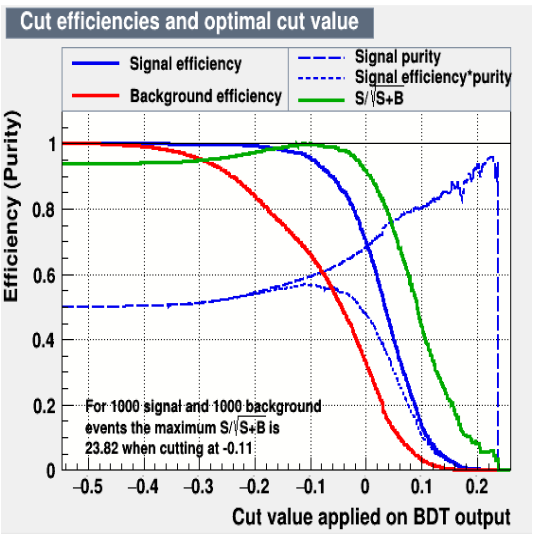
(c) M_{T2}^{W} - ML1



(d) M_{T2}^{W} - ML1

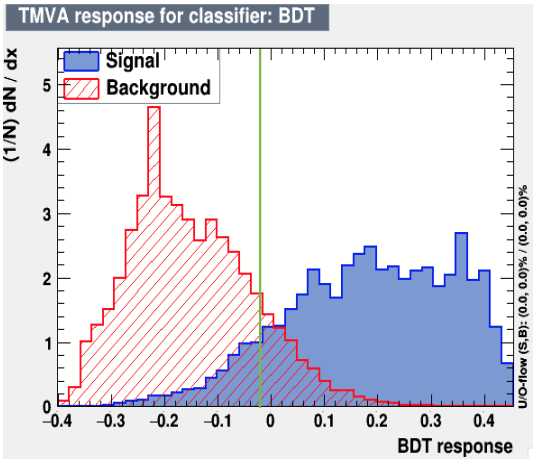


(e) $\alpha_{M_{T2}^{W}}$ - ML1

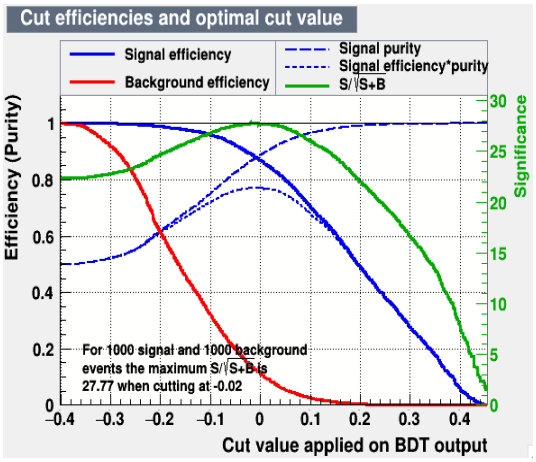


(f) $\alpha_{M_{T2}^{W}}$ - ML1

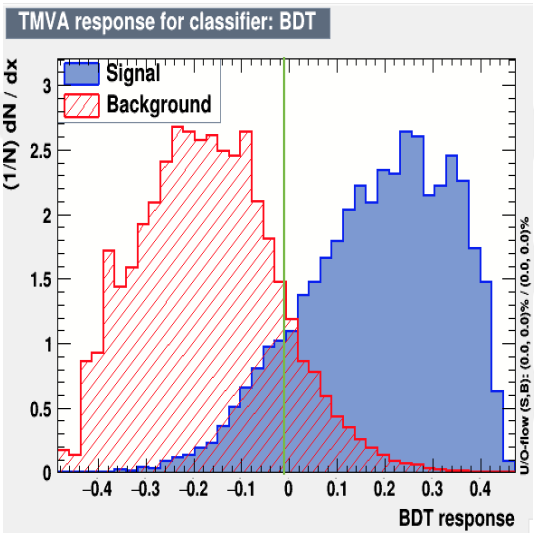
Figure 4.6 : Left column (a, c, e) shows TMVA responses for BDT method, for signal ML1 at 50PU, 300 fb^{-1}. Right column (b, d, f) presents cut efficiencies and optimal cut values.



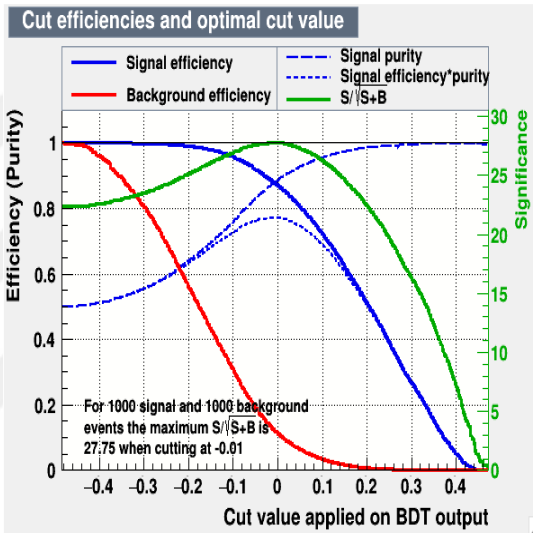
(a) Topness - ML5



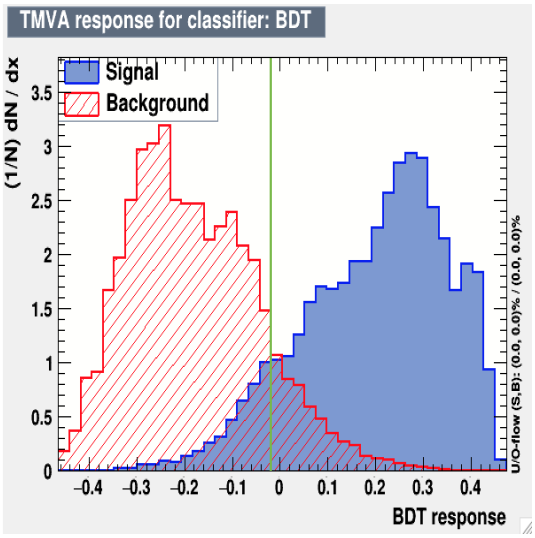
(b) Topness - ML5



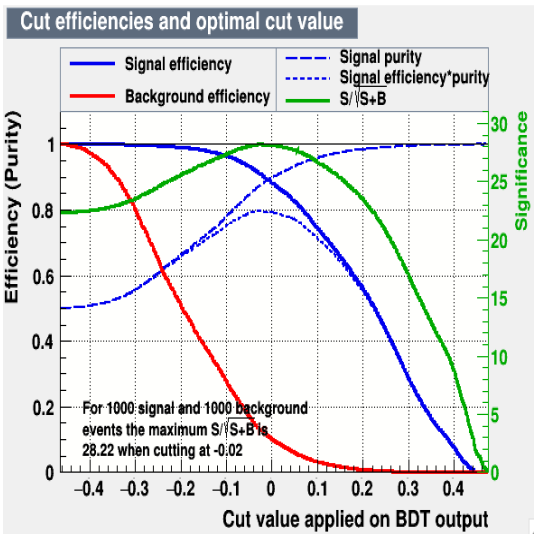
(c) M_{T2}^{W} - ML5



(d) M_{T2}^{W} - ML5

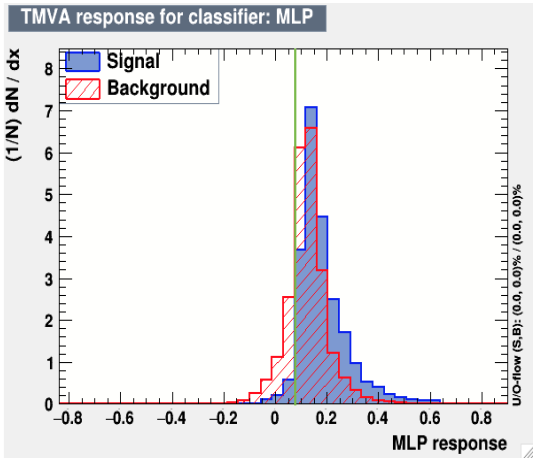


(e) $\alpha_{M_{T2}^{W}}$ - ML5

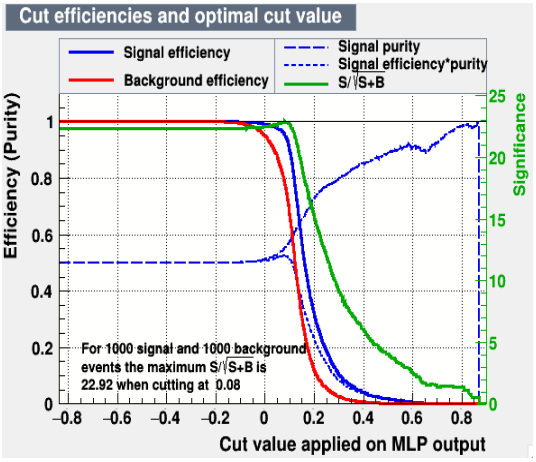


(f) $\alpha_{M_{T2}^{W}}$ - ML5

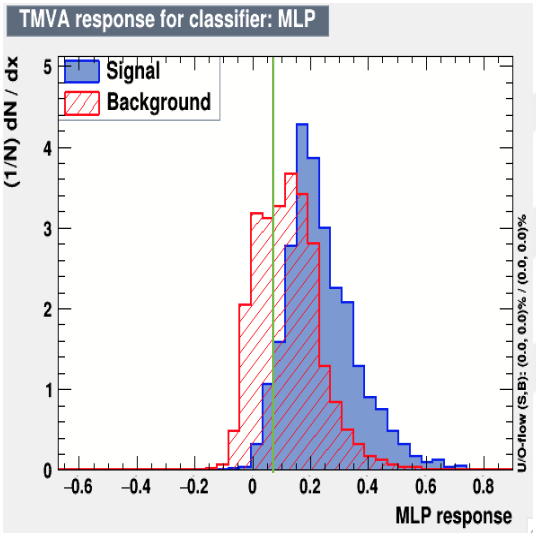
Figure 4.7 : Left column (a, c, e) showsTMVA responses for BDT method, for signal ML1 at 50PU, 300 fb^{-1}. Right column (b, d, f) presents cut efficiencies and optimal cut values.



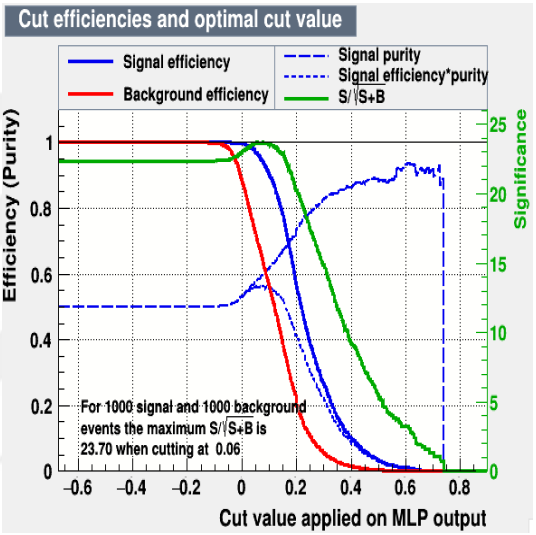
(a) Topness - ML1



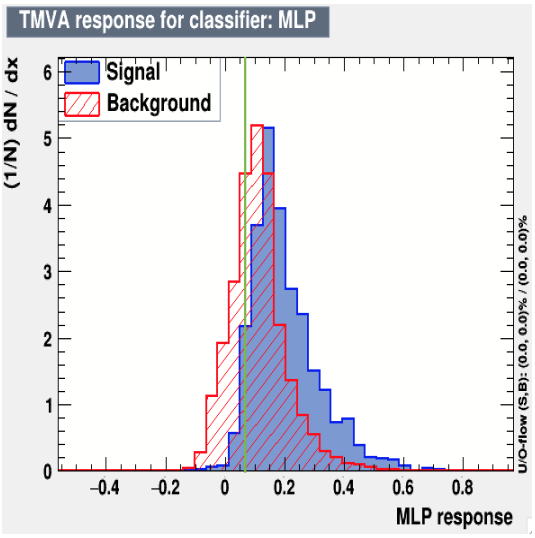
(b) Topness - ML1



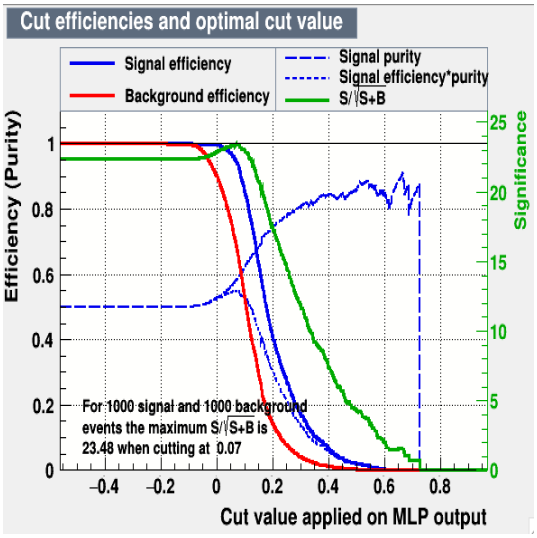
(c) M_{T2}^{W} - ML1



(d) M_{T2}^{W} - ML1

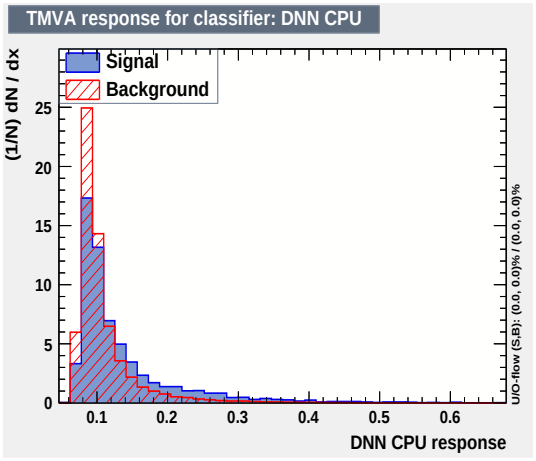


(e) $\alpha_{M_{T2}^{W}}$ - ML1

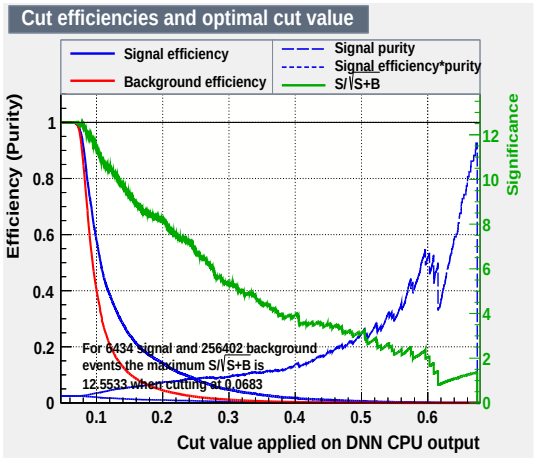


(f) $\alpha_{M_{T2}^{W}}$ - ML1

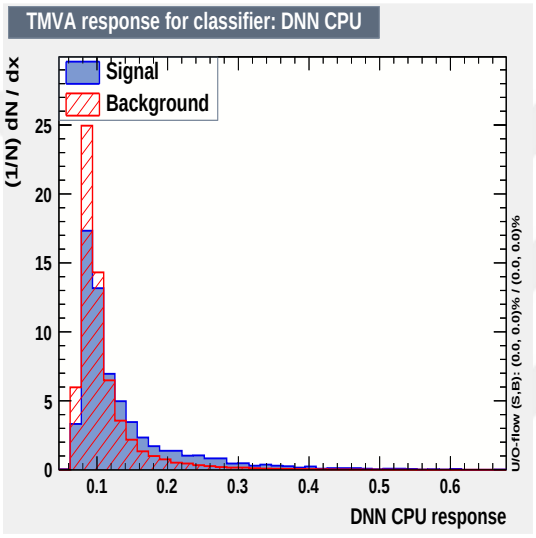
Figure 4.8 : Left column (a, c, e) shows TMVA responses for MLP method, for signal ML1 at 50PU, 300 fb^{-1}. Right column (b, d, f) presents cut efficiencies and optimal cut values.



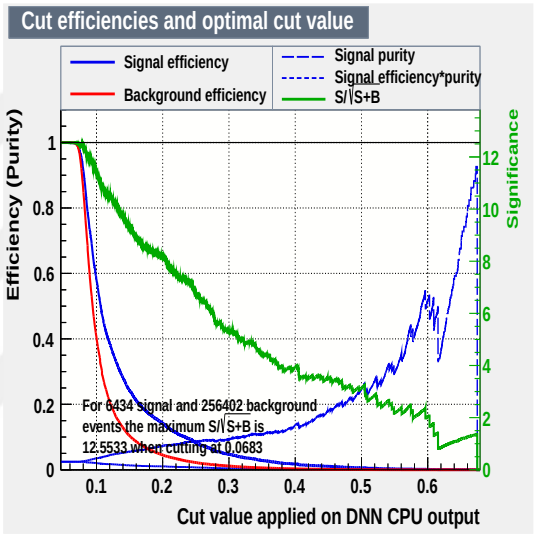
(a) ML1 - M_{T2}^{W}



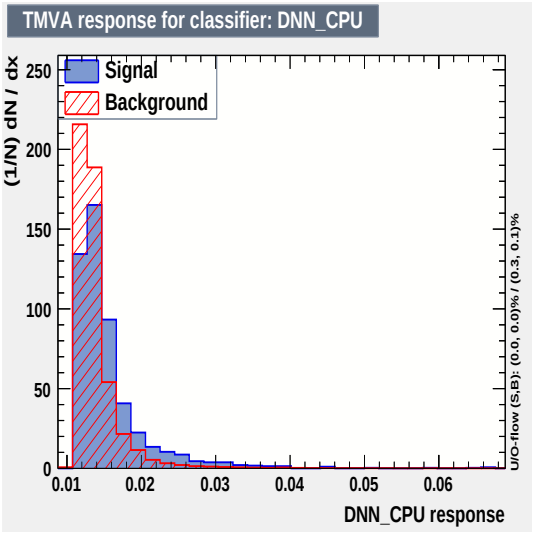
(b) ML1 - M_{T2}^{W}



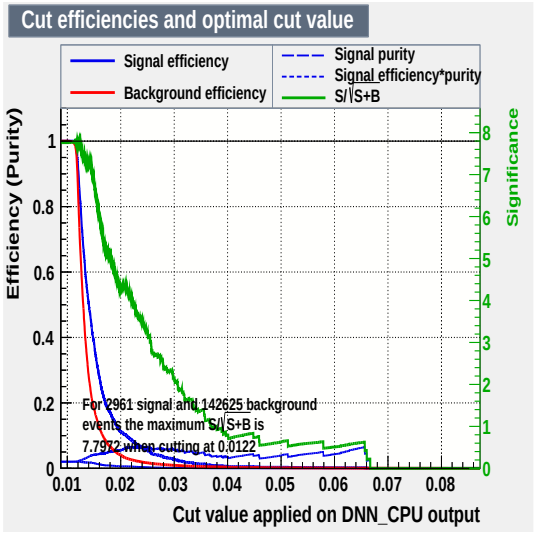
(c) ML1 - Topness



(d) ML1 - Topness



(e) ML1 - $\alpha_{M_{T2}^{W}}$



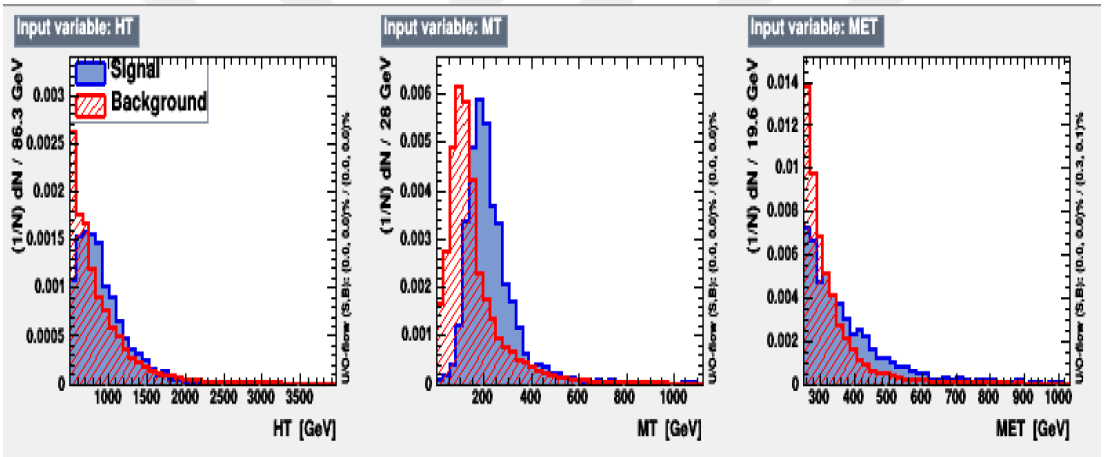
(f) ML1 - $\alpha_{M_{T2}^{W}}$

Figure 4.9 : Left column (a, c) shows TMVA responses for DNN method, for signal ML1 at 50PU, 300 fb^{-1}. Right column (b, d) presents cut efficiencies and optimal cut values.

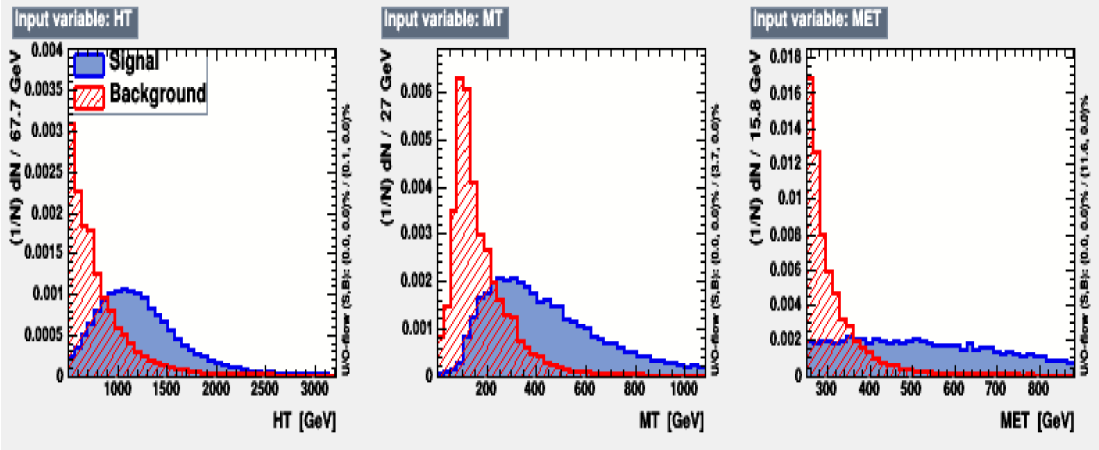
The results for the correlation between input variables for ML1 signal and background samples can be seen in Figure 4.11. Linear correlation coefficients is given in percentages in matrix form. Correlation varies from +100% to -100%. The red diagonal obviously implies that variables are fully correlated with themselves.

Discussing the signal values for the topness variable, highest correlation is with M_T by 15%, and lowest with E_T^{miss} by 1%. M_{T2}^W variable is most correlated with H_T with a 16%, and least correlated with E_T^{miss} with 11%. $\alpha_{M_{T2}^W}$ variable has a -10% correlation with M_T, and -8% with E_T^{miss}. Highest correlation is always between E_T^{miss} and H_T with 40+%.

Discussing the background values for the topness variable, highest correlation is with M_T by 9%, and lowest with E_T^{miss} and H_T by 4%. M_{T2}^W variable is most correlated with

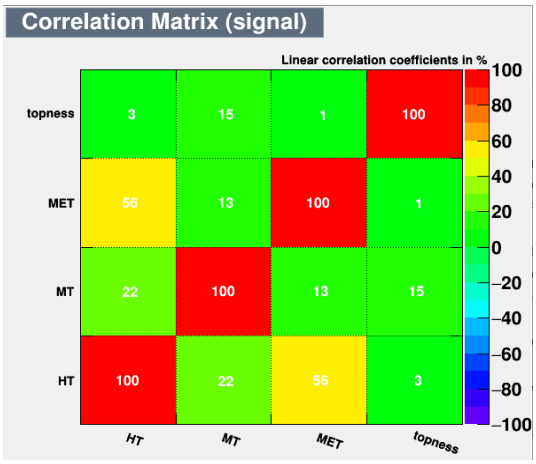


(a) ML1

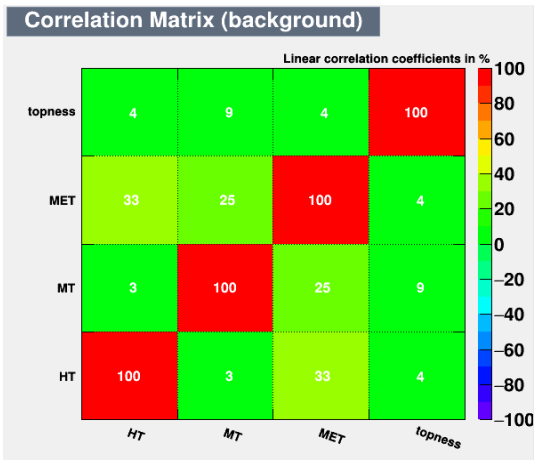


(b) ML5

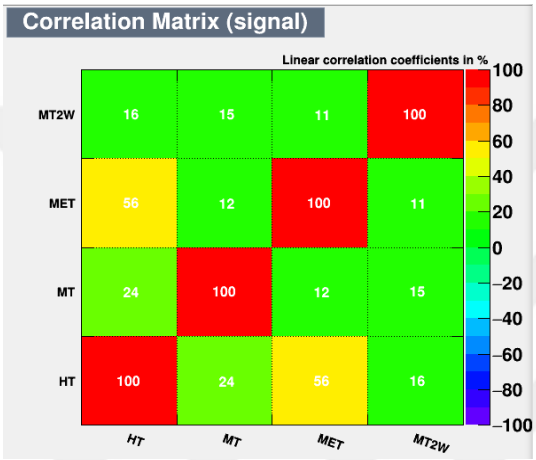
Figure 4.10 : Signal and background distributions of main input variables H_T, M_T and E_T^{miss} fed for each training.



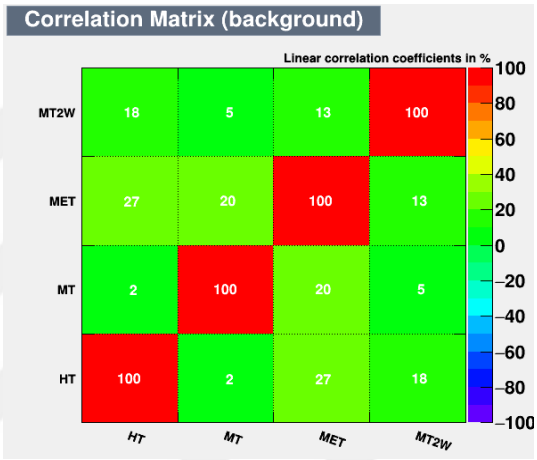
(a) Topness - ML1



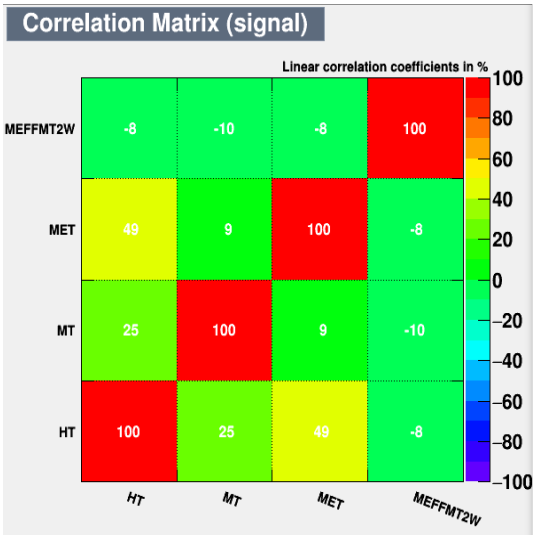
(b) Topness - ML1



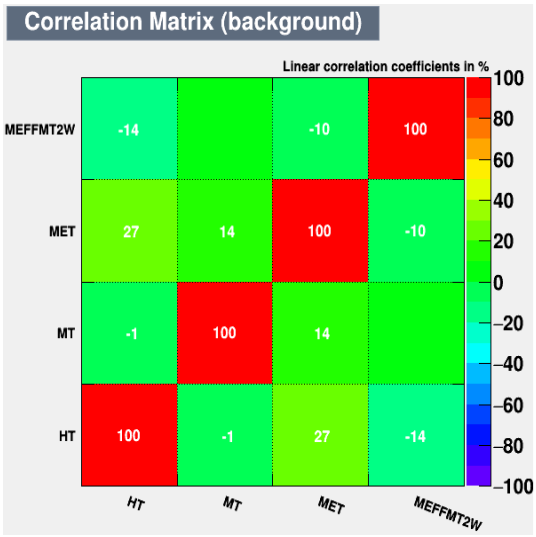
(c) M_{T2}^{W} - ML1



(d) M_{T2}^{W} - ML1



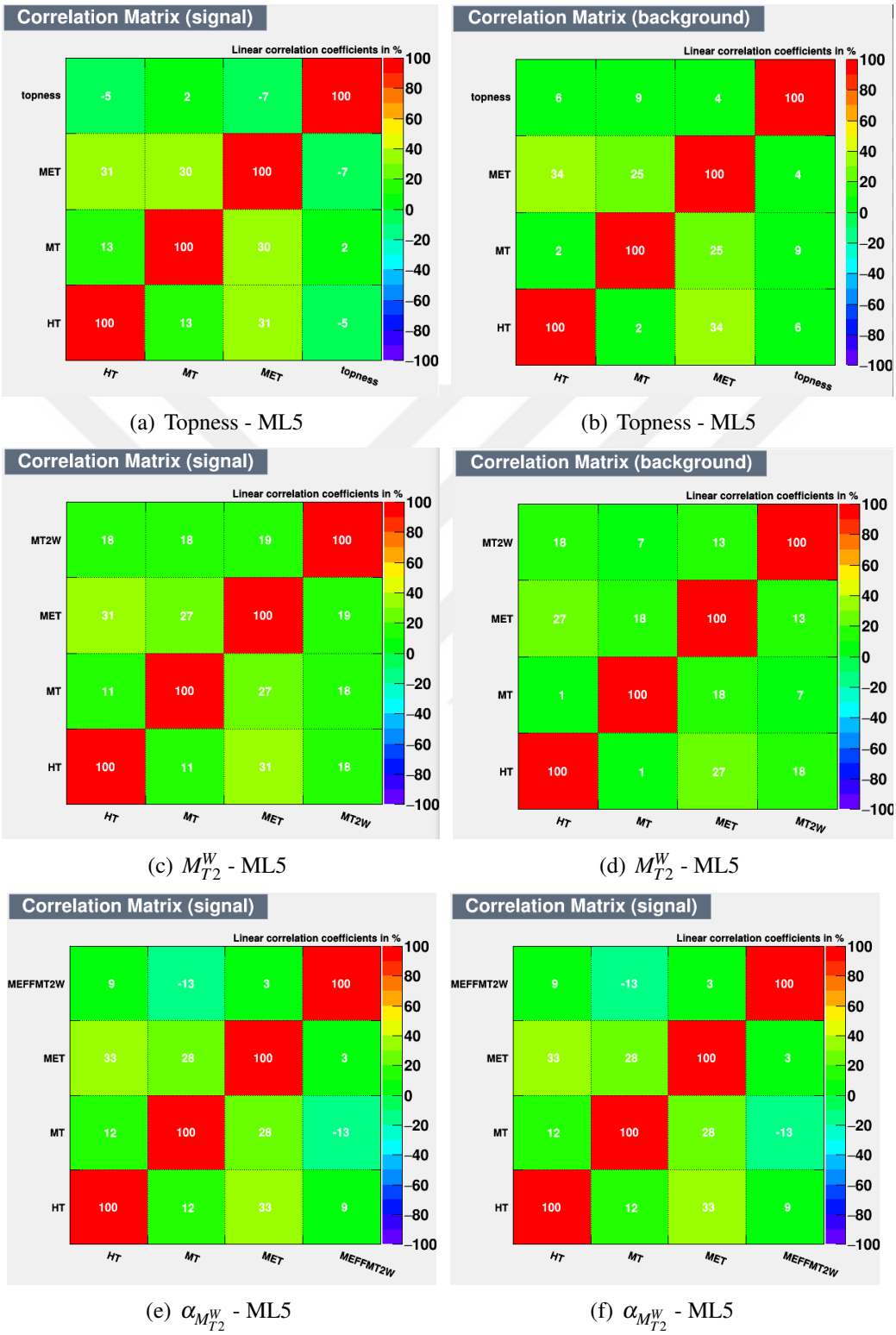
(e) $\alpha_{M_{T2}^{W}}$ - ML1



(f) $\alpha_{M_{T2}^{W}}$ - ML1

Figure 4.11 : Correlation between input variables illustrated in matrices for 140PU, 3000 fb^{-1}, signal ML1. First column of plots (a, c, e) presents results for signal, second row (b, d, f) shows the background.

H_T with a 18%, and least correlated with M_T, 5%. $\alpha_{M_{T2}^W}$ variable has a 5% correlation with M_T, and 18% with H_T. Highest correlation is always between E_T^{miss} and H_T with 20+%.



(a) Topness - ML5

(b) Topness - ML5

(c) M_{T2}^W - ML5

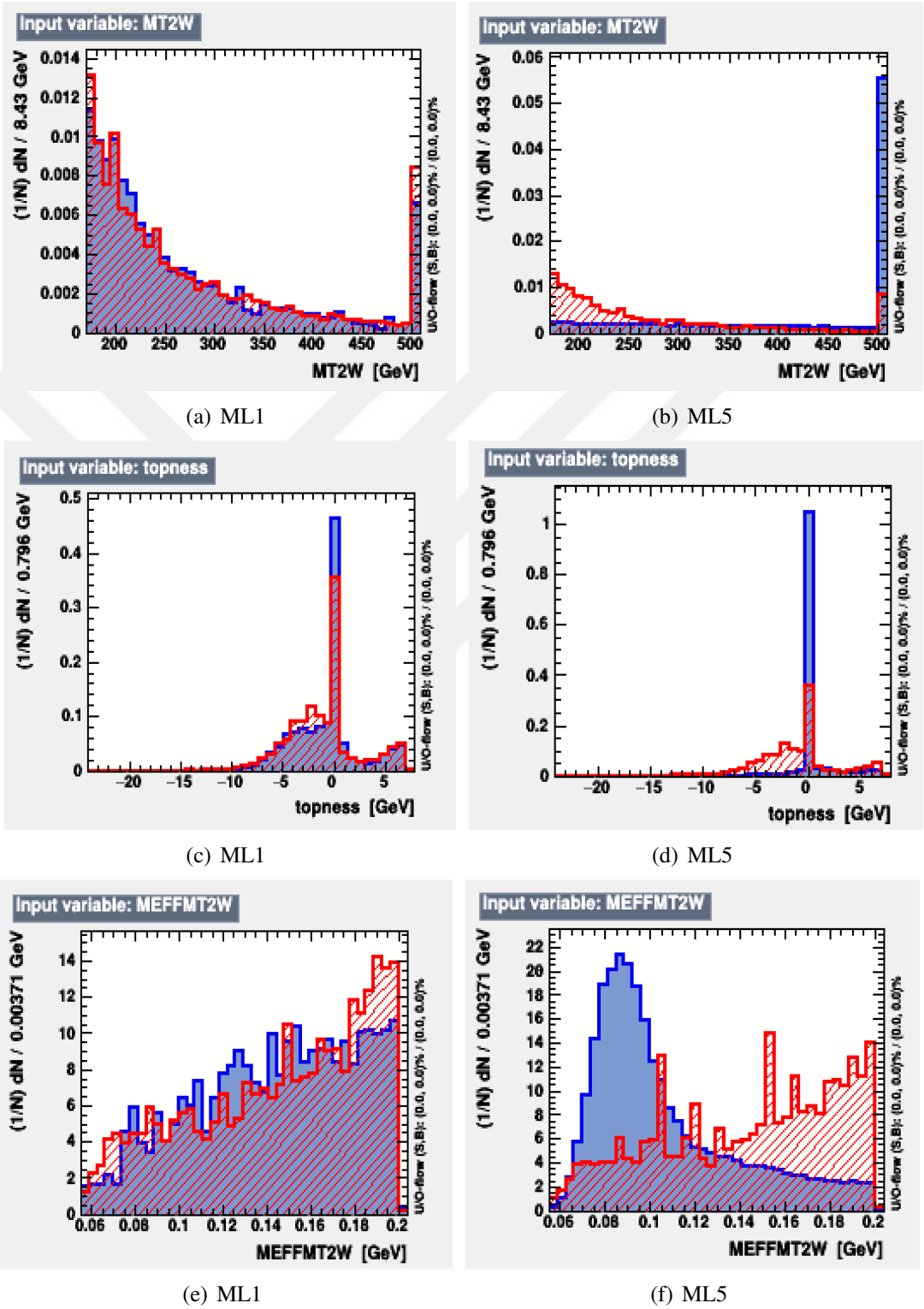
(d) M_{T2}^W - ML5

(e) $\alpha_{M_{T2}^W}$ - ML5

(f) $\alpha_{M_{T2}^W}$ - ML5

Figure 4.12 : Correlation between input variables illustrated in matrices for 140PU, 3000 fb^{-1}, signal ML5. First column of plots (a, c, e) presents results for signal, second row (b, d, f) shows the background.

The results for the correlation between input variables for ML5 signal and background samples can be seen in Figure 4.12. Discussing the signal values for the topness variable, highest correlation is with M_T by 2%, and lowest with E_T^{miss} by 2%.



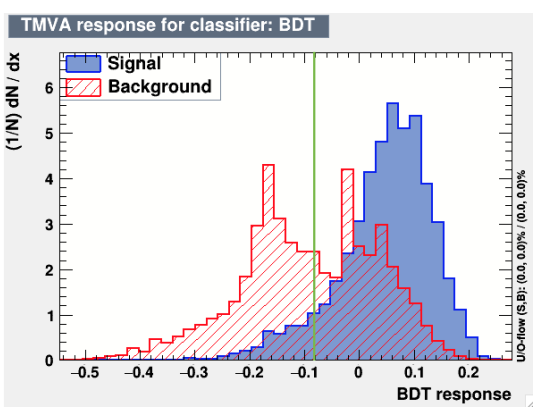
(a) ML1 (b) ML5

(c) ML1 (d) ML5

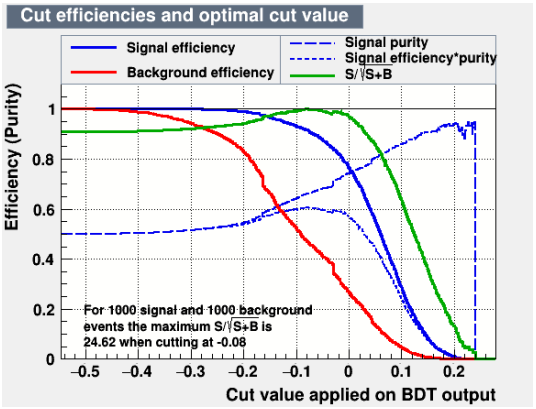
(e) ML1 (f) ML5

Figure 4.13 : Interchangeable input variables at 140PU, 3000 fb^{-1}. Left column (a, c, e) presents ML1 signal, right column (b, d, f) shows ML5.

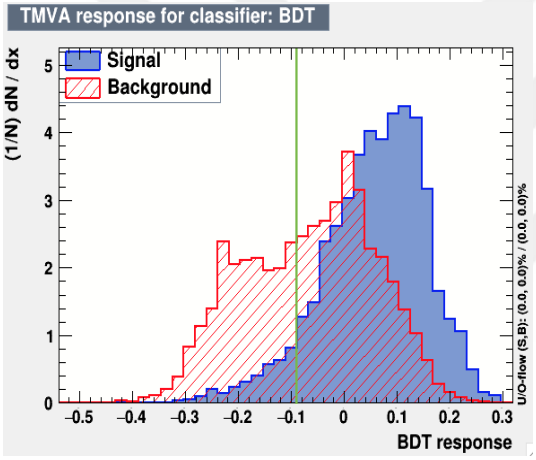
M_{T2}^{W} variable is most correlated with E_T^{miss} with a 19%, and least correlated with m_T, with 18%. $\alpha_{M_{T2}^{W}}$ variable has a 9% correlation with H_T, and -13% with M_T. Highest correlation is always between E_T^{miss} and H_T with 30+%.



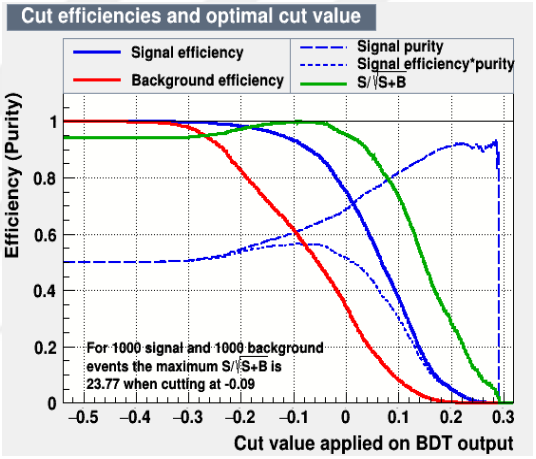
(a) ML1 -M_{T2}^{2}



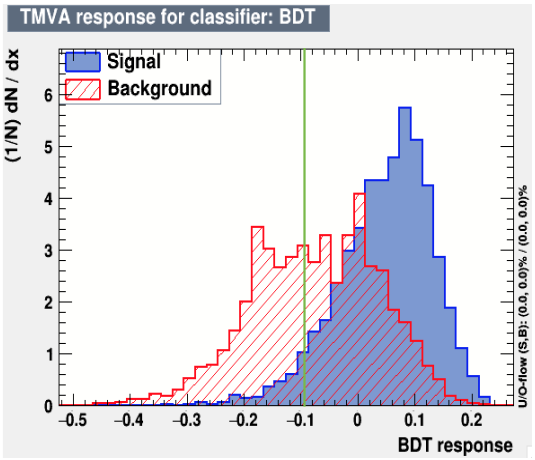
(b) ML1 -M_{T2}^{2}



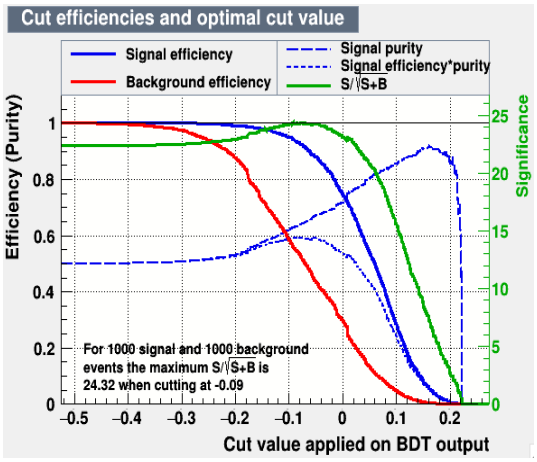
(c) ML1 - Topness



(d) ML1 - Topness

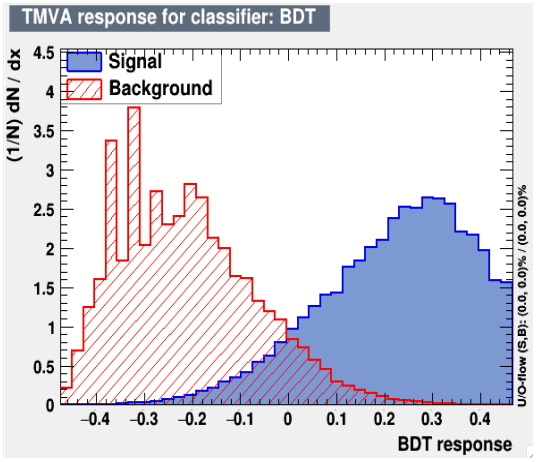


(e) ML1 - $\alpha_{M_{T2}^{W}}$

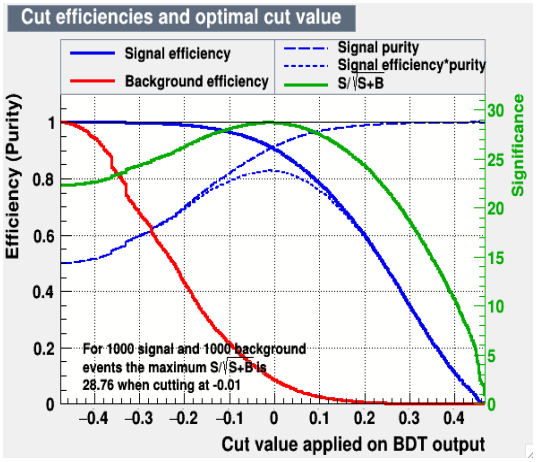


(f) ML1 - $\alpha_{M_{T2}^{W}}^{W}$

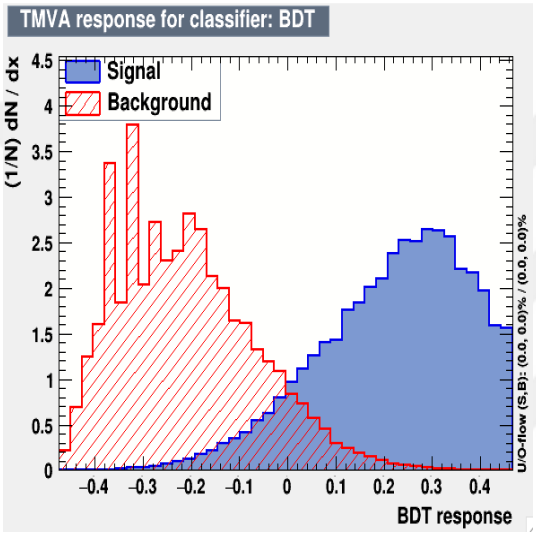
Figure 4.14 : Left column (a, c, e) shows TMVA responses for BDT method, for signal ML1 at 140PU, 3000 fb^{-1}. Right column (b, d, f) presents cut efficiencies and optimal cut values.



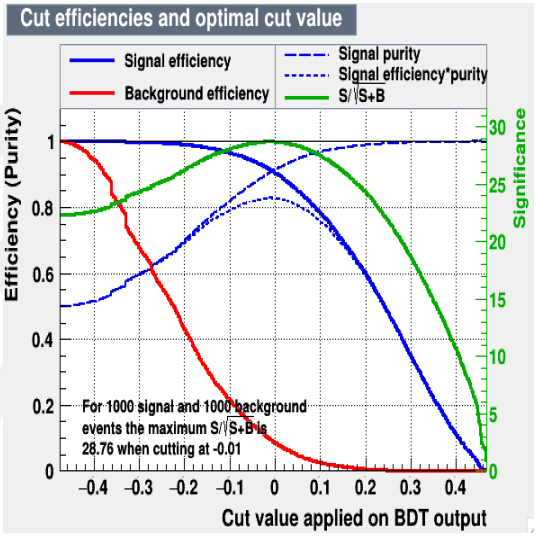
(a) ML5 -M_{T2}^2



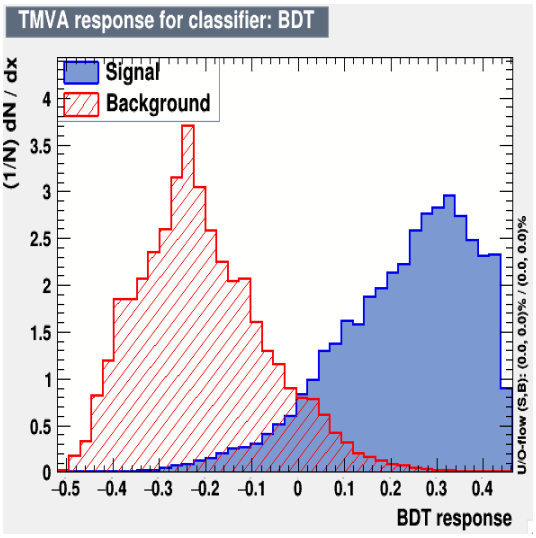
(b) ML5 -M_{T2}^2



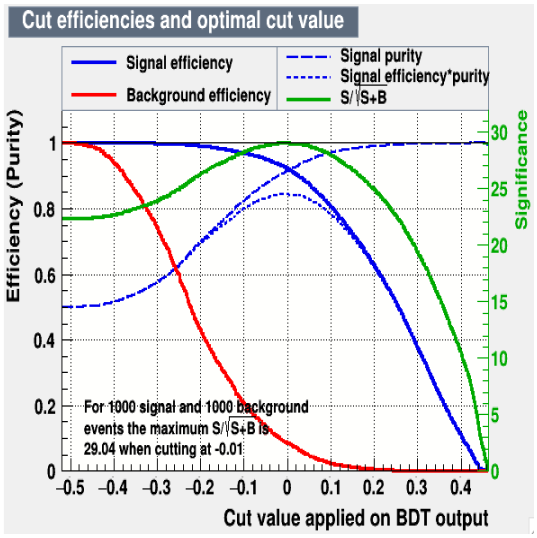
(c) ML5 - Topness



(d) ML5 - Topness



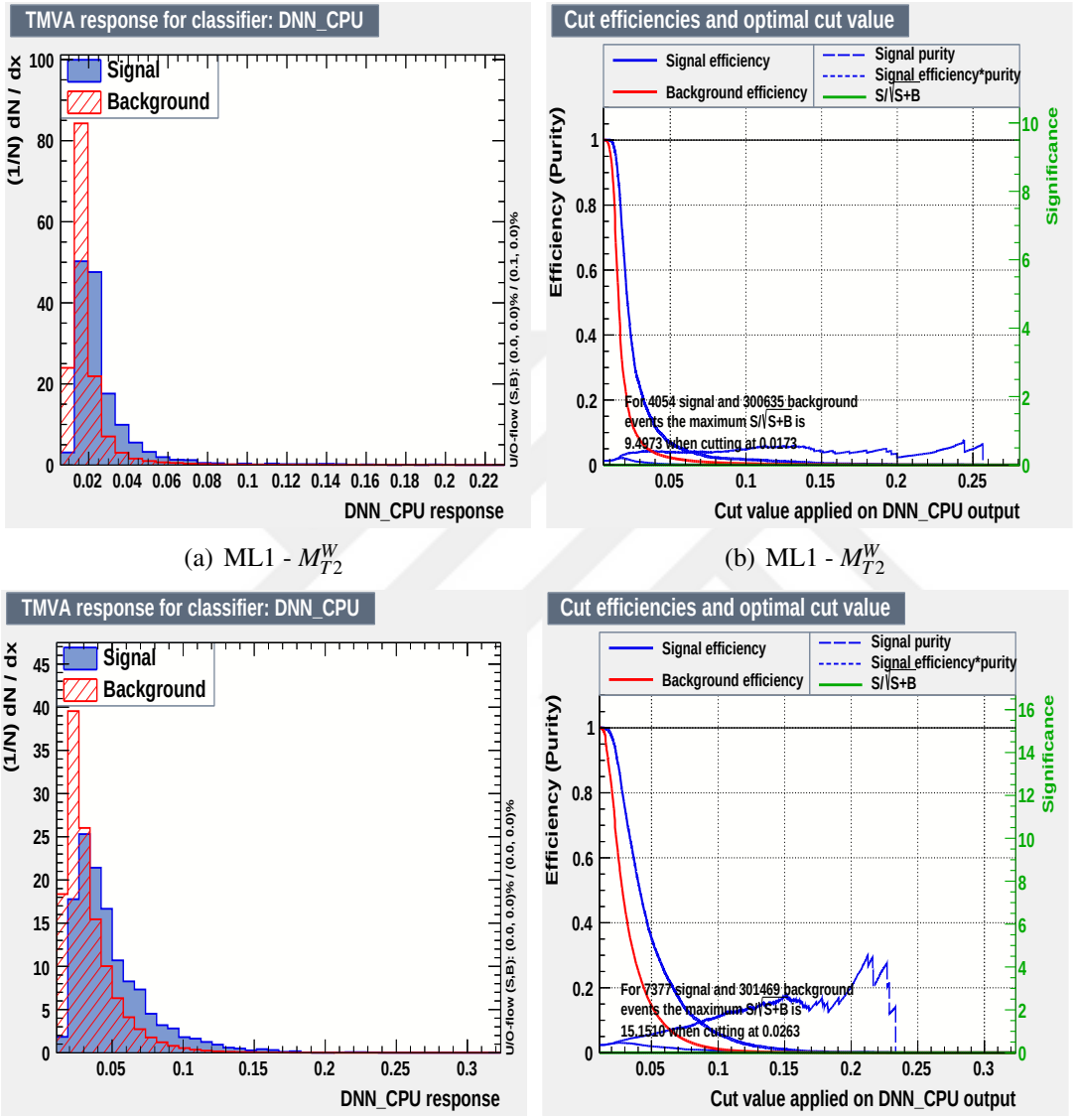
(e) ML5 - $\alpha_{M_{T2}^W}$



(f) ML5 - $\alpha_{M_{T2}^W}^W$

Figure 4.15 : Left column (a, c, e) shows TMVA responses for BDT method, for signal ML5 at 140PU, 3000 fb^{-1}. Right column (b, d, f) presents cut efficiencies and optimal cut values.

Discussing the background values for the topness variable, highest correlation is with H_T by 6%, and lowest with E_T^{miss} by 4%. M_{T2}^W variable is most correlated with H_T with a 18%, and least correlated with M_T, 7%. $\alpha_{M_{T2}^W}$ variable has a -13% correlation with M_T, and -9% with H_T and E_T^{miss}. Highest correlation is always between E_T^{miss} and H_T with 25+%.



(a) ML1 - M_{T2}^W

(b) ML1 - M_{T2}^W

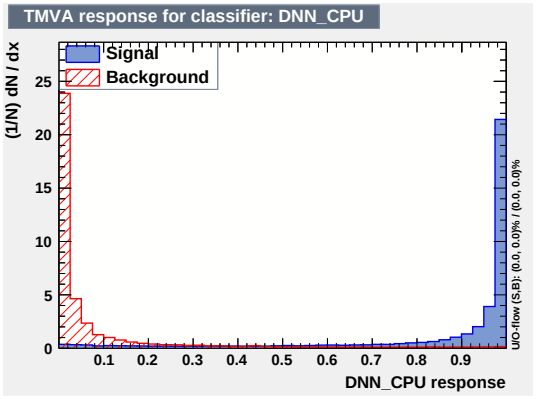
(c) ML1 - Topness

(d) ML1 - Topness

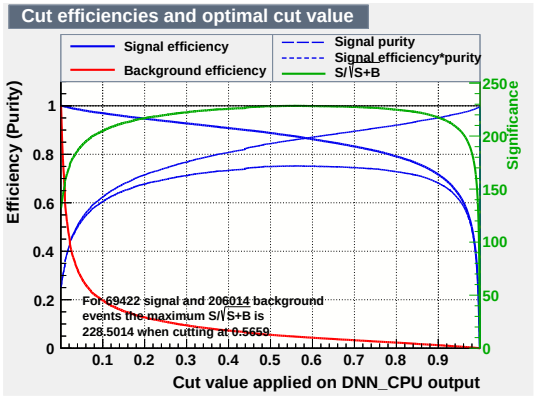
Figure 4.16 : Left column (a, c) shows TMVA responses for DNN method, for signal ML1 at 140PU, 3000 fb^{-1}. Right column (b, d) presents cut efficiencies and optimal cut values.

Classifier output distributions, and efficiency for the topness variable can be observed from Figure 4.14. Figure on the left descibes the signal and background distributions for given variables. On the right, cut efficiencies with optimal cut value can be observed. For discrimination, the discriminator values given with BDT is used.

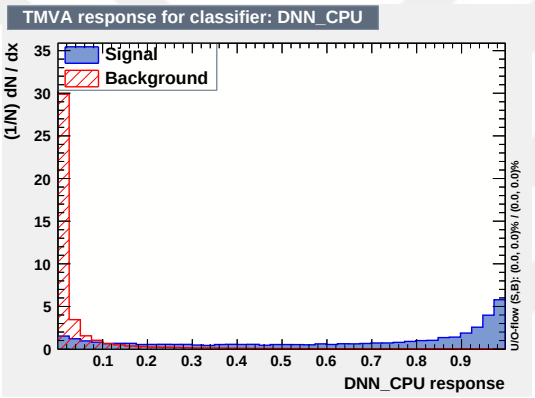
Discriminator is determined by examining the shape that gives $S/\sqrt{S+B} = 5$. More clearly, the region of interest should be greater than 5σ whereas 3σ is considered to be excess.



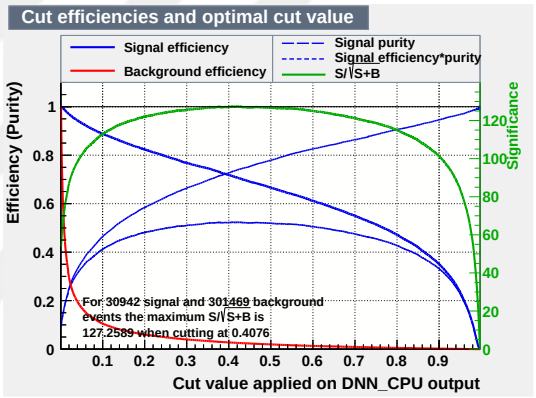
(a) ML5 - $\alpha_{M_{T2}^W}$



(b) ML5 - $\alpha_{M_{T2}^W}$



(c) ML5 - Topness



(d) ML5 - Topness

Figure 4.17 : Left column (a, c) shows TMVA responses for DNN method, for signal ML5 at 140PU, 3000 fb^{-1}. Right column (b, d) presents cut efficiencies and optimal cut values.

Summary of these trainings, BDT, MLP and DNN response corresponding to different scenarios, signals and variables can be seen in Tables 4.1, 4.2, 1, 2, 3 for ML1, ML2, ML3, ML4 and ML5 respectively. For 140PU, M_{T2}^W variable often fails to give healthy results for MLP and DNN trainings.

Results from these tables are illustrated in Figures 4.18, 4.19, 4.20.

Left column of Figure 4.18 investigates the optimal cut with different training methods in every signal, when trained with each interchangeable variable at 50PU scenario. Right column illustrates the significance of these trainings. For signal ML2, optimal cut has similar values for all training methods regardless of the variables, whereas other signals display shifts in the discriminator. For signal ML3, DNN gives lower

significance compared to BDT, MLP. Left column of Figure 4.19 repeats the same investigation for 140PU scenario. For $\alpha_{M_{T2}^W}$ and M_{T2}^W calculations for the optimal value is troublesome for ML2 and ML4.

Table 4.1 : BDT, MLP and DNN training results for signal sample ML1, for 50PU and 140PU.

ML1 - BDT						
PU	Variable	(#signal, #backgr.)	Optimal-cut	S/sqrt(S+B)	NSig	NBkg
NoPU	M_{T2}^W	(2082, 408750)	-0.0651	3.64921	1808	243662
	$\alpha_{M_{T2}^W}$	(1920, 253115)	-0.0406	4.32997	1546	125936
	Topness	(2875, 487990)	-0.0371	4.58619	2295	248077
50PU	M_{T2}^W	(2970, 225888)	-0.0215	7.71626	2428	96583
	$\alpha_{M_{T2}^W}$	(2961, 142625)	-0.0294	9.6991	2467	62220
	Topness	(6434, 256402)	-0.0051	14.3549	4640	99841
140PU	M_{T2}^W	(4054, 300635)	0.0198	11.2829	2928	64416
	$\alpha_{M_{T2}^W}$	(4028, 206014)	0.0408	13.0582	2474	33421
	Topness	(7377, 301469)	0.0290	17.8627	4791	67137
ML1 - MLP						
PU	Variable	(#signal, #backgr.)	Optimal-cut	S/sqrt(S+B)	NSig	NBkg
NoPU	M_{T2}^W	(2082, 408750)	0.0040	3.81714	1692 194791	
	$\alpha_{M_{T2}^W}$	(1920, 253115)	0.0072	4.45072	1426	101228
	Topness	(2875, 487990)	0.0038	4.61966	2480	285785
50PU	M_{T2}^W	(2970, 225888)	0.1292	7.85914	2382	89479
	$\alpha_{M_{T2}^W}$	(2961, 142625)	0.0187	9.91909	2355	54004
	Topness	(6434, 256402)	0.1059	14.1453	5350	137698
140PU	M_{T2}^W	(4054, 300635)	0.0239	11.7194	2560	45157
	$\alpha_{M_{T2}^W}$	(4028, 206014)	0.0413	13.4709	2302	26900
	Topness	(7377, 301469)	0.0224	17.8014	5761	98966
ML1 - DNN						
PU	Variable	(#signal, #backgr.)	Optimal-cut	S/sqrt(S+B)	NSig	NBkg
NoPU	M_{T2}^W	(2082, 408750)	0.0039	3.45069	1754	256619.1
	$\alpha_{M_{T2}^W}$	(1920, 253115)	0.0056	4.11508	1697.769	168517.9
	Topness	(2875, 487990)	0.0058	4.30703	2728.847	398694.9
50PU	M_{T2}^W	(2970, 225888)	0.0618	6.47942	2838	189008
	$\alpha_{M_{T2}^W}$	(2961, 142625)	0.0122	7.79721	2624	110628
	Topness	(2970, 225888)	0.0618	6.47942	2838	189008
140PU	M_{T2}^W	(7377, 301469)	0.0263	15.151	6335	168485
	$\alpha_{M_{T2}^W}$	(4028, 206014)	0.0000	8.78894	4028	206014
	Topness	(4054, 300635)	0.0173	9.49729	3202.74	110519

Figure 4.20 compares significance between all possible scenarios with all variables; for ML1, ML4 and ML5. For ML1, highest significance is obtained via BDT method at 140PU when trained with the variable topness. $\alpha_{M_{T2}^W}$ variable shows a higher significance than M_{T2}^W in all methods at 50PU scenario, and at 140PU with BDT and

MLP. However, there is no great significance difference between any method in ML1, as can be observed from the flat distributions.

For ML4, highest significance is obtained via DNN method at 140PU when trained with the variable $\alpha_{M_{T2}^W}$. Unlike ML1, topness here gives the lowest significance at 50PU scenario. Even though the methods for the same variable does not show great difference, when compared variables themselves it can be seen how differently they act on the signal.

Table 4.2 : BDT, MLP and DNN training results for signal sample ML1, for 50PU and 140PU.

ML5 - BDT						
PU	Variable	(#signal, #backgr.)	Optimal-cut	S/sqrt(S+B)	NSig	NBkg
NoPU	M_{T2}^W	(25881, 408750)	0.1563	90.1088	14172	10565
	$\alpha_{M_{T2}^W}$	(23794, 253115)	0.1588	99.2377	14062	6017
	Topness	(3588, 487990)	0.2148	14.0587	622	1335
50PU	M_{T2}^W	(50789, 225888)	0.0876	174.953	37693	8724
	$\alpha_{M_{T2}^W}$	(48393, 142625)	0.0506	182.326	39895	7983
	Topness	(18450, 256402)	0.1434	93.8605	11572	3628
140PU	M_{T2}^W	(73952, 300635)	0.0767	225.856	60426	11153
	$\alpha_{M_{T2}^W}$	(69422, 206014)	0.0580	229.122	59726	8225
	Topness	(30942, 30169)	-0.0062	158.965	27980	3001
ML5 - MLP						
PU	Variable	(#signal, #backgr.)	Optimal-cut	S/sqrt(S+B)	NSig	NBkg
NoPU	M_{T2}^W	(25881, 408750)	0.2890	91.3809	14301	10189
	$\alpha_{M_{T2}^W}$	(23794, 253115)	0.3462	99.8881	14766	7086
	Topness	(3588, 487990)	0.0092	13.1575	1017	4953
50PU	M_{T2}^W	(50789, 225888)	0.0552	174.604	38871	10690
	$\alpha_{M_{T2}^W}$	(48393, 142625)	0.0375	181.586	39337	7592
	Topness	(18450, 256402)	0.0251	86.7085	11182	5449
140PU	M_{T2}^W	(73952, 300635)	0.4101	226.076	61470	12459
	$\alpha_{M_{T2}^W}$	(69422, 206014)	0.5000	132.278	69422	206014
	Topness	(30942, 30169)	0.0810	158.783	28090	3207
ML5 - DNN						
PU	Variable	(#signal, #backgr.)	Optimal-cut	S/sqrt(S+B)	NSig	NBkg
NoPU	M_{T2}^W	(25881, 408750)	0.3657	87.5118	14045	11712
	$\alpha_{M_{T2}^W}$	(23794, 253115)	0.4241	97.2661	14972	8722
	Topness	(3588, 487990)	0.0514	12.336	1190	8116
50PU	M_{T2}^W	(50789, 225888)	0.0517	170.384	37045	10226
	$\alpha_{M_{T2}^W}$	(48393, 142625)	0.0270	178.455	38585	8164
	Topness	(18450, 256402)	0.0141	81.3882	10792	6790
140PU	M_{T2}^W	(73952, 300635)	0.0000	120.83	73952	300635
	$\alpha_{M_{T2}^W}$	(69422, 206014)	0.5659	228.501	60570	9695
	Topness	(30942, 30169)	0.0630	157.044	28470	4395

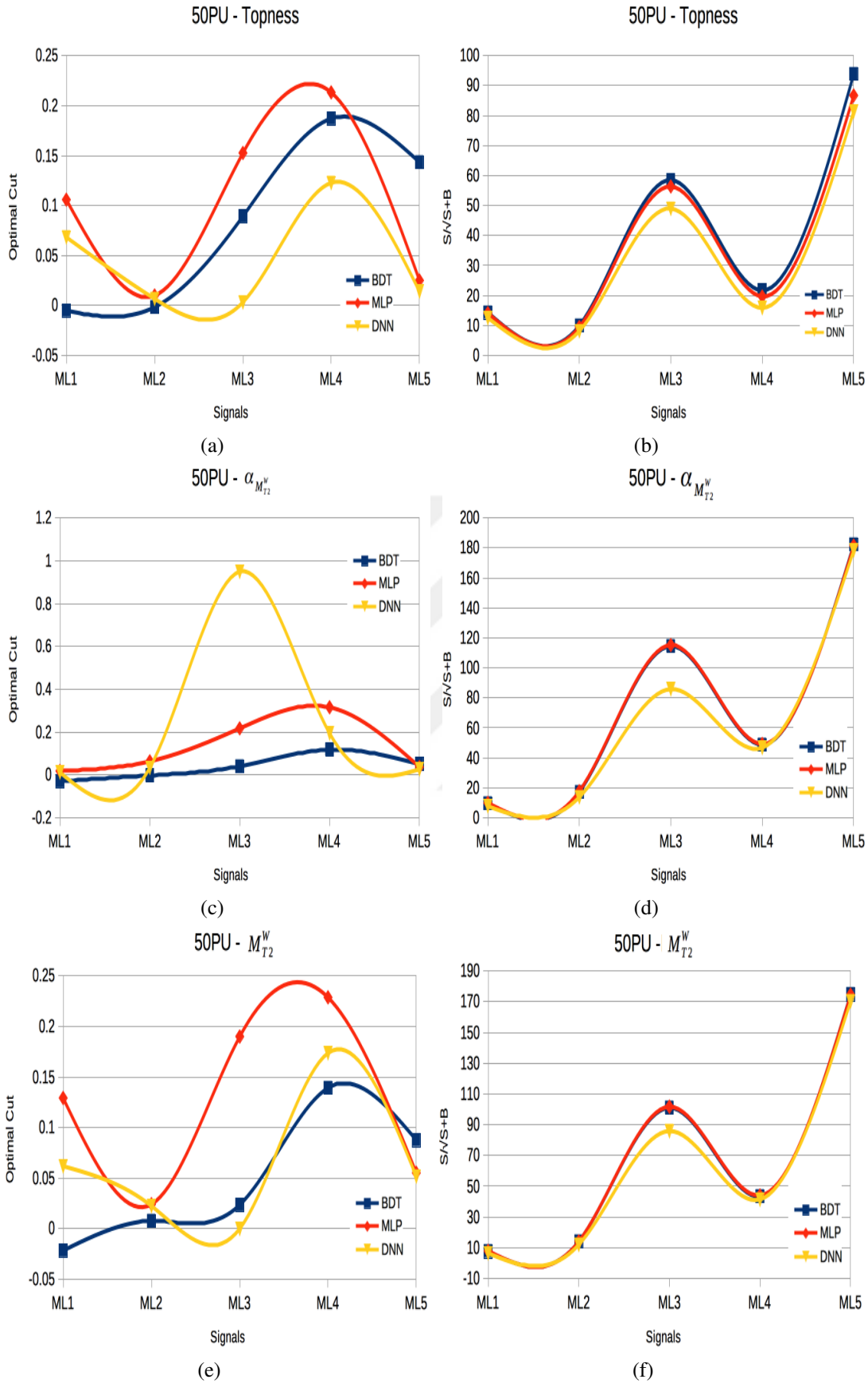


Figure 4.18 : Optimal cut values (a), (c), (e) and (b), (d), (f) $S/\sqrt{S+B}$ for BDT, MLP, and DNN methods for all signals trained with the SM background, for 50PU 300 fb^{-1}.

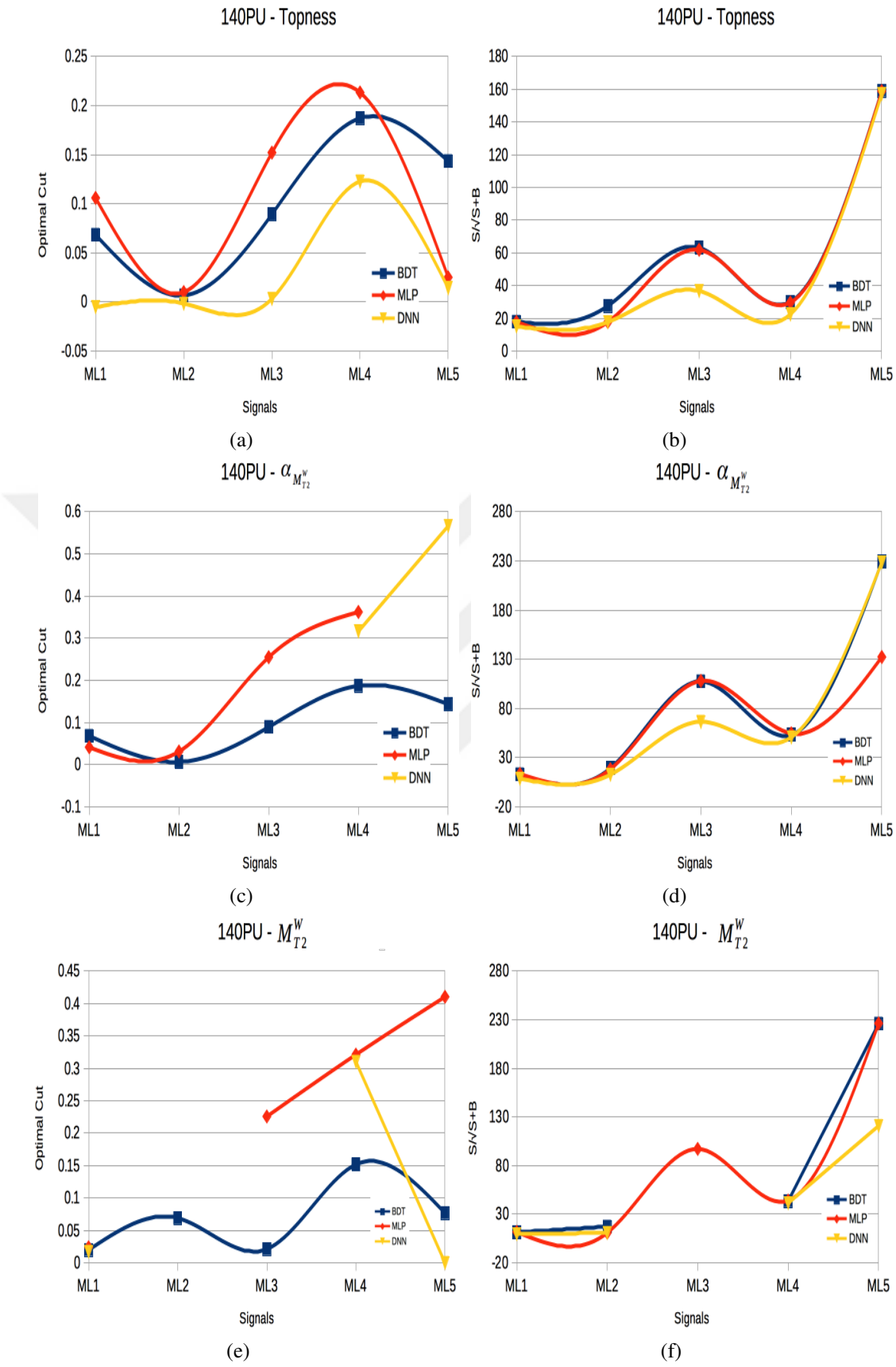
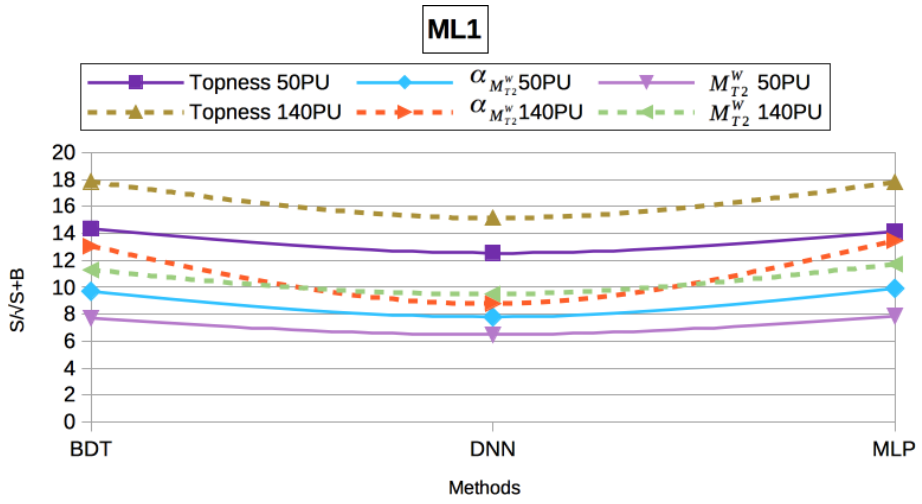
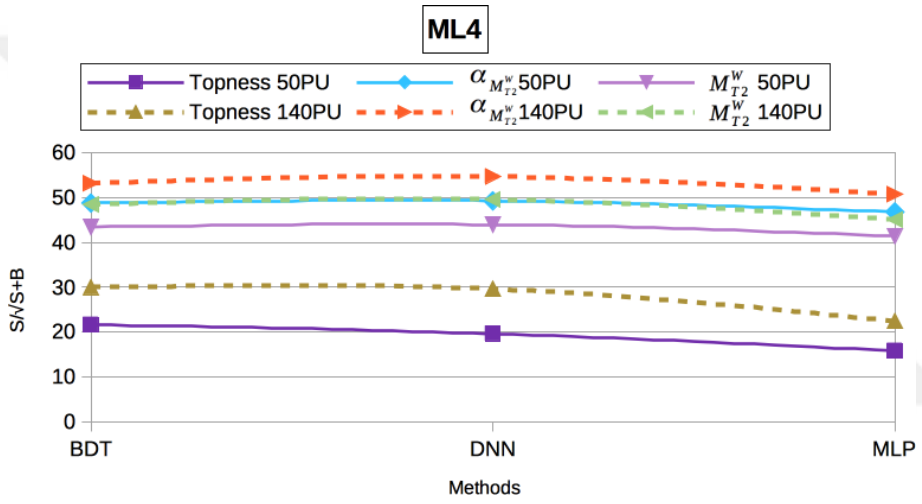


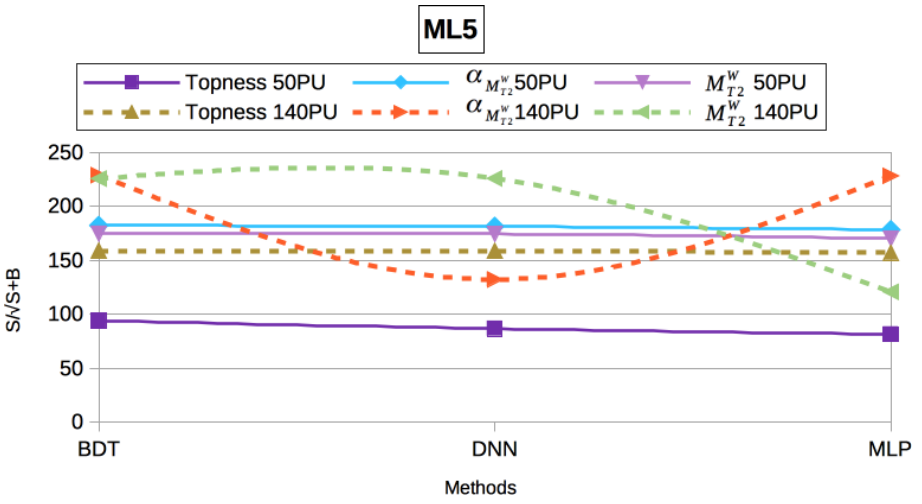
Figure 4.19 : Optimal cut values (a), (c), (e) and (b), (d), (f) $S/\sqrt{S+B}$ for BDT, MLP, and DNN methods for all signals trained with the SM background, for 140PU 3000 fb^{-1}.



(a)



(b)



(c)

Figure 4.20 : Summary of the significance for all training methods, variables, and pileup scenarios. BDT, MLP and DNN methods with variables $\alpha_{M_{T2}^W}$, M_{T2}^W, and topness at 50PU, 300 fb^{-1} and 140PU, 3000 fb^{-1}.

5. CONCLUSION AND OUTLOOK

SUSY offers theoretical solutions to the SM's insufficiencies. So far, supersymmetric particles have not been experimentally observed. Monte Carlo simulations and data analysis help tremendously in this search that has been going on for years. A SUSY model is chosen with certain variables and then the physics can be simulated computationally. It is a tedious job to distinguish SUSY signals amongst a sea of SM backgrounds. Therefore, new approaches and/or analyses in rare regions are always welcomed.

This work covers the direct production of top quark's supersymmetric partner, stop. The production occurs at 14 TeV via proton-proton collision, and single-lepton channel was taken of interest. Deep learning techniques are used at LHC and HL-LHC conditions. LHC is simulated at 50PU with 300 fb^{-1}, whereas HL-LHC conditions are 140PU with 300 fb^{-1}. The importance of these scenarios is to comprehend how pile-up and luminosity makes a difference in SUSY analyses. TMVA has a blackbox approach, the aim is to draw conclusions by looking at the BDT, MLP and DNN responses and interpreting the plots.

5 SUSY signals (ML1-5) falling into different mass regions were examined in this thesis. The signals are subjected to both pile-up and luminosity scenarios. The pile-up is simulated with DELPHES software via various cards. SM background is provided by the Snowmass effort, whereas signal samples are simulated with Madgraph and Pythia.

Signal and background samples are subjected to various baseline and signal selections. The number of signal and background events left are counted and documented in tables for each sample and scenario. Three generic variables H_T, M_T and E_T^{miss} are used in every training. In addition, three kinematic variables are used interchangeably in order to understand which of them performs better in which mass regions. Two of these variables are M_{T2}^W and $Topness$, that have been coiled by scientists and proven to be effective in SUSY analyses. Third interchangeable variable is the $\alpha_{M_{T2}^W}$, a new

variable that is introduced in this thesis. The trainings are made over a total of four variables. The cutflow Tables 5 and 7 are used to deduce variable efficiencies. Among these variables, M_{T2}^{W} and $\alpha_{M_{T2}^{W}}$ show similar performance with LHC and HL-LHC both, mostly coming first and second best in background elimination, respectively. Topness performs best when suppressing VV+jets, and drastically lags behind with $t\bar{t} + Jets$. TMVA offers the chance to use five or more variables as well, though it leads to poor performance with mid range computers.

Figure 1 displays the difference between signal samples. The training are done with the same variable at the same scenario, however BDT performance varies in each region. ML1 proves to be the hardest region in separating signal from the background, whereas performance betters in ML3, ML4 and ML5 respectively. Figure 4.13 (e) and (f) also provides a good comparison between ML1 and ML5. The input variables have different S/B distributions in both signals, ML5 showing more variety whereas ML1 shows the same distribution with the SM background.Trying different and more robust settings when working with ML1 may result in improvement. Signal comparisons show consistency in both 50PU and 140PU scenarios. Main difference between LHC and HL-LHC scenarios is the statistics. Higher luminosity of HL-LHC provides higher statistics. This leads to an increase in the signal events, however, the SM background also rises accordingly.

Tables 4.1 and 3 summarizes results of trainings for all possible scenarios, methods and variables, and some of these results are interpreted in Figures 4.18, 4.19, 4.20. It is understood from these plots that there is important differences between each training, even when applied on the same signal, when the variables, methods and scenarios are changed. The signals can also be compared with each other under same conditions. It is clear that for an efficient training, it is crucial to combine the correct variable, signal, method, and pileup scenarios with each other.

Eventually, no signals could be detected. Possible causes for this may include wrong analysis techniques, or chosing variables that are not effective enough for the used signals. Pileup effects may have caused great uncertanties on signal and background. It is a challenging task to control the missing energy contribution at different pileup scenarios, and it is a question that has been deeply worked on. It is important to find a variable that is independent of pileup. LHC's proton-proton collisions result in

large amounts of background, that might not be suppressed enough. It results in signal getting lost within background, making it even harder to detect.

In the future, it is strongly recommended to apply the same motivation with different final states such as dilepton, trilepton, 4 leptons and hadronic channels. Different classification methods, different signal and background samples with other Monte Carlo tools can be employed. Resulting from these, other variables that are sutiable may be designed accordingly.

REFERENCES

[1] Stephen P. Martin. A supersymmetry primer. In *Advanced Series on Directions in High Energy Physics*, pages 1–153. WORLD SCIENTIFIC, apr 2010.

[2] John Ellis, Mary K. Gaillard, and Dimitri V. Nanopoulos. A historical profile of the higgs boson. In *Advanced Series on Directions in High Energy Physics*, pages 255–274. WORLD SCIENTIFIC, aug 2016.

[3] Quantum diaries. `http://www.quantumdiaries.org/2012/12/10/gluino-higgsino-bingo-2/`. (Accessed on 07/19/2018).

[4] The large hadron collider | cern. `https://home.cern/topics/large-hadron-collider`. (Accessed on 07/11/2018).

[5] Cms | cern. `https://home.cern/about/experiments/cms`. (Accessed on 07/16/2018).

[6] Detector | cms experiment. `https://cms.cern/detector`. (Accessed on 07/16/2018).

[7] M. A. Dobbs et al. Les houches guidebook to monte carlo generators for hadron collider physics. In *Physics at TeV colliders. Proceedings, Workshop, Les Houches, France, May 26-June 3, 2003*, pages 411–459, 2004.

[8] Cms detector layout. `http://www.hephy.at/user/friedl/diss/html/node8.html`. (Accessed on 07/11/2018).

[9] Andreas Hocker et al. Tmva - toolkit for multivariate data analysis. 2007.

[10] David J. Griffiths. *Introduction to Elementary Particles*. Wiley, 2 edition, 2008.

[11] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Phys. Rev.*, 105:1413–1415, Feb 1957.

[12] P Langacker. *The standard model and beyond*. CRC Press, Taylor & Francis Group, Boca Raton, FL, 2017.

[13] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

[14] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.

[15] John F. Donoghue. The effective field theory treatment of quantum gravity. AIP, 2012.

[16] Sundance O Bilson-Thompson, Fotini Markopoulou, and Lee Smolin. Quantum gravity and the standard model. *Classical and Quantum Gravity*, 24(16):3975–3993, jul 2007.

[17] Adam G. Riess, Alexei V. Filippenko, et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 116(3):1009–1038, sep 1998.

[18] Q. R. Ahmad, R. C. Allen, et al. Direct evidence for neutrino flavor transformation from neutral-current interactions in the sudbury neutrino observatory. *Physical Review Letters*, 89(1), jun 2002.

[19] André de Gouvêa. Neutrino mass models. *Annual Review of Nuclear and Particle Science*, 66(1):197–217, oct 2016.

[20] H.P. Nilles. Supersymmetry, supergravity and particle physics. *Physics Reports*, 110(1-2):1–162, aug 1984.

[21] H Haber. The search for supersymmetry: Probing physics beyond the standard model. *Physics Reports*, 117(2-4):75–263, jan 1985.

[22] R. Barbieri, S. Ferrara, and C.A. Savoy. Gauge models with spontaneously broken local supersymmetry. *Physics Letters B*, 119(4-6):343–347, dec 1982.

[23] Sally Dawson, E. Eichten, and C. Quigg. Search for supersymmetric particles in hadron-hadron collisions. *Physical Review D*, 31(7):1581–1637, apr 1985.

[24] J. Wess and B. Zumino. Supergauge transformations in four dimensions. *Nuclear Physics B*, 70(1):39–50, feb 1974.

[25] Professor Howard Baer and Xerxes Tata. *Weak Scale Supersymmetry: From Superfields to Scattering Events*. Cambridge University Press, 2010.

[26] The large hadron collider | cern. `https://home.cern/topics/large-hadron-collider`. (Accessed on 07/11/2018).

[27] Lhc beams. `https://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/beam.htm`. (Accessed on 07/11/2018).

[28] S. Chatrchyan, V. Khachatryan, A.M. Sirunyan, et al. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Physics Letters B*, 716(1):30 – 61, 2012.

[29] G. Aad, T. Abajyan, B. Abbott, et al. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, 716(1):1 – 29, 2012.

[30] Detector | cms experiment. `https://cms.cern/detector`. (Accessed on 07/16/2018).

[31] Michael H. Seymour and Marilyn Marx. Monte carlo event generators. In *LHC Phenomenology*, pages 287–319. Springer International Publishing, aug 2014.

[32] CMS Collaboration. Search for top-squark pair production in the single lepton final state in pp collisions at 8 tev. CMS Physics Analysis Summary CMS-PAS-SUS-13-011, 2013.

[33] S. Ovyn, X. Rouby, and V. Lemaitre. DELPHES, a framework for fast simulation of a generic collider experiment. 2009.

[34] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 0605:026, 2006.

[35] Aram Avetisyan et al. Methods and Results for Standard Model Event Generation at $\sqrt{s}$ = 14 TeV, 33 TeV and 100 TeV Proton Colliders (A Snowmass Whitepaper). In *Proceedings, 2013 Community Summer Study on the Future of U.S. Particle Physics: Snowmass on the Mississippi (CSS2013): Minneapolis, MN, USA, July 29-August 6, 2013*.

[36] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. MadGraph 5: going beyond. *Journal of High Energy Physics*, 2011(6), jun 2011.

[37] Jacob Anderson et al. Snowmass energy frontier simulations. In *Proceedings, 2013 Community Summer Study on the Future of U.S. Particle Physics: Snowmass on the Mississippi (CSS2013): Minneapolis, MN, USA, July 29-August 6, 2013*, 2013.

[38] Aram Avetisyan et al. Snowmass energy frontier simulations using the open science grid (a snowmass 2013 whitepaper). In *Proceedings, 2013 Community Summer Study on the Future of U.S. Particle Physics: Snowmass on the Mississippi (CSS2013): Minneapolis, MN, USA, July 29-August 6, 2013*, 2013.

[39] Rene Brun and Fons Rademakers. ROOT — an object oriented data analysis framework. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 389(1-2):81–86, apr 1997.

[40] Glossary | cms experiment. `http://cms.web.cern.ch/content/glossary#P`. (Accessed on 07/09/2018).

[41] Yang Bai, Hsin-Chia Cheng, Jason Gallicchio, and Jiayin Gu. Stop the top background of the stop search. *Journal of High Energy Physics*, 2012(7), jul 2012.

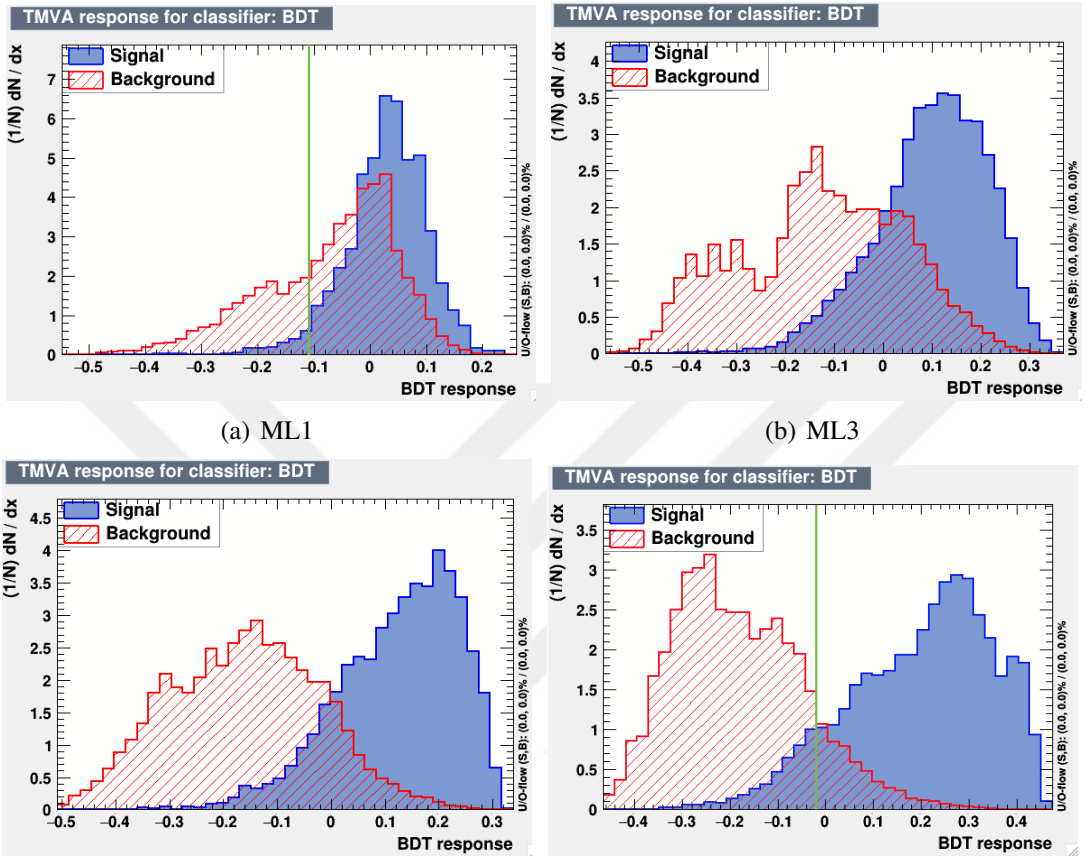
[42] Michael L. Graesser and Jessie Shelton. Hunting mixed top squark decays. *Physical Review Letters*, 111(12), sep 2013.

APPENDICES

APPENDIX A.1 : BDT Response for ML1, ML3, ML4 and ML5.
APPENDIX A.2 : Summary tables for all signal samples, variables and methods.
APPENDIX A.3 : DNN Response for ML1.
APPENDIX A.4 : Results for the SM background variable comparisons.
APPENDIX A.5 : Results for MLP2, MLP3 and MLP4 at 50PU $300 fb^{-1}$
APPENDIX A.6 : Results for MLP2, MLP3 and MLP4 at 140PU $3000 fb^{-1}$
APPENDIX A.7 : Results for the signal sample variable comparisons.
APPENDIX A.8 : Pileup Simulation with Delphes 3.0.10

Appendix A.1

BDT training responses for signal samples ML1, ML3, ML4 and ML5. Trainings are done with $\alpha_{M_{T2}^W}$ variable, for the scenario 50PU at 300 fb^{-1}.



(a) ML1 (b) ML3

(c) ML4 (d) ML5

Figure 1 : BDT responses of ML1, ML3, ML4 and ML5 trained with $\alpha_{M_{T2}^W}$ at 50PU 300fb^{-1}.

Appendix A.2

Summary of the results obtained from the training methods BDT, MLP and DNN for the scenarios 50PU and 140PU, with respect to all variables $\alpha_{M_{T2}^W}$, M_{T2}^W, and topness, for all signal samples ML1, ML2, ML3, ML4 and ML5.

Table 1 : BDT, MLP and DNN training results for signal sample ML2, for 50PU and 140PU.

ML2 - BDT						
PU	Variable	(#signal, #backgr.)	Optimal-cut	S/sqrt(S+B)	NSig	NBkg
50PU	M_{T2}^{W}	(5064, 225888)	0.0077	14.0261	3702	65960
	$\alpha_{M_{T2}^{W}}$	(5070, 142625)	-0.0021	17.2178	3808	45107
	Topness	(4161, 256402)	-0.0016	9.99908	2987	86234
140PU	M_{T2}^{W}	(5993, 300635)	0.0689	17.2178	3353	34561
	$\alpha_{M_{T2}^{W}}$	(5888, 206014)	0.0518	19.3358	3492	29123
	Topness	(9888, 301469)	0.0673	27.5927	5608	35600
ML2 - MLP						
PU	Variable	(#signal, #backgr.)	Optimal-cut	S/sqrt(S+B)	NSig	NBkg
50PU	M_{T2}^{W}	(5064, 225888)	0.0242	14.0454	3620	62808
	$\alpha_{M_{T2}^{W}}$	(5070, 142625)	0.0638	17.7829	2640	19399
	Topness	(4161, 256402)	0.0100	9.51744	3641	142702
140PU	M_{T2}^{W}	(5993, 300635)	0.5000	10.8228	5993	300635
	$\alpha_{M_{T2}^{W}}$	(5888, 206014)	0.0312	18.338	4782	63219
	Topness	(9888, 301469)	0.5000	17.7206	9888	301469
ML2 - DNN						
50PU	M_{T2}^{W}	(5064, 225888)	0.0225	12.0037	3682	90407
	$\alpha_{M_{T2}^{W}}$	(5070, 142625)	0.0338	13.821	3290	53375
	Topness	(4161, 256402)	0.0067	8.16951	4159	255011
140PU	M_{T2}^{W}	(5993, 300635)	0.0000	10.8228	5993	300635
	$\alpha_{M_{T2}^{W}}$	(5888, 206014)	0.0013	12.7992	5888	205737
	Topness	(9888, 301469)	0.0000	17.7206	9888	301469

Table 2 : BDT, MLP and DNN training results for signal sample ML3, for 50PU and 140PU.

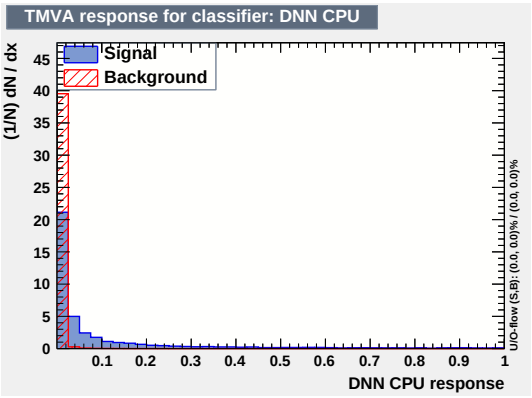
ML3 - BDT						
PU	Variable	(#signal, #backgr.)	Optimal-cut	S/sqrt(S+B)	NSig	NBkg
NoPU	M_{T2}^{W}	(20577, 408750)	0.0171	39.7108	13345	99582
	$\alpha_{M_{T2}^{W}}$	(19908, 253115)	0.0445	50.4344	12102	45477
	Topness	(6618, 487990)	-0.0273	10.7452	4754	190992
50PU	M_{T2}^{W}	(36176, 225888)	0.0236	101.102	27160	45007
	$\alpha_{M_{T2}^{W}}$	(35647, 142625)	0.0404	114.556	25969	25420
	Topness	(21335, 256402)	0.0893	58.5348	10370	21018
140PU	M_{T2}^{W}					
	$\alpha_{M_{T2}^{W}}$	(30654, 206014)	0.0714	107.521	21406	18230
	Topness	(20878, 301469)	0.1019	63.1248	11588	22111
ML3 - MLP						
PU	Variable	(#signal, #backgr.)	Optimal-cut	S/sqrt(S+B)	NSig	NBkg
NoPU	M_{T2}^{W}	(20577, 408750)	0.0633	40.3371	12449	82794
	$\alpha_{M_{T2}^{W}}$	(19908, 253115)	0.0987	51.1966	12524	47318
	Topness	(6618, 487990)	0.0117	10.6607	4532	176190
50PU	M_{T2}^{W}	(36176, 225888)	0.1899	101.759	25850	38682
	$\alpha_{M_{T2}^{W}}$	(35647, 142625)	0.2176	115.235	28227	31773
	Topness	(21335, 256402)	0.1526	56.3892	11439	29709
140PU	M_{T2}^{W}	(31264, 300635)	0.2256	97.0638	20864	25340
	$\alpha_{M_{T2}^{W}}$	(30654, 206014)	0.2549	107.988	21900	19228
	Topness	(20878, 301469)	0.1727	61.8761	11670	239018
ML3 - DNN						
PU	Variable	(#signal, #backgr.)	Optimal-cut	S/sqrt(S+B)	NSig	NBkg
NoPU	M_{T2}^{W}	(20577, 408750)	0.0904	38.7181	12712.62	95093
	$\alpha_{M_{T2}^{W}}$	(19908, 253115)	0.0382	47.7797	12662	57567
	Topness	(6618, 487990)	0.0130	10.2506	4248	167480
50PU	M_{T2}^{W}	(36176, 225888)	0.0001	85.947	25356	61680
	$\alpha_{M_{T2}^{W}}$	(35647, 142625)	0.9494	85.8901	29083	85571
	Topness	(21335, 256402)	0.0033	48.9983	17151	105366
140PU	M_{T2}^{W}					
	$\alpha_{M_{T2}^{W}}$	(30654, 206014)	0.0000	66.4585	28252	152464
	Topness	(20878, 301469)	0.0000	36.7728	20878	301469

Table 3 : BDT, MLP and DNN training results for signal sample ML4, for 50PU and 140PU.

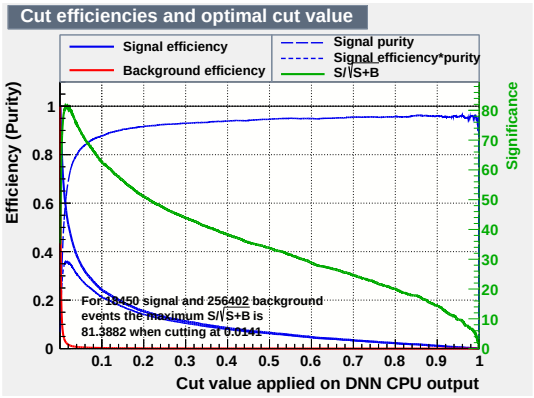
ML4 - BDT						
PU	Variable	(#signal, #backgr.)	Optimal-cut	S/sqrt(S+B)	NSig	NBkg
NoPU	M_{T2}^{W}	(3635, 408750)	0.1192	13.7345	1418	9247
	$\alpha_{M_{T2}^{W}}$	(3362, 253115)	0.1116	17.8701	1626	6653
	Topness	(638, 487990)	0.0555	1.40035	184	17080
50PU	M_{T2}^{W}	(7543, 225888)	0.1391	43.4114	3666	3467
	$\alpha_{M_{T2}^{W}}$	(7210, 142625)	0.1188	48.7642	4178	3163
	Topness	(2998, 256402)	0.1871	21.6804	906	840
140PU	M_{T2}^{W}	(8065, 300635)	0.1516	48.4719	43401	3678
	$\alpha_{M_{T2}^{W}}$	(7591, 206014)	0.1268	53.1963	4777	3286
	Topness	(3919, 301469)	0.1794	30.0025	1506	1015
ML4 - MLP						
PU	Variable	(#signal, #backgr.)	Optimal-cut	S/sqrt(S+B)	NSig	NBkg
NoPU	M_{T2}^{W}	(3635, 408750)	0.0649	14.6325	1480	8756
	$\alpha_{M_{T2}^{W}}$	(3362, 253115)	0.1276	18.7275	1372	3995
	Topness	(638, 487990)	0.0059	1.75606	228	16629
50PU	M_{T2}^{W}	(7543, 225888)	0.2286	43.9662	4015	4323
	$\alpha_{M_{T2}^{W}}$	(7210, 142625)	0.3160	49.2466	4072	2765
	Topness	(2998, 256402)	0.2132	19.6795	958	1412
140PU	M_{T2}^{W}	(8065, 300635)	0.3210	49.5758	4251	3100
	$\alpha_{M_{T2}^{W}}$	(7591, 206014)	0.3620	54.6918	4487	2243
	Topness	(3919, 301469)	0.2886	29.6537	1538	1153
ML4 - DNN						
PU	Variable	(#signal, #backgr.)	Optimal-cut	S/sqrt(S+B)	NSig	NBkg
NoPU	M_{T2}^{W}	(3635, 408750)	0.0680	12.7921	1683	15616
	$\alpha_{M_{T2}^{W}}$	(3362, 253115)	0.0770	17.3012	1472	5767
	Topness	(638, 487990)	0.0019	1.56396	259	27126
50PU	M_{T2}^{W}	(7543, 225888)	0.1736	41.4126	3694	4264
	$\alpha_{M_{T2}^{W}}$	(7210, 142625)	0.1944	46.7993	3948	3169
	Topness	(2998, 256402)	0.1228	15.8843	798	1726
140PU	M_{T2}^{W}	(8065, 300635)	0.3102	45.0852	4165	4368
	$\alpha_{M_{T2}^{W}}$	(7591, 206014)	0.3162	50.7599	4539	3456
	Topness	(3919, 301469)	0.1726	22.4586	1398	2478

Appendix A.3

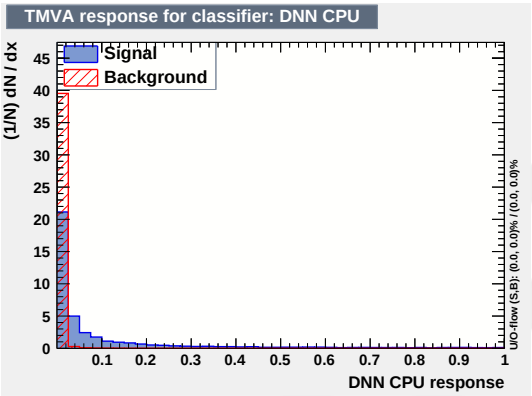
DNN training responses, and significance for given signal and background events for signal sample ML1, trained with the variables M_{T2}^{W}, topness and $\alpha_{M_{T2}^{W}}$.



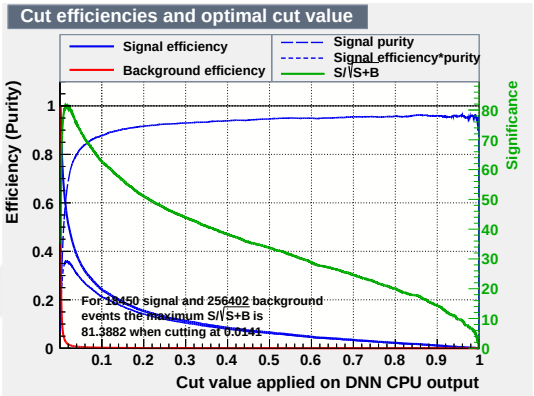
(a) ML5 - M_{T2}^{W}



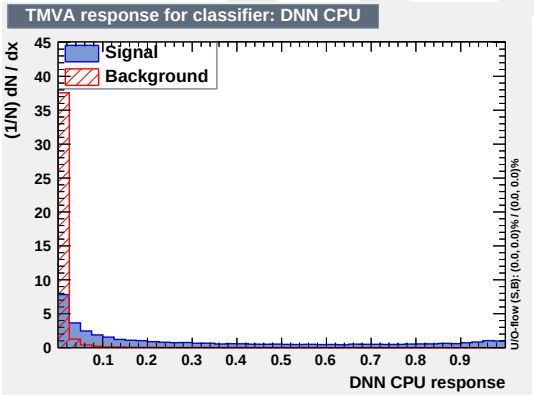
(b) ML5 - M_{T2}^{W}



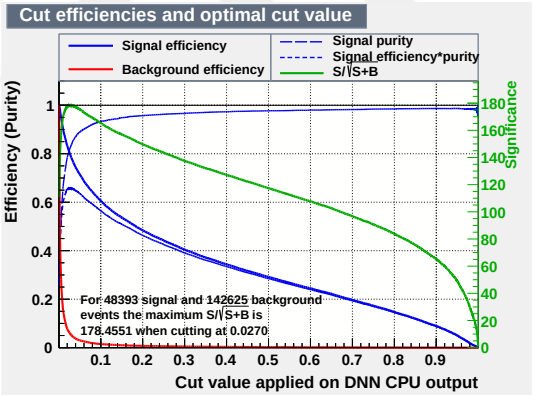
(c) ML5 - Topness



(d) ML5 - Topness



(e) ML5 - $\alpha_{M_{T2}^{W}}$

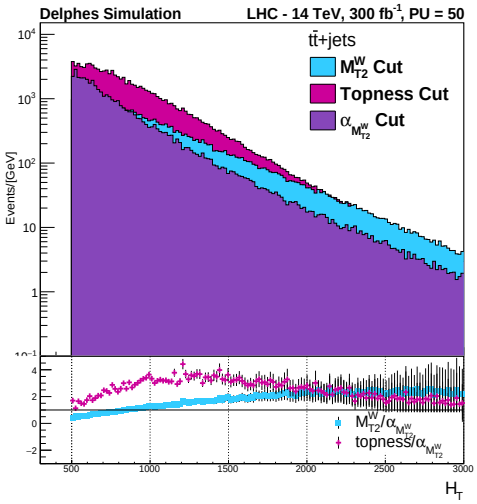


(f) ML5 - $\alpha_{M_{T2}^{W}}$

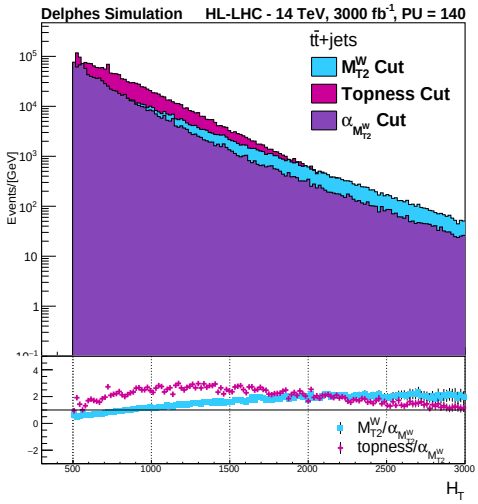
Figure 2 : Left column (a, c) shows TMVA responses for DNN method, for signal ML5 at 50PU, 300 fb^{-1}. Right column (b, d) presents cut efficiencies and optimal cut values.

Appendix A.4

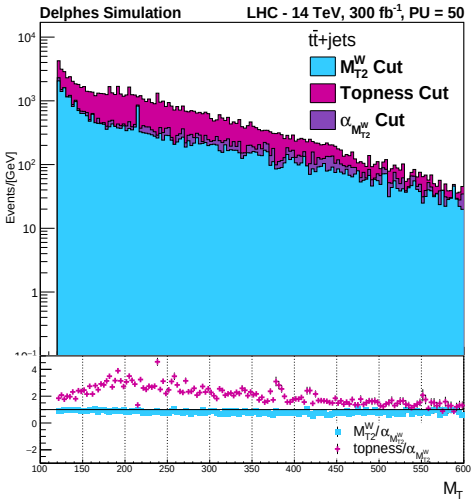
Results for the SM Background Variable Comparisons.



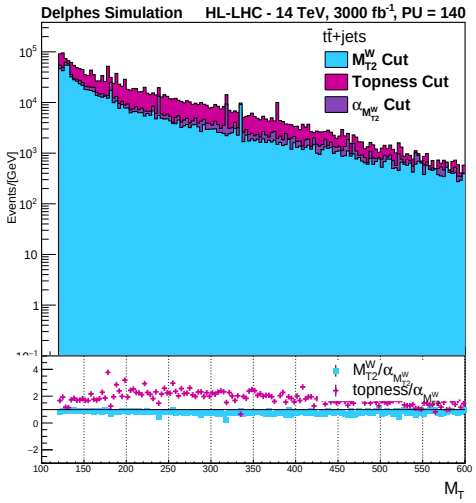
(a) H_T, LHC



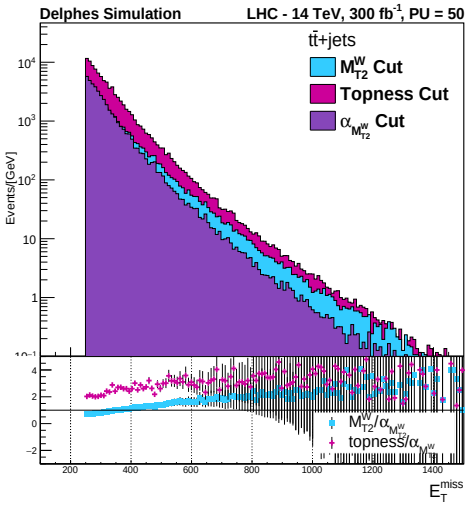
(b) H_T, HL-LHC



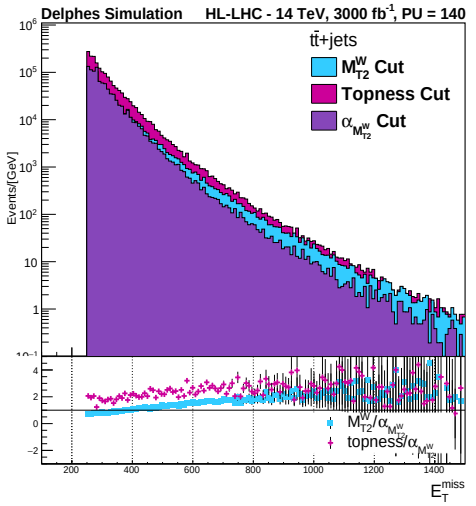
(c) M_T, LHC



(d) M_T, HL-LHC

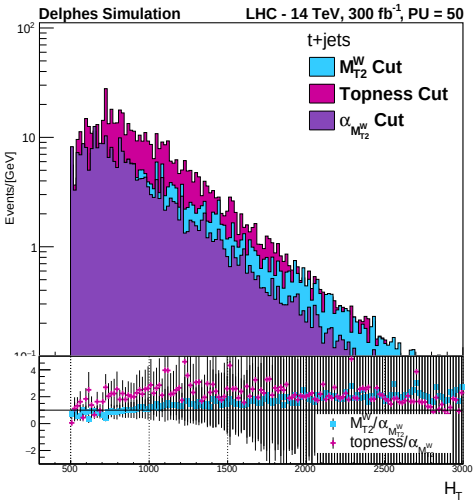


(e) E_T^{miss}, LHC

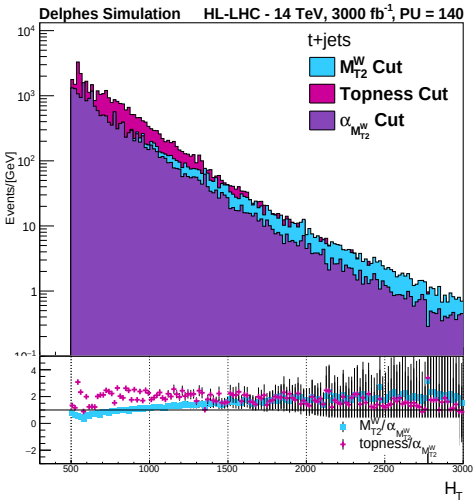


(f) E_T^{miss}, HL-LHC

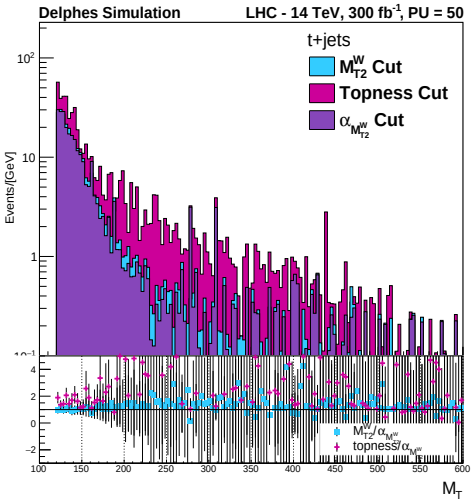
Figure 3 : The effect of the interchangeable variable selections on training variables, for background $t\bar{t} + jets$ at 50PU, 300 fb^{-1} and 140PU, 3000 fb^{-1}.



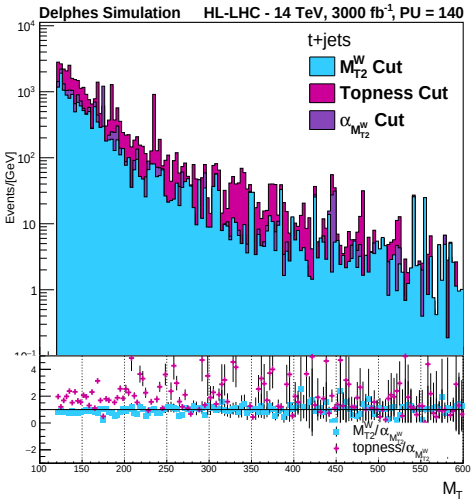
(a) H_T, LHC



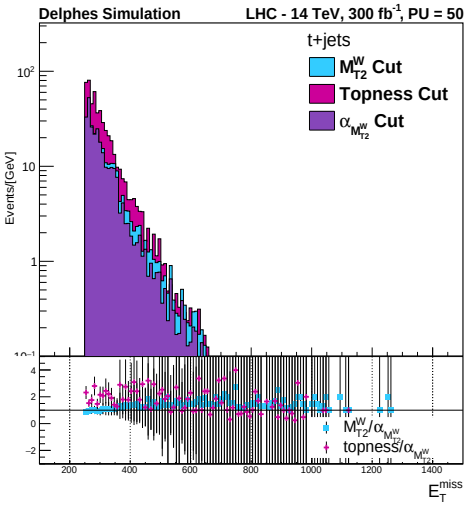
(b) H_T, HL-LHC



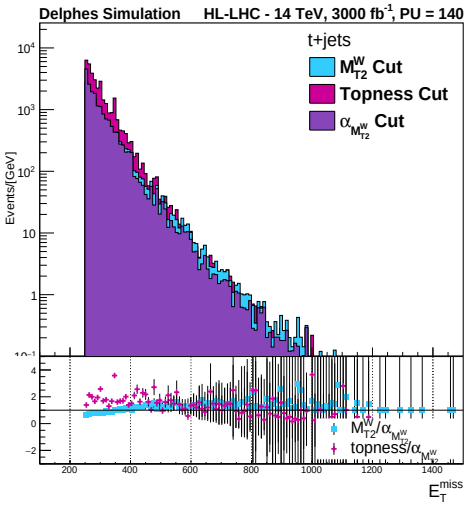
(c) M_T, LHC



(d) M_T, HL-LHC

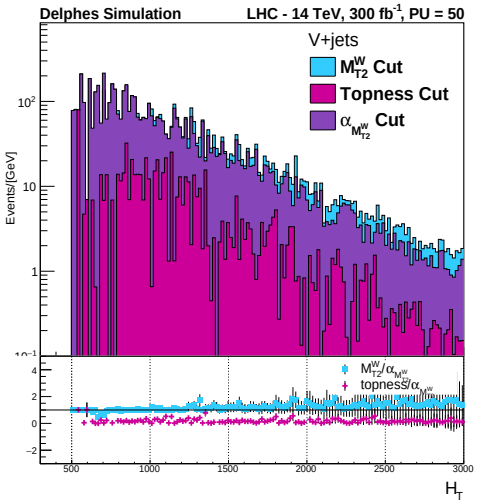


(e) E_T^{miss}, LHC

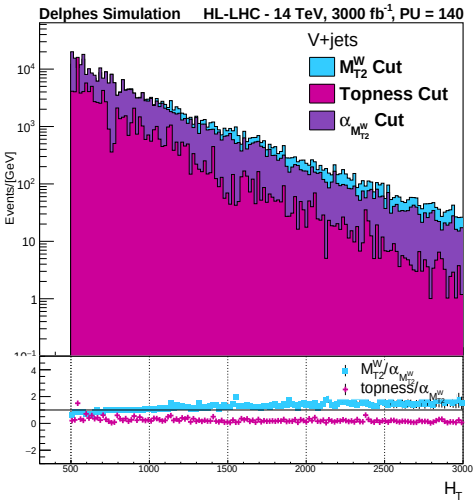


(f) E_T^{miss}, HL-LHC

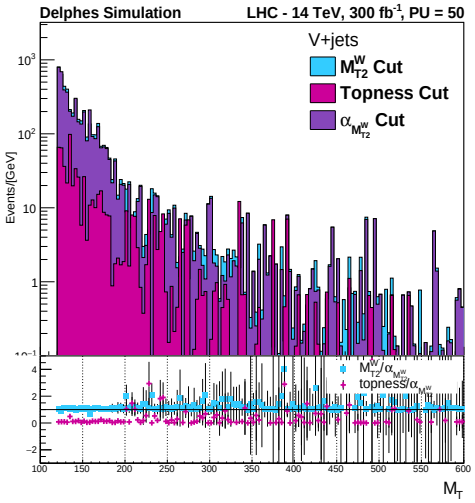
Figure 4 : The effect of the interchangeable variable selections on training variables, for background $t + jets$ at 50PU, 300 fb^{-1} and 140PU, 3000 fb^{-1}.



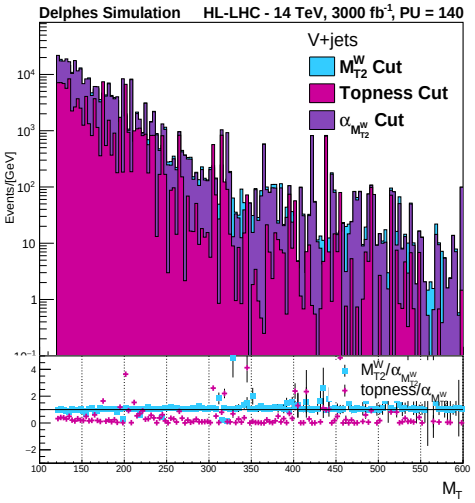
(a) H_T, LHC



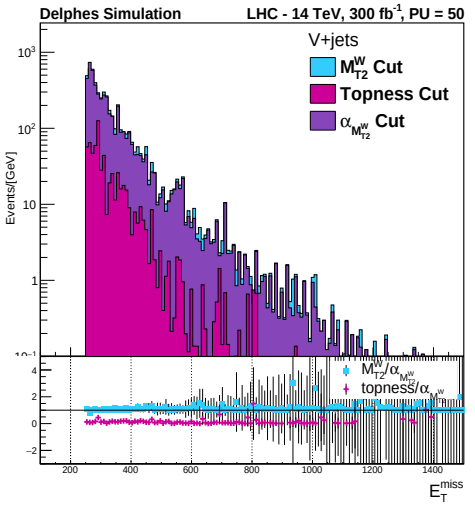
(b) H_T, HL-LHC



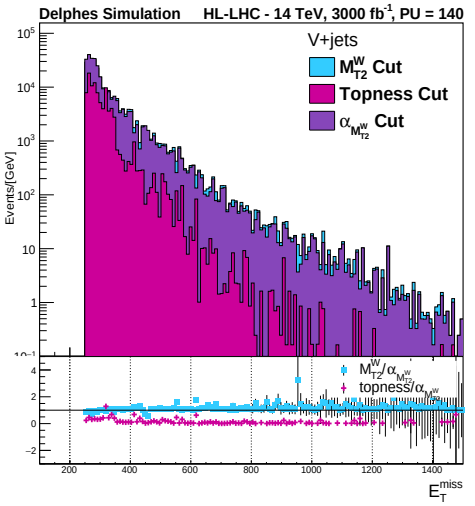
(c) M_T, LHC



(d) M_T, HL-LHC

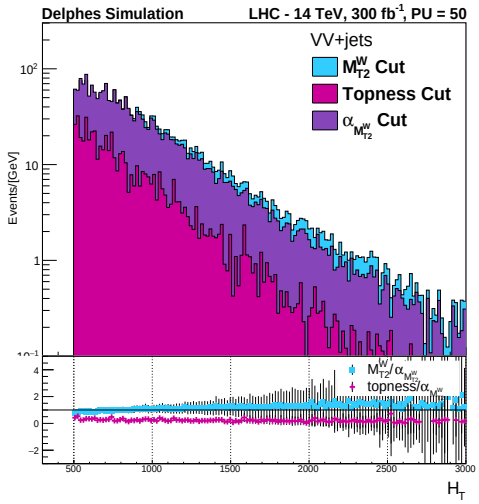


(e) E_T^{miss}, LHC

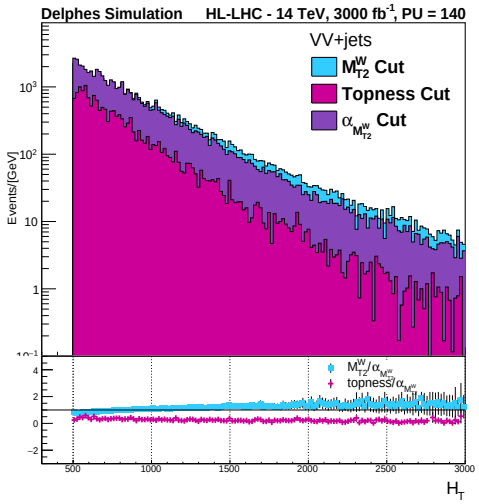


(f) E_T^{miss}, HL-LHC

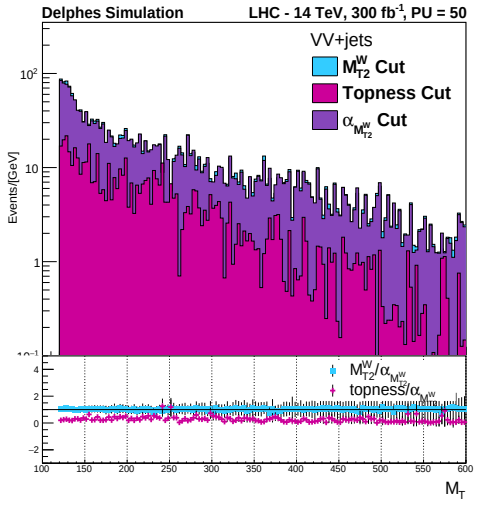
Figure 5 : The effect of the interchangeable variable selections on training variables, for background $V + jets$ at 50PU, 300 fb^{-1} and 140PU, 3000 fb^{-1}.



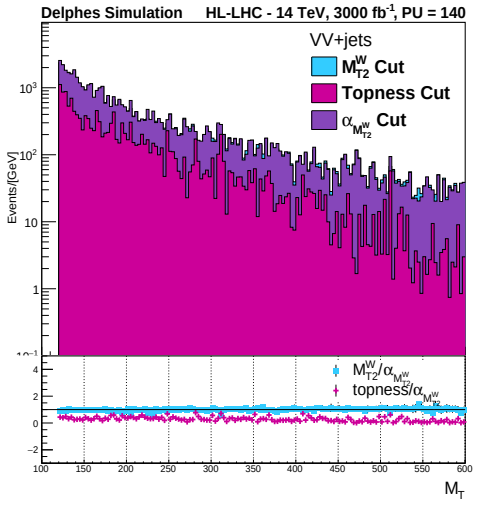
(a) H_T, LHC



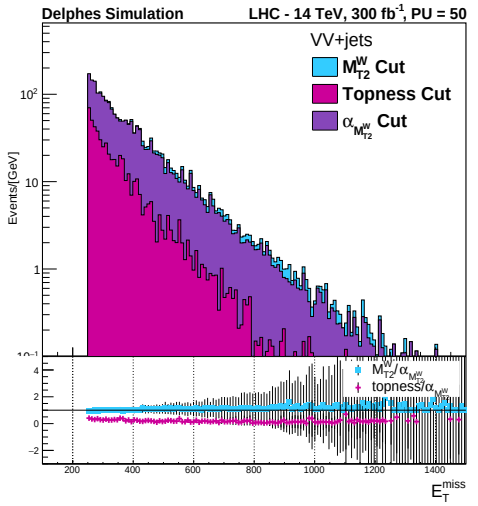
(b) H_T, HL-LHC



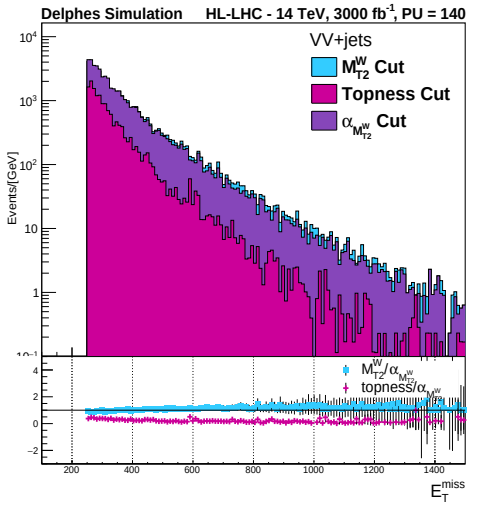
(c) M_T, LHC



(d) M_T, HL-LHC



(e) E_T^{miss}, LHC

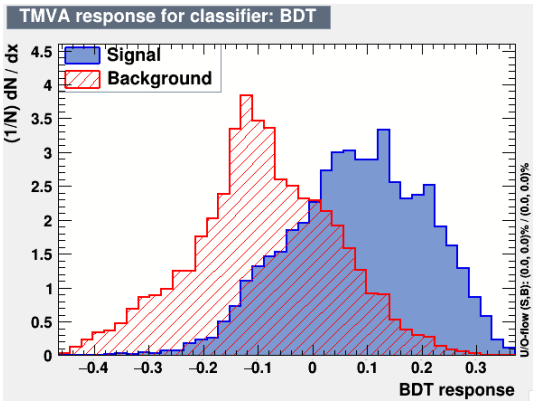


(f) E_T^{miss}, HL-LHC

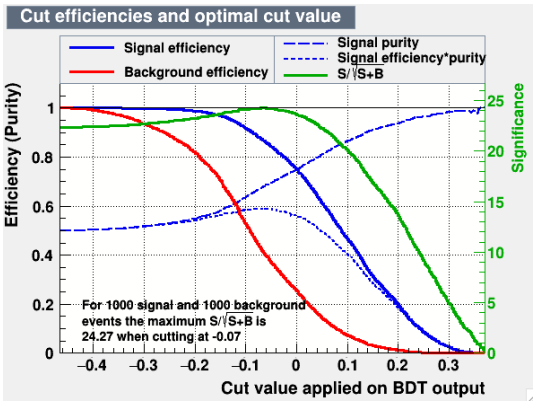
Figure 6 : The effect of the interchangeable variable selections on training variables, for background $VV + jets$ at 50PU, 300 fb^{-1} and 140PU, 3000 fb^{-1}.

Appendix A.5

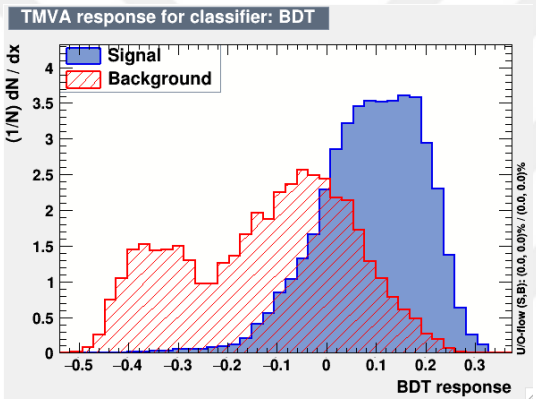
Results for MLP2, MLP3 and MLP4 at 50PU $300 fb^{-1}$



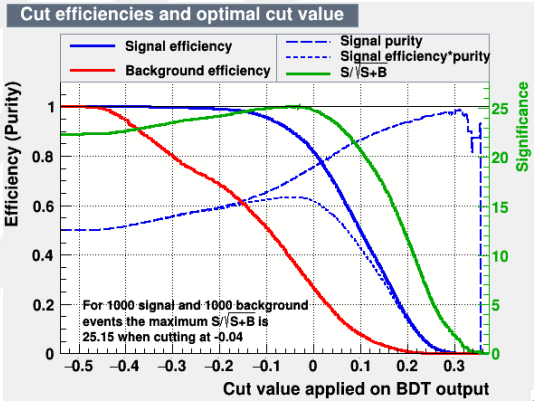
(a) M_{T2}^{W} - ML3



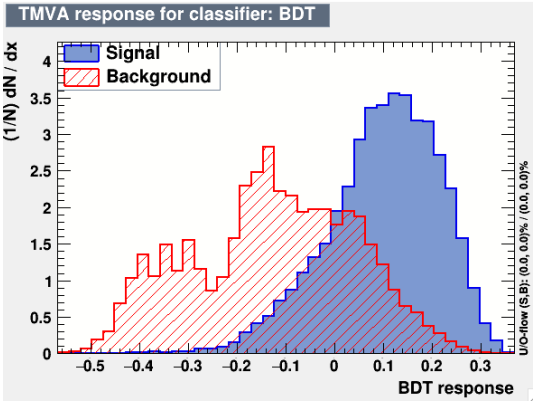
(b) M_{T2}^{W} - ML3



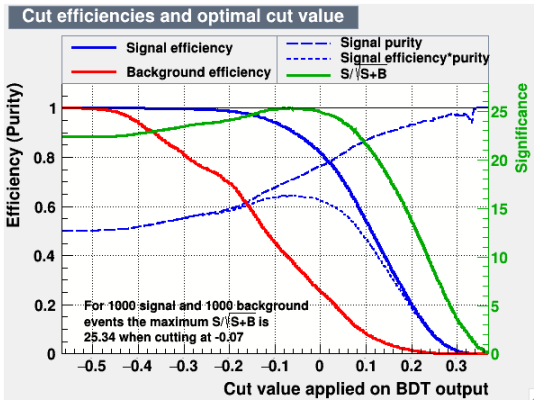
(c) Topness - ML3



(d) Topness - ML3

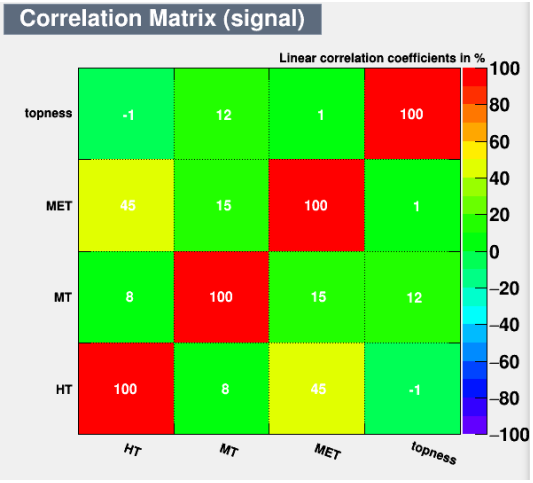


(e) $\alpha_{M_{T2}^{W}}$ - ML3

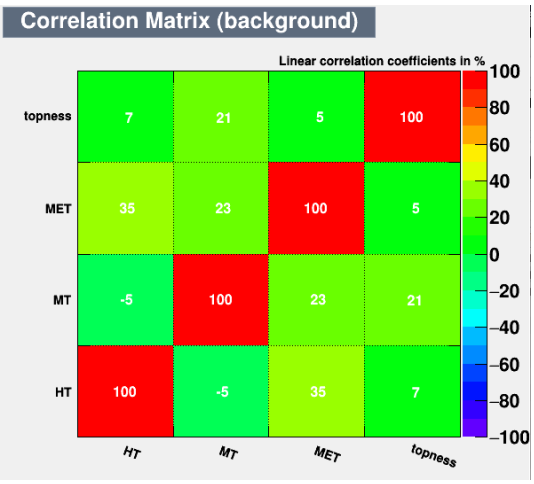


(f) $\alpha_{M_{T2}^{W}}$ - ML3

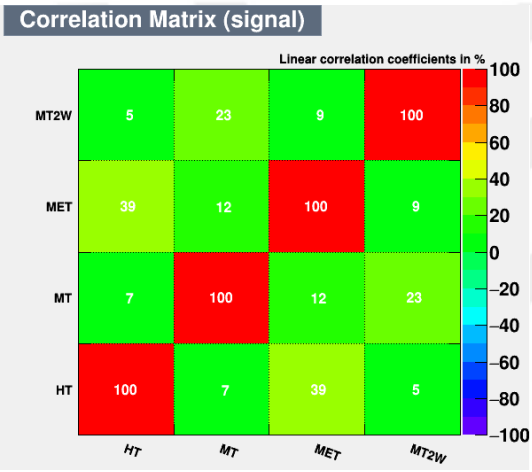
Figure 7 : BDT training results and cut efficiencies for ML3, for 50PU and $300 fb^{-1}$.



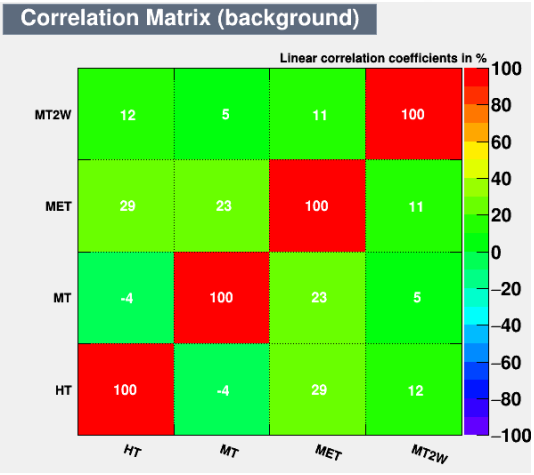
(a) Topness - ML3



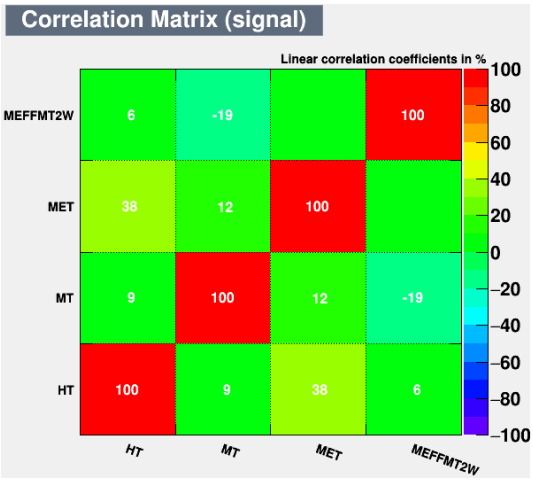
(b) Topness - ML3



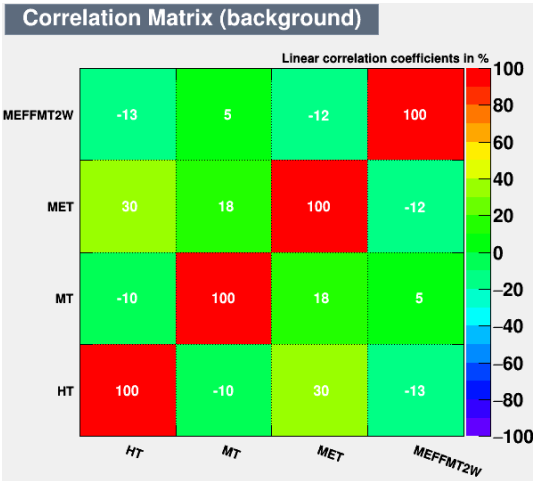
(c) M_{T2}^{W} - ML3



(d) M_{T2}^{W} - ML3

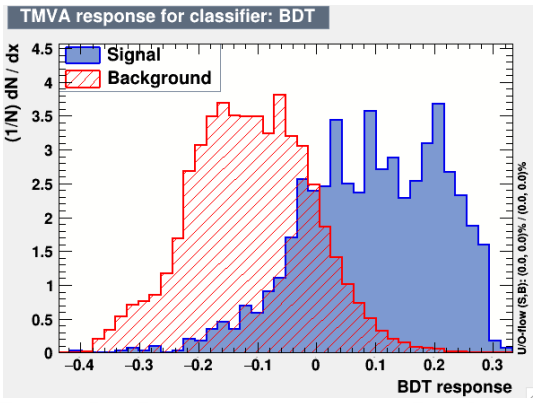


(e) $\alpha_{M_{T2}^{W}}$ - ML3

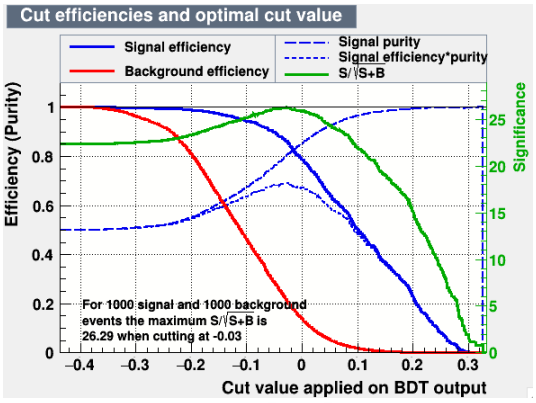


(f) $\alpha_{M_{T2}^{W}}$ - ML3

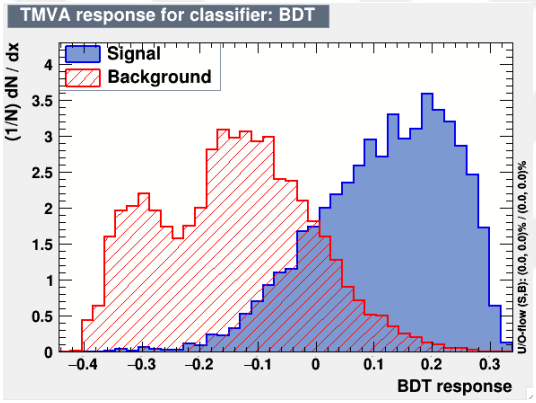
Figure 8 : Correlation between input variables illustrated in matrices for 50PU, 300 fb^{-1}, signal ML3.



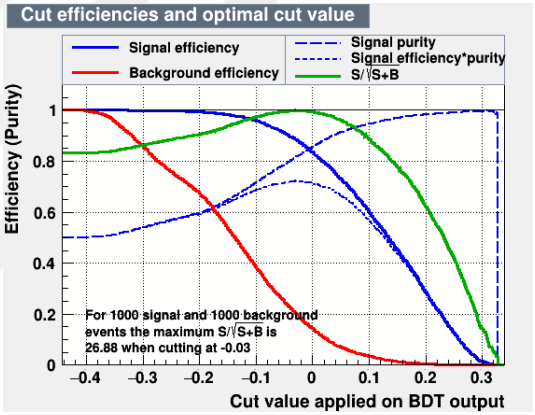
(a) M_{T2}^{W} - ML4



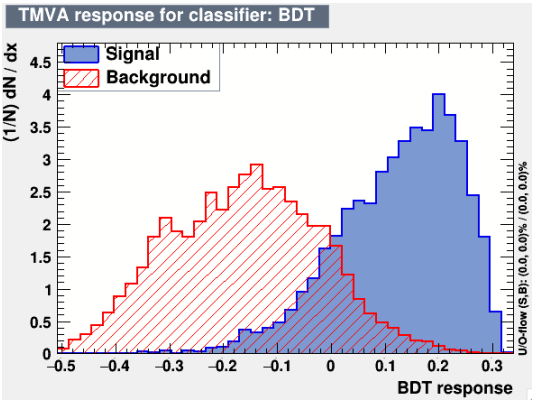
(b) M_{T2}^{W} - ML4



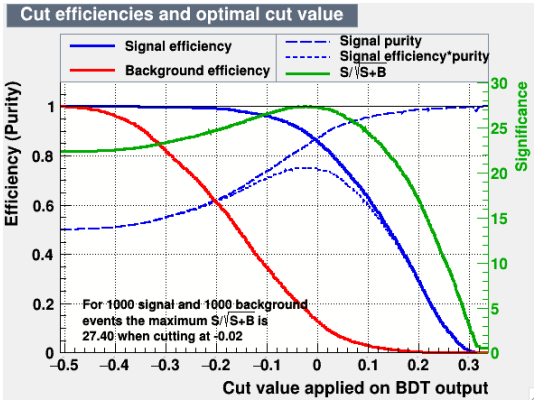
(c) Topness - ML4



(d) Topness - ML4

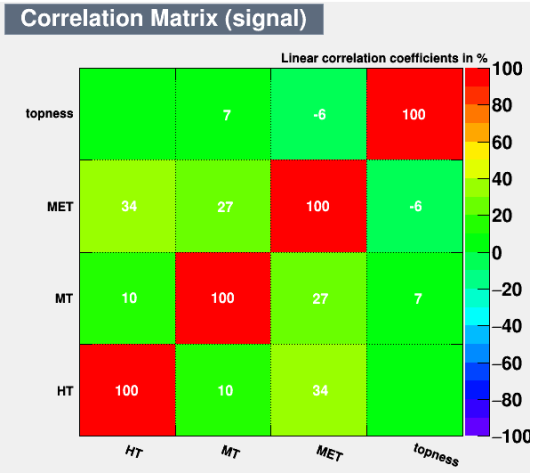


(e) $\alpha_{M_{T2}^{W}}$ - ML4

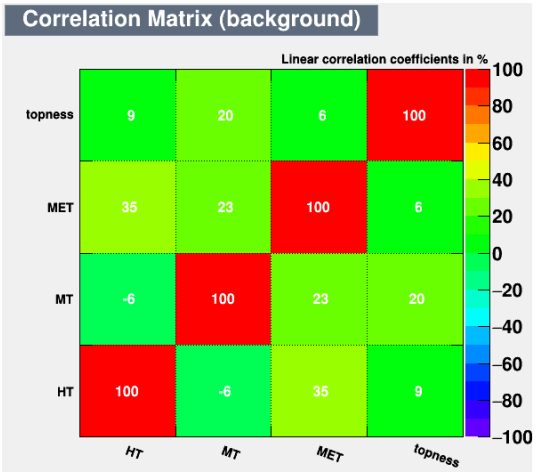


(f) $\alpha_{M_{T2}^{W}}$ - ML4

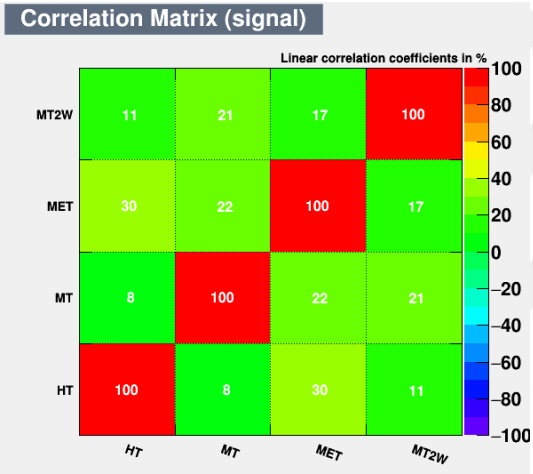
Figure 9 : BDT training results and cut efficiencies for ML4, for 50PU and $300 fb^{-1}$.



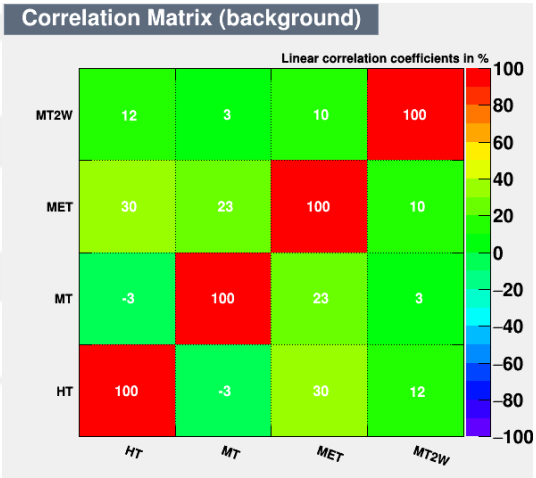
(a) Topness - ML4



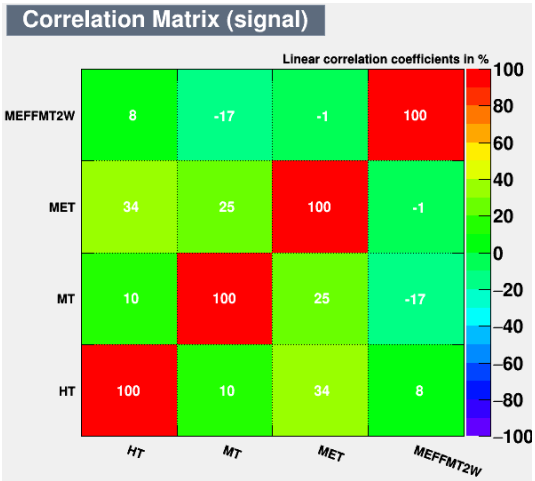
(b) Topness - ML4



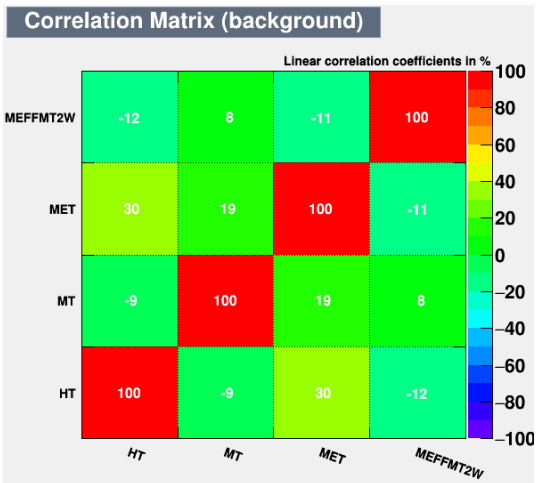
(c) M_{T2}^{W} - ML4



(d) M_{T2}^{W} - ML4



(e) $\alpha_{M_{T2}^{W}}$ - ML4

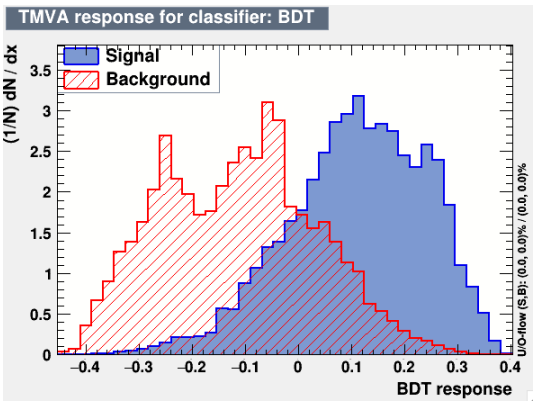


(f) $\alpha_{M_{T2}^{W}}$ - ML4

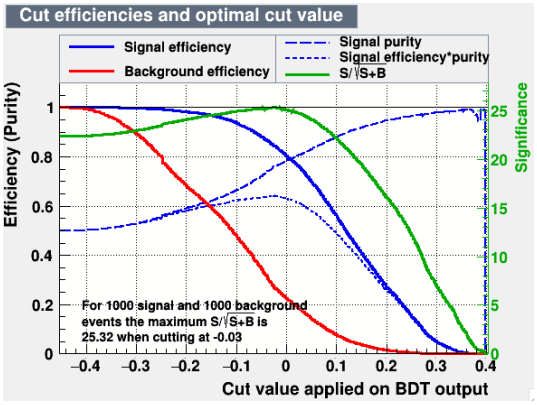
Figure 10 : Correlation between input variables illustrated in matrices for 50PU, 300 fb^{-1}, signal ML4.

Appendix A.6

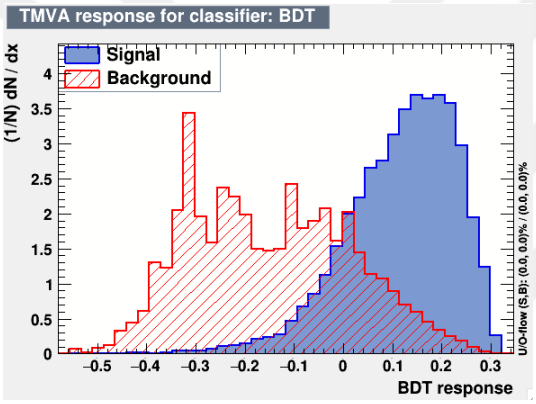
Results for MLP2, MLP3 and MLP4 at 140PU 3000fb^{-1}



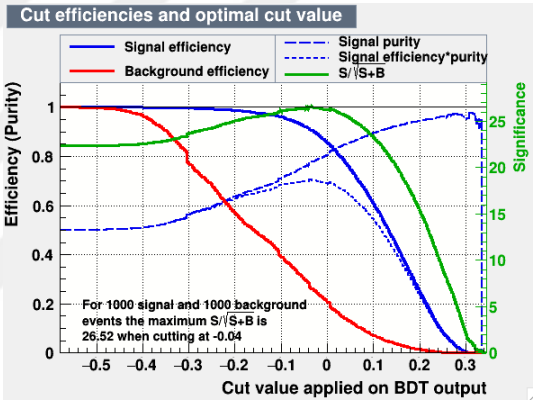
(a) M_{T2}^{W} - ML3



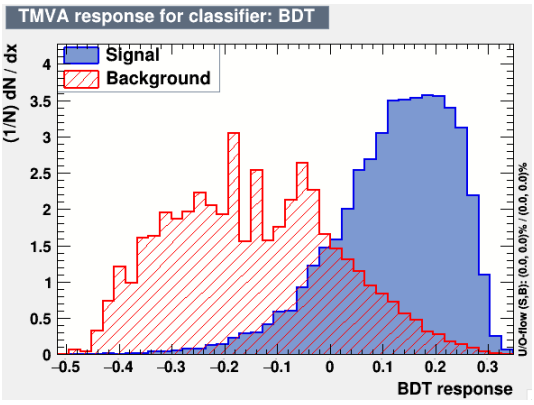
(b) M_{T2}^{W} - ML3



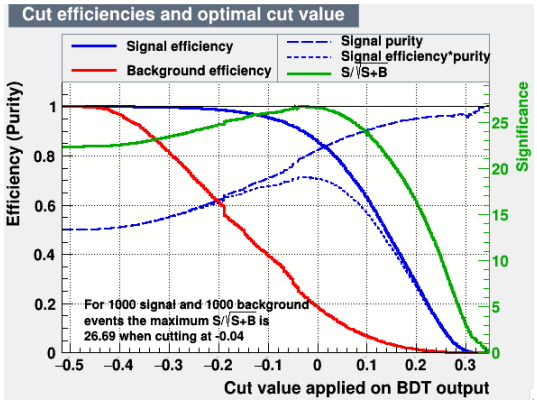
(c) Topness - ML3



(d) Topness - ML3

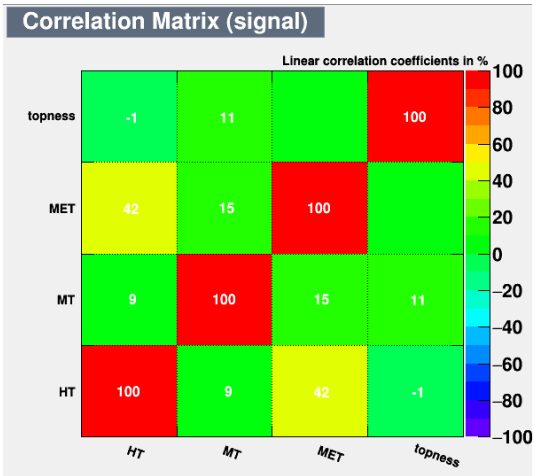


(e) $\alpha_{M_{T2}^{W}}$ - ML3

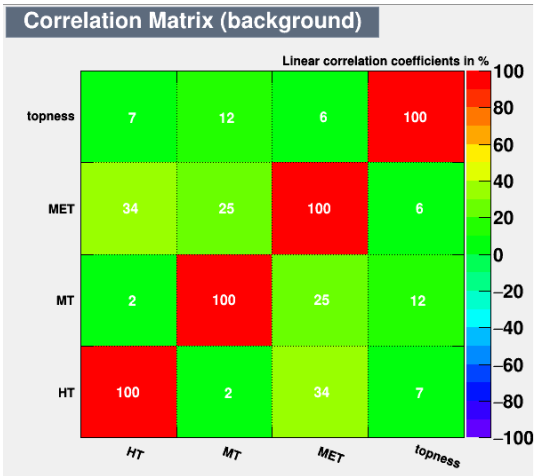


(f) $\alpha_{M_{T2}^{W}}$ - ML3

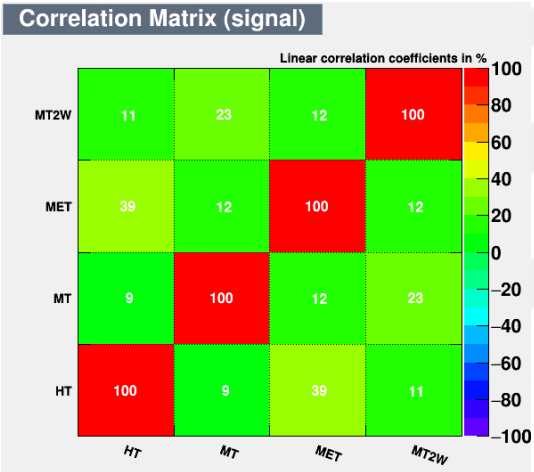
Figure 11 : BDT training results and cut efficiencies for ML3, for 140PU and 3000fb^{-1}.



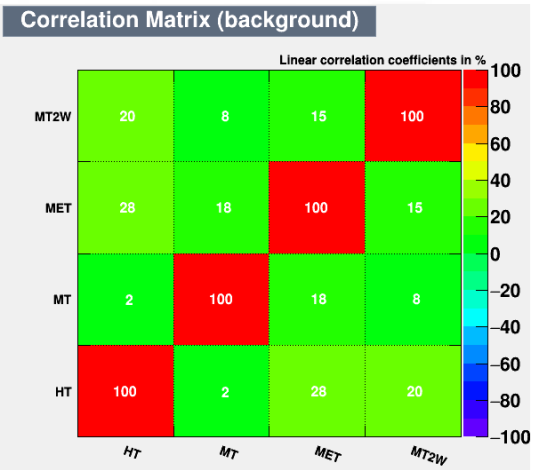
(a) Topness - ML3



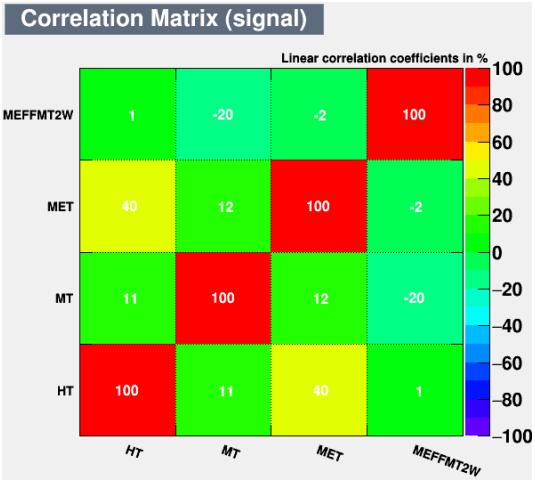
(b) Topness - ML3



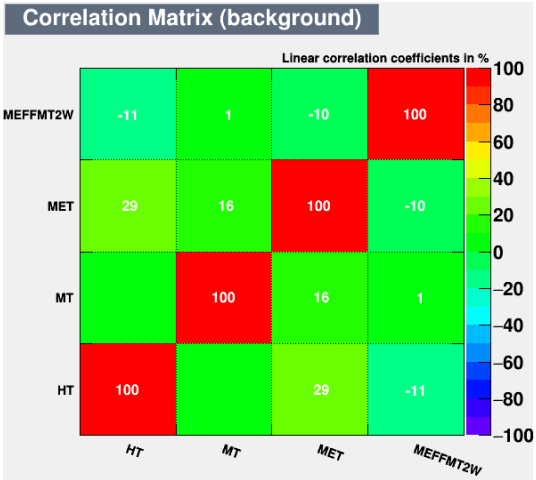
(c) M_{T2}^{W} - ML3



(d) M_{T2}^{W} - ML3

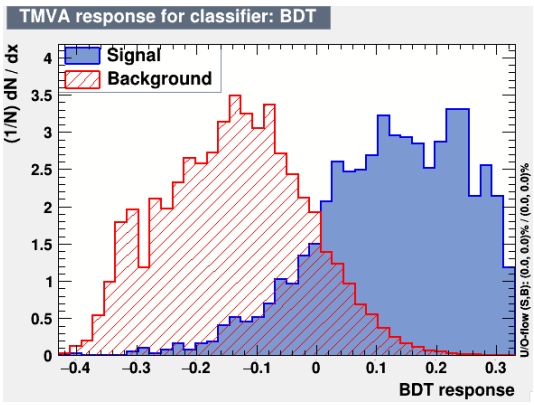


(e) $\alpha_{M_{T2}^{W}}$ - ML3

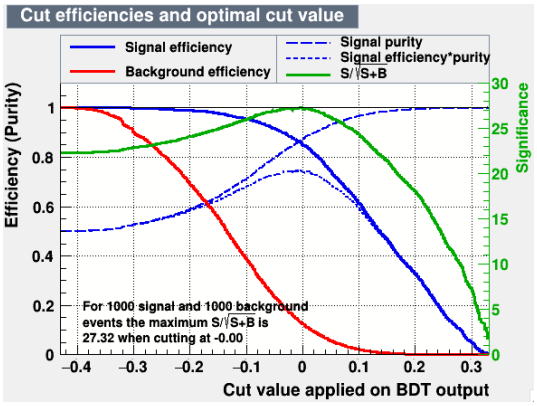


(f) $\alpha_{M_{T2}^{W}}$ - ML3

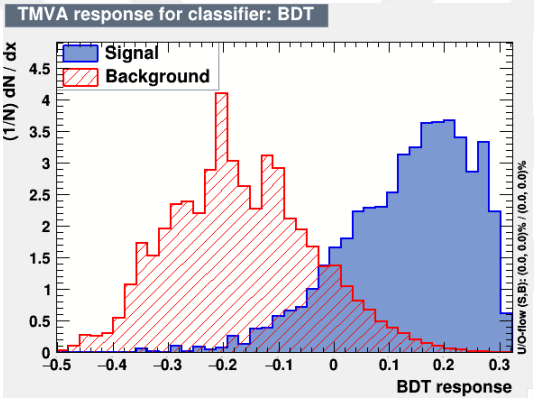
Figure 12 : Correlation between input variables illustrated in matrices for 140PU, 3000 fb^{-1}, signal ML3.



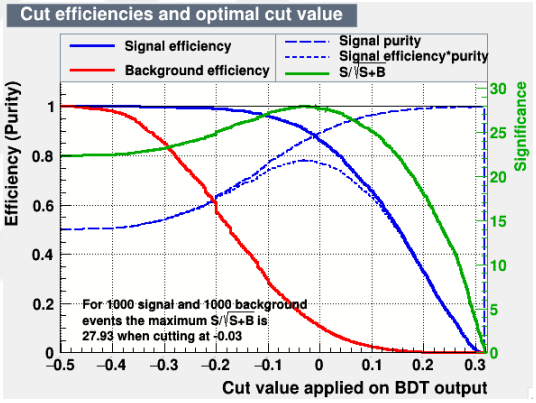
(a) M_{T2}^{W} - ML4



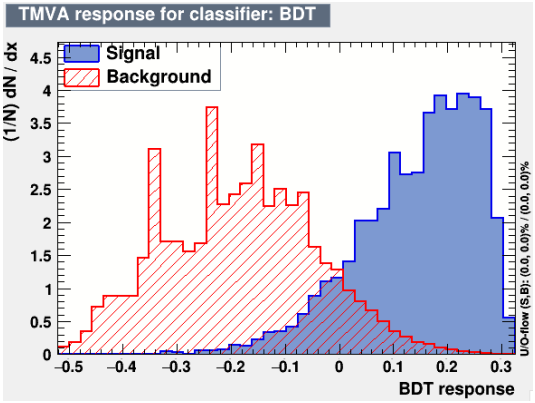
(b) M_{T2}^{W} - ML4



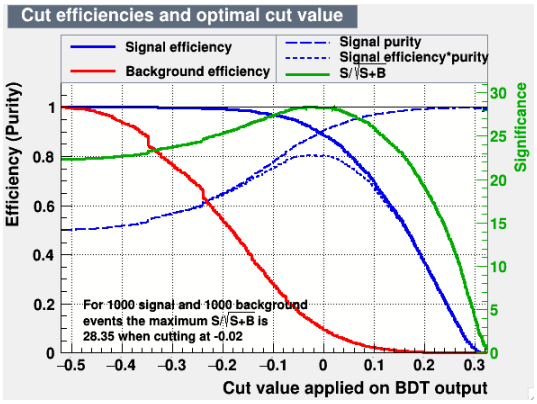
(c) Topness - ML4



(d) Topness - ML4

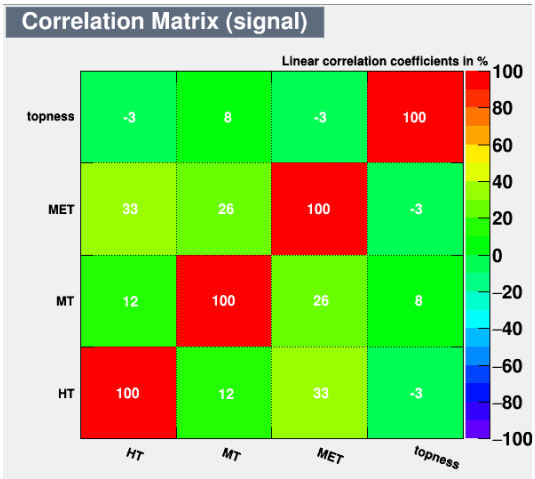


(e) $\alpha_{M_{T2}^{W}}$ - ML4

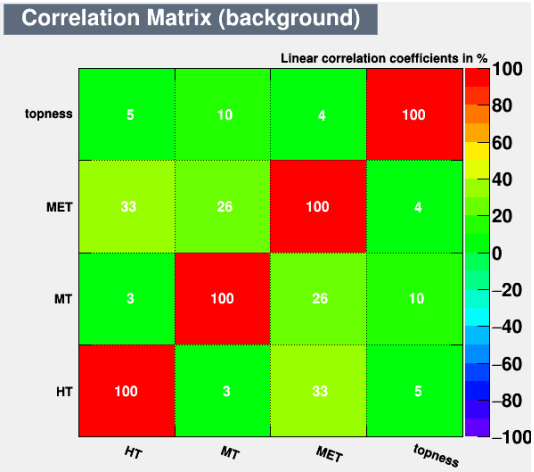


(f) $\alpha_{M_{T2}^{W}}$ - ML4

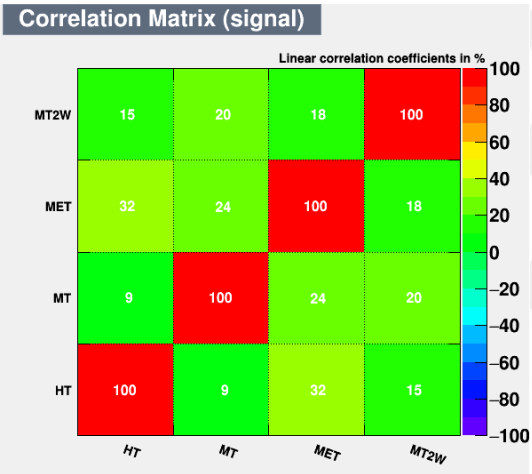
Figure 13 : BDT training results and cut efficiencies for ML4, for 140PU and $3000 fb^{-1}$.



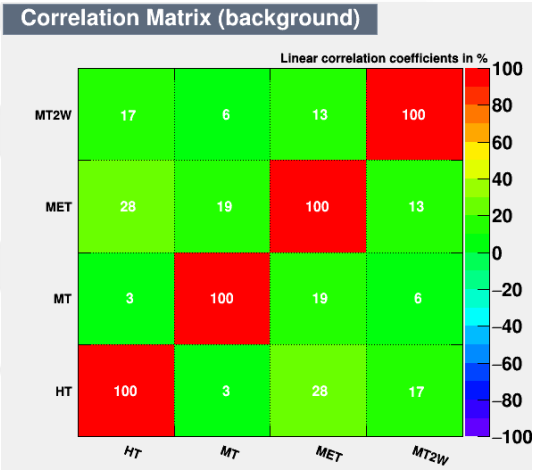
(a) Topness - ML4



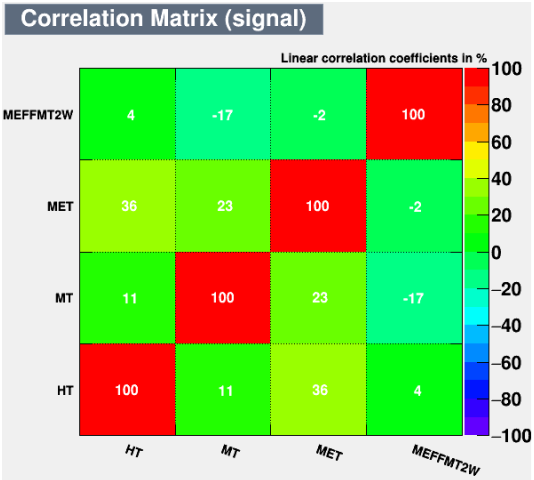
(b) Topness - ML4



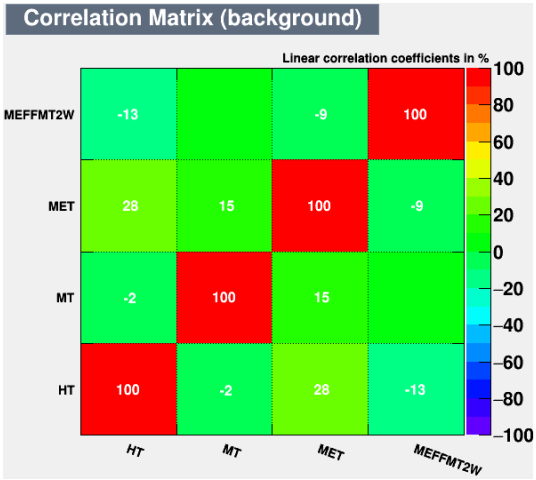
(c) M_{T2}^{W} - ML4



(d) M_{T2}^{W} - ML4



(e) $\alpha_{M_{T2}^{W}}$ - ML4

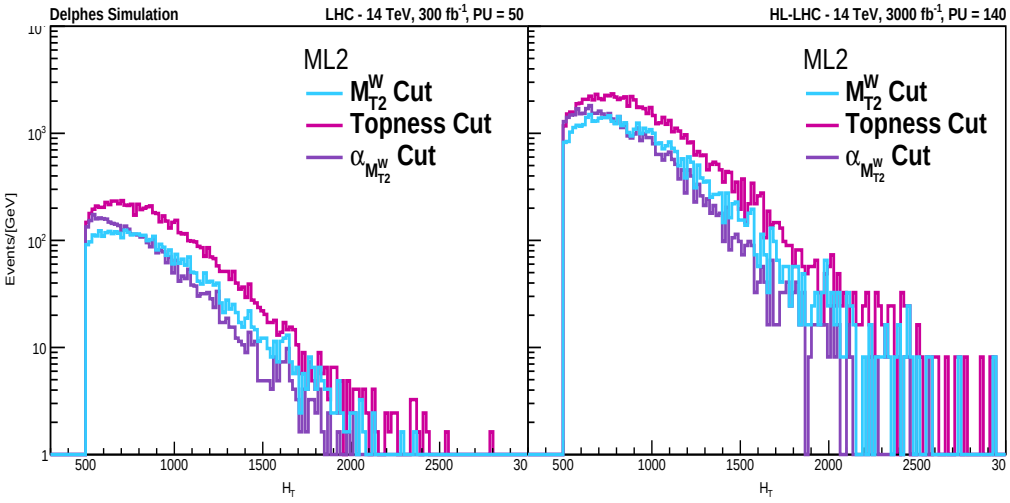


(f) $\alpha_{M_{T2}^{W}}$ - ML4

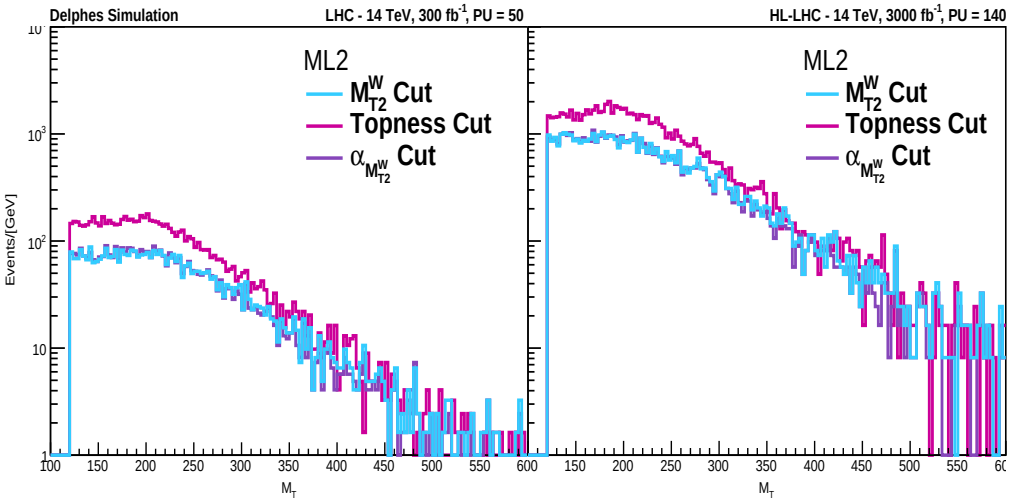
Figure 14 : Correlation between input variables illustrated in matrices for 140PU, 3000 fb^{-1}, signal ML4.

Appendix A.7

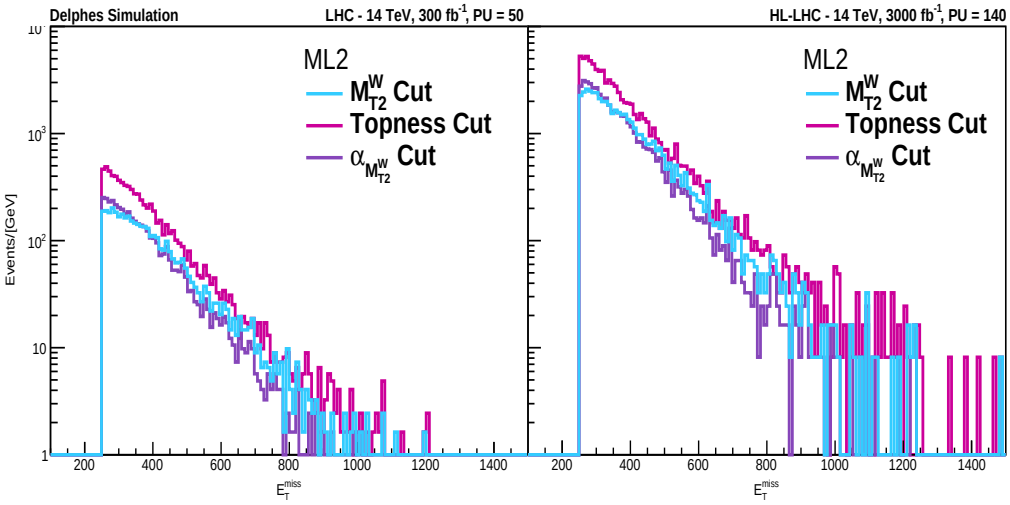
Comparison results for signals ML2, ML3 and ML4, and background samples $V + jets$, $VV + jets$, $t\bar{t} + Jets$ and $Top + jets$.



(a) H_T,

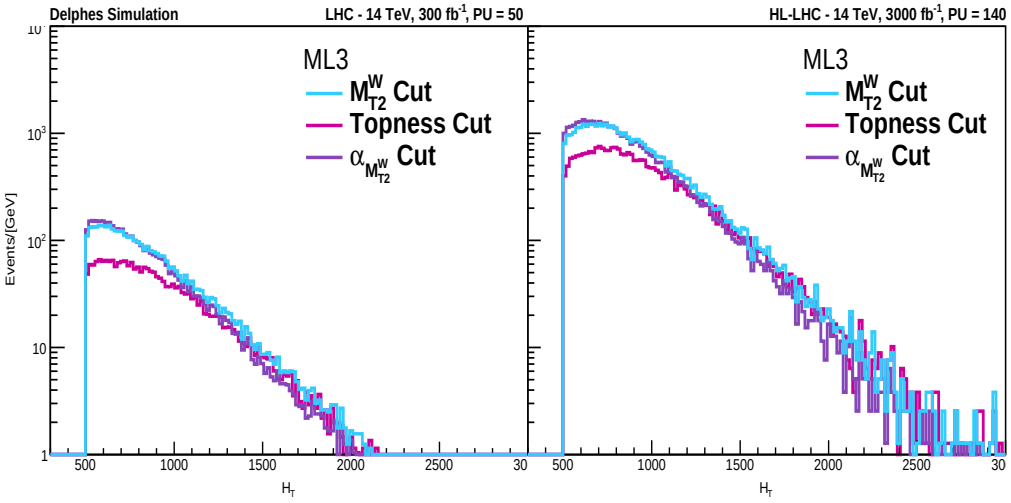


(b) M_T

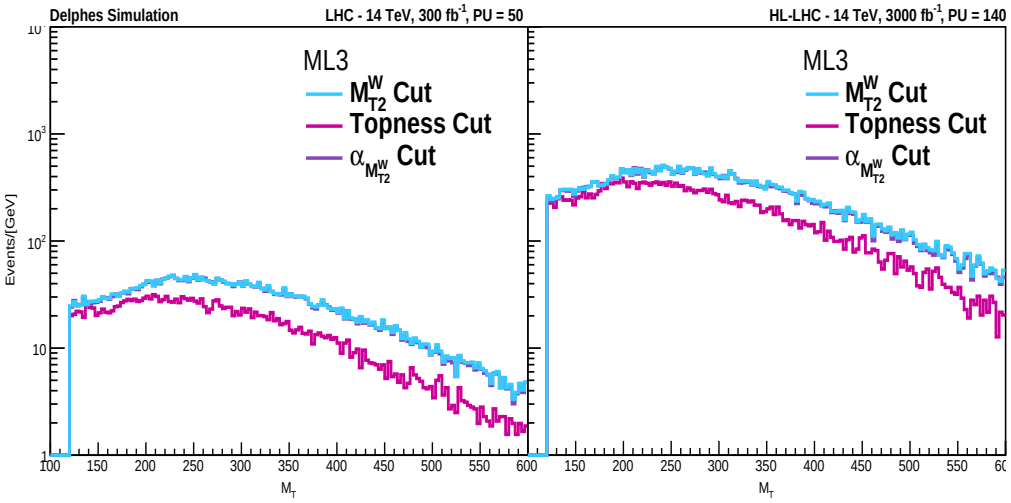


(c) E_T^{miss}

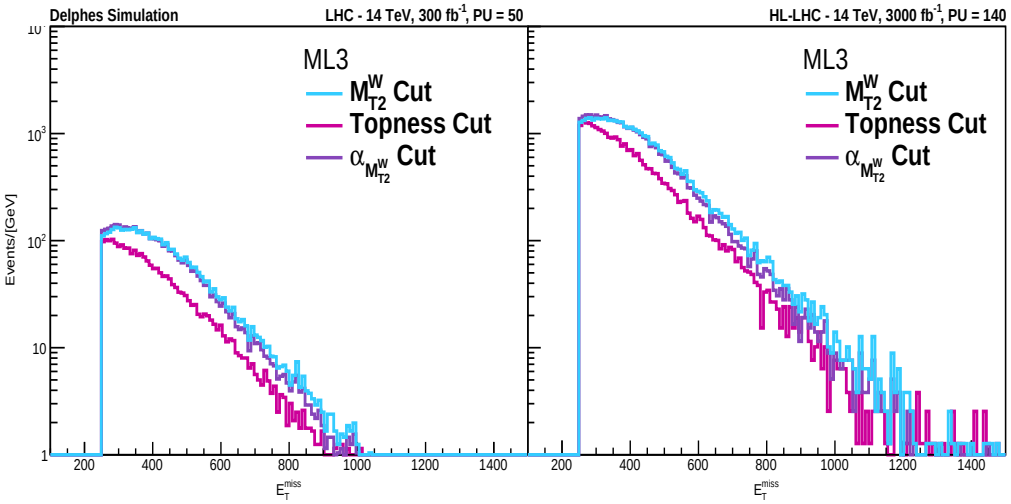
Figure 15 : The effect of the interchangeable variable selections on training variables, for signal ML2 at 50PU, 300 fb^{-1} and 140PU, 3000 fb^{-1}.



(a) H_T,

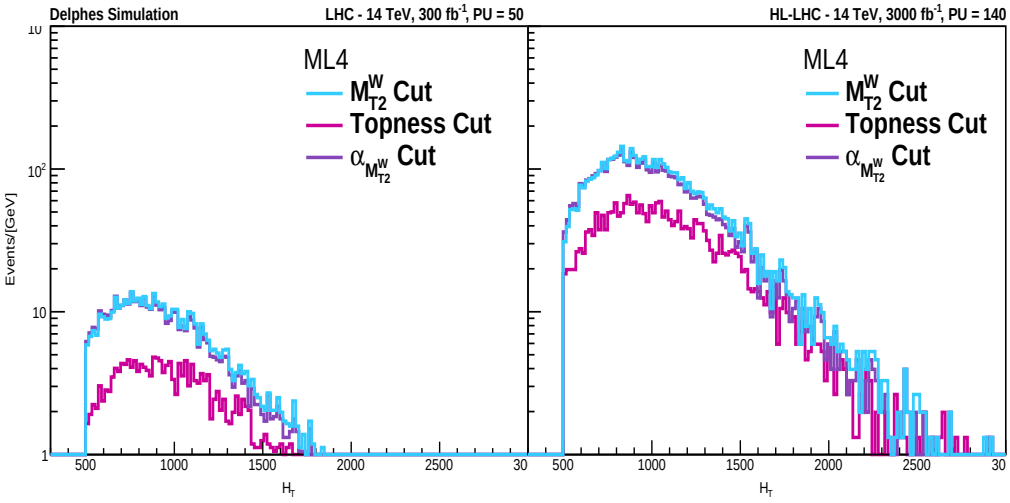


(b) M_T

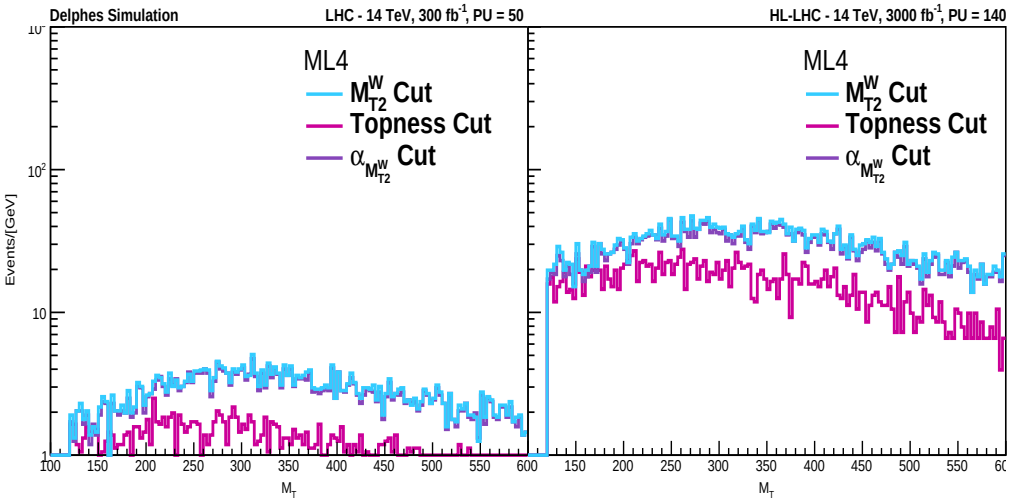


(c) E_T^{miss}

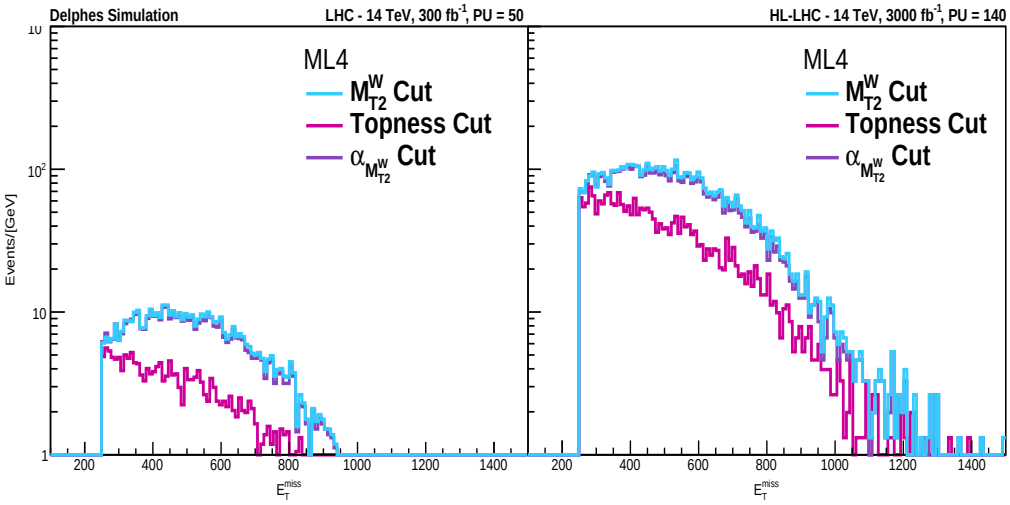
Figure 16 : The effect of the interchangeable variable selections on training variables, for signal ML3 at 50PU, 300 fb^{-1} and 140PU, 3000 fb^{-1}.



(a) H_T,

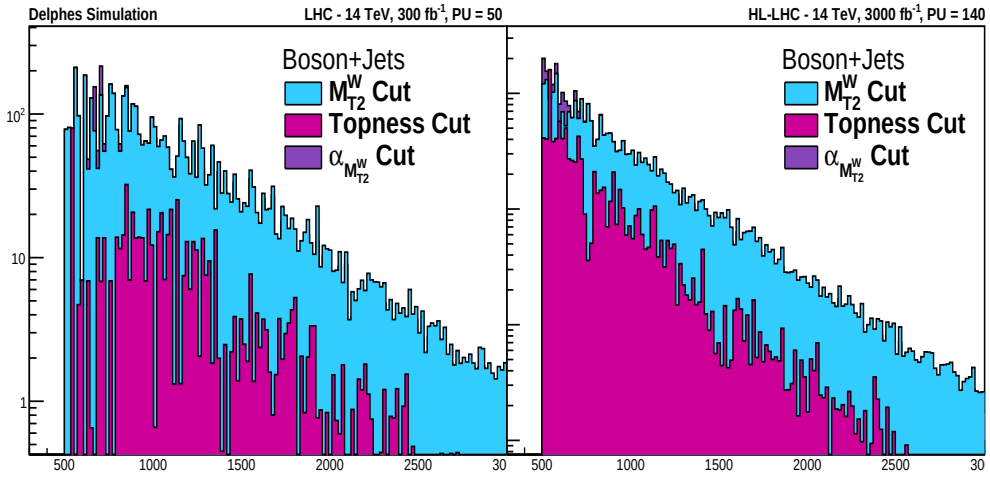


(b) M_T

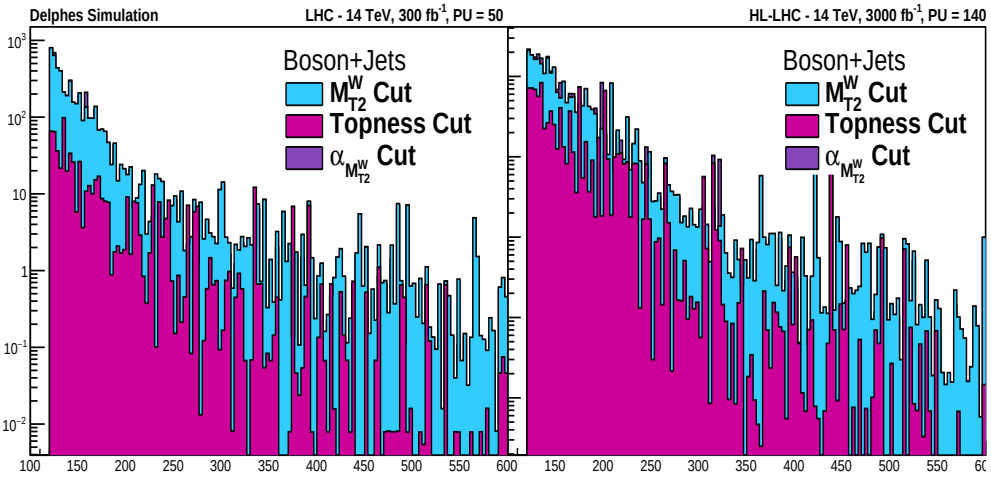


(c) E_T^{miss}

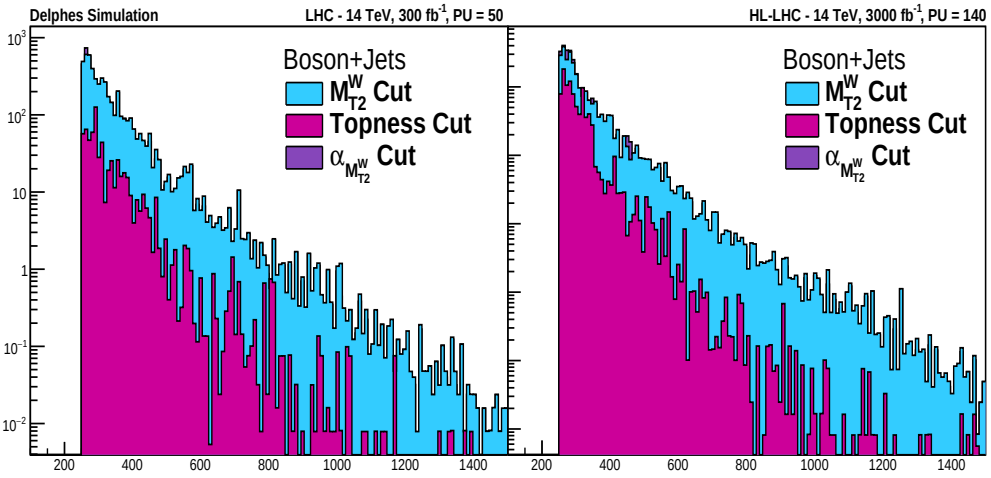
Figure 17 : The effect of the interchangeable variable selections on training variables, for signal ML4 at 50PU, 300 fb^{-1} and 140PU, 3000 fb^{-1}.



(a) H_T,

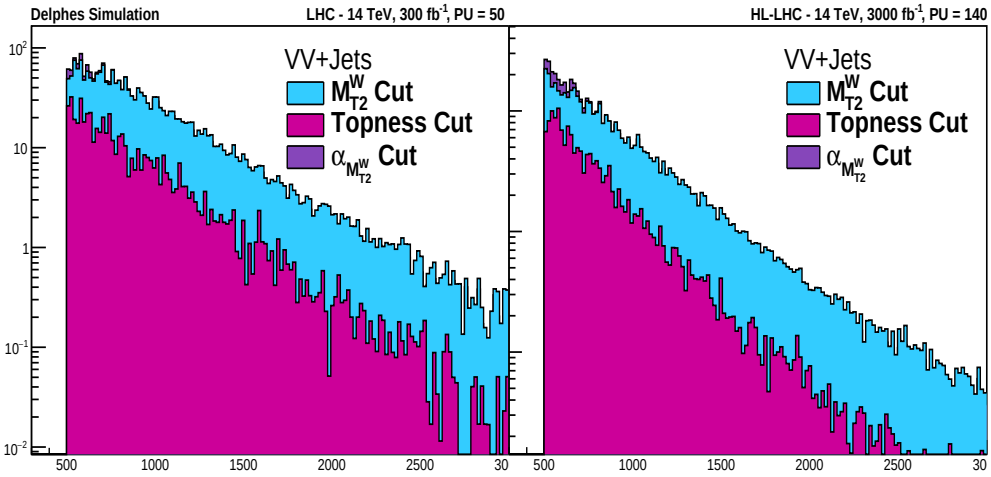


(b) M_T

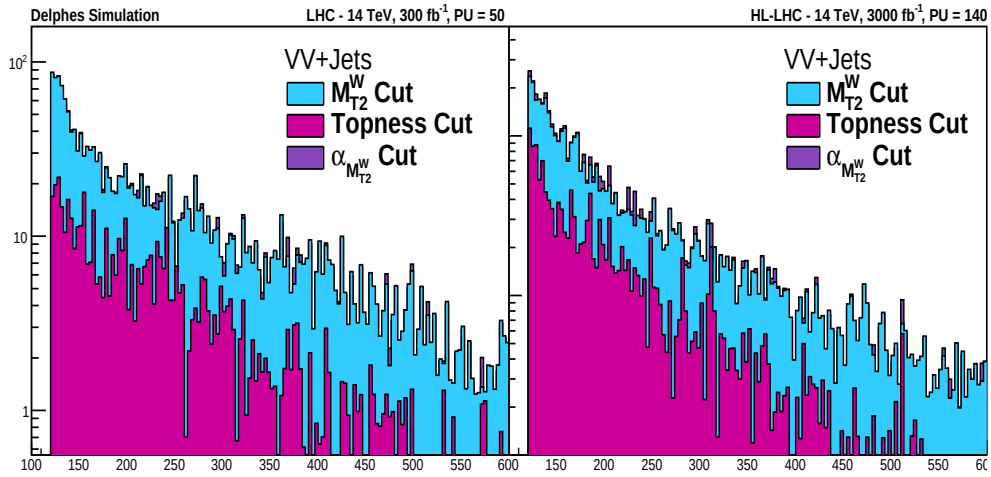


(c) E_T^{miss}

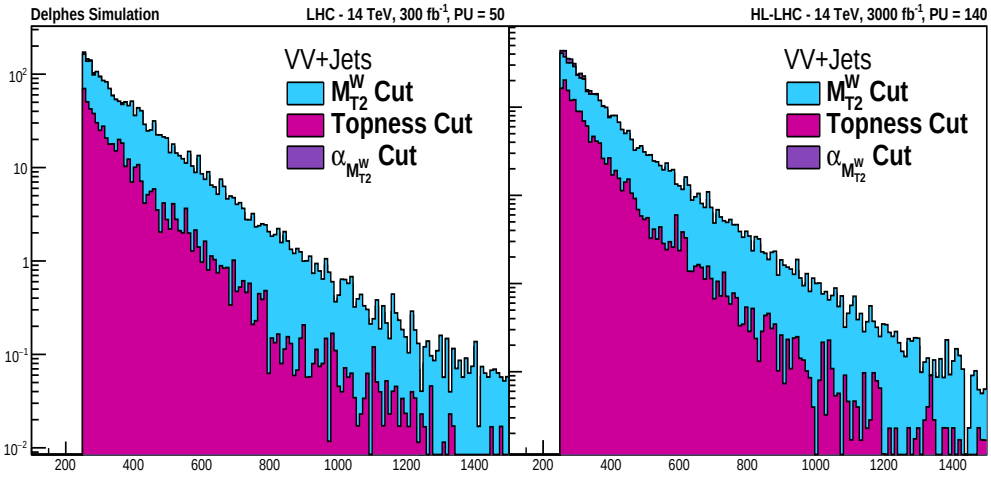
Figure 18 : The effect of the interchangeable variable selections on training variables, for signal $V + jets$ at 50PU, 300 fb^{-1} and 140PU, 3000 fb^{-1}.



(a) H_T,

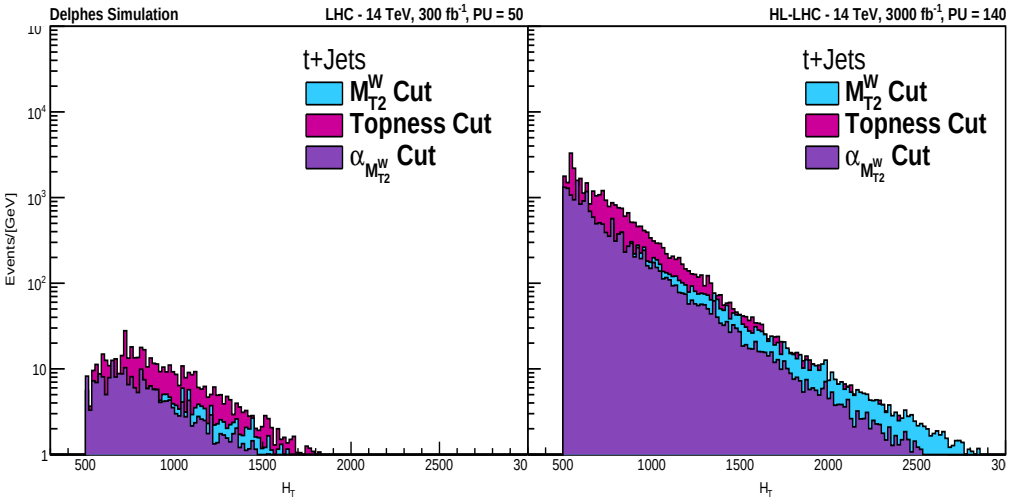


(b) M_T

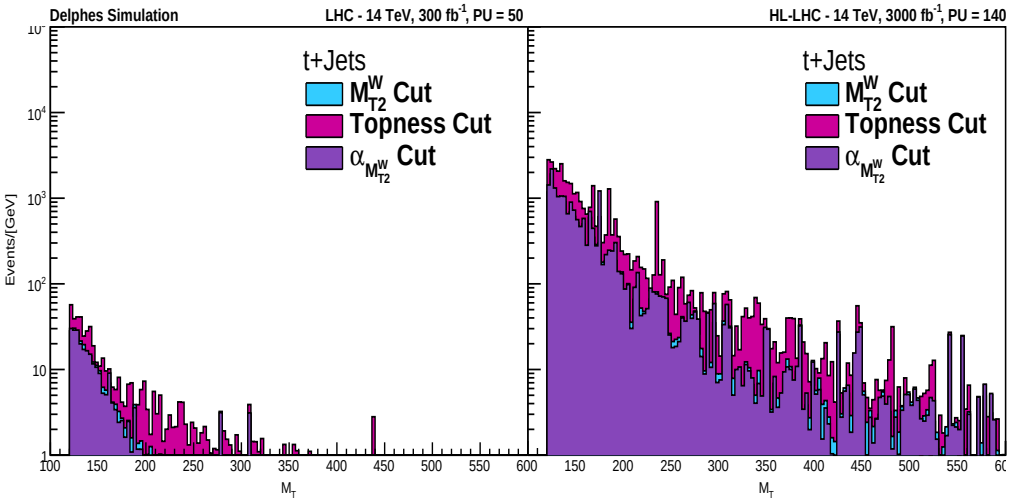


(c) E_T^{miss}

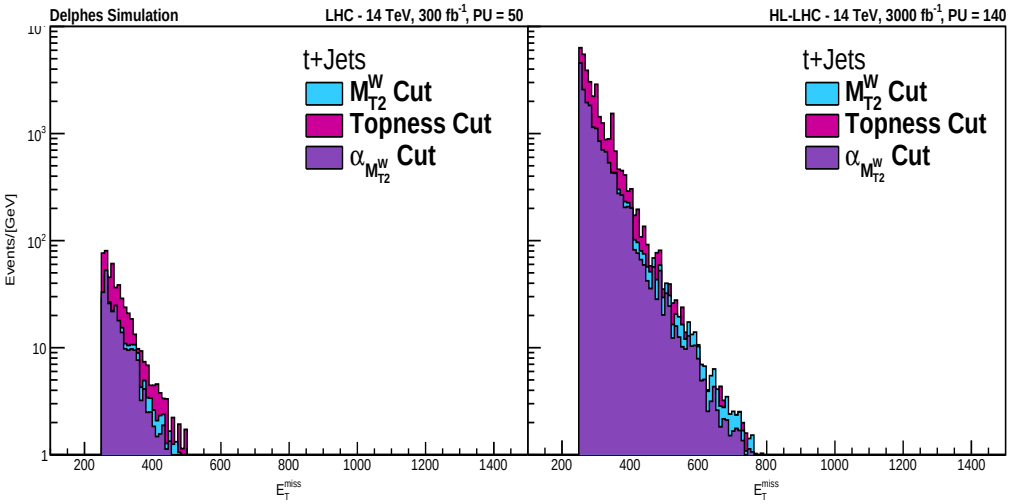
Figure 19 : The effect of the interchangeable variable selections on training variables, for signal $VV + jets$ at 50PU, 300 fb^{-1} and 140PU, 3000 fb^{-1}.



(a) H_T,

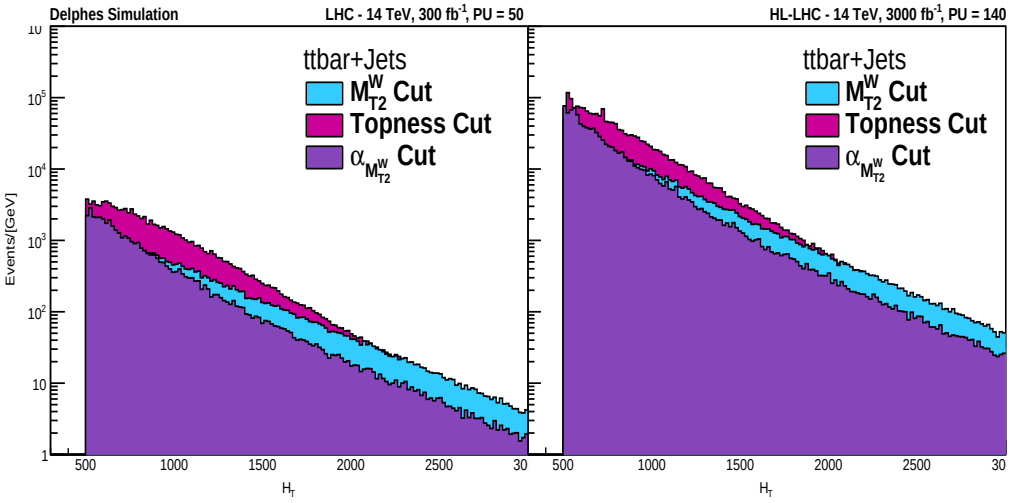


(b) M_T

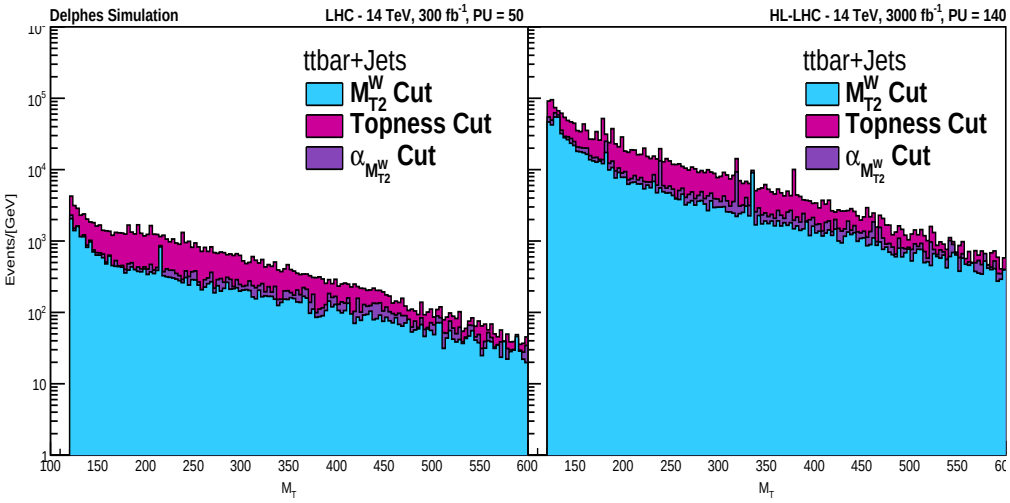


(c) E_T^{miss}

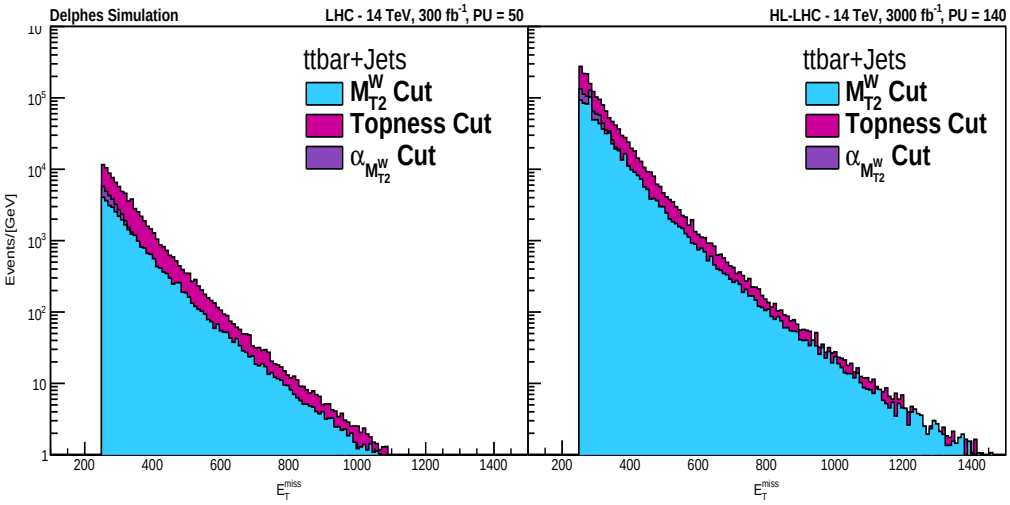
Figure 20 : The effect of the interchangeable variable selections on training variables, for signal $t + jets$ at 50PU, 300 fb^{-1} and 140PU, 3000 fb^{-1}.



(a) H_T,

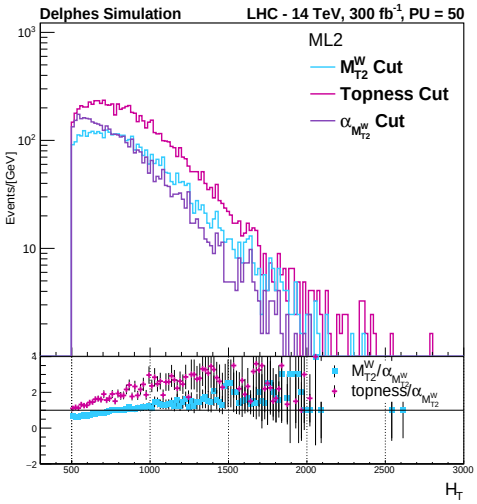


(b) M_T

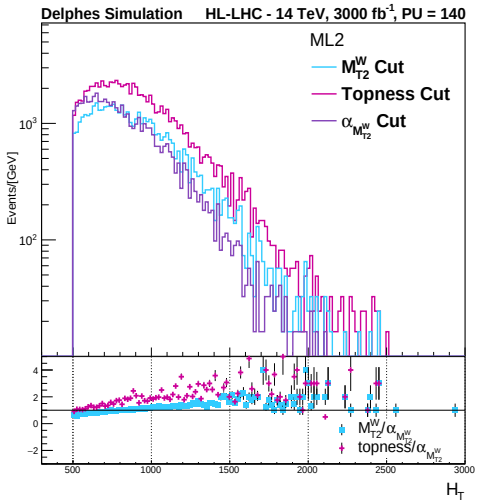


(c) E_T^{miss}

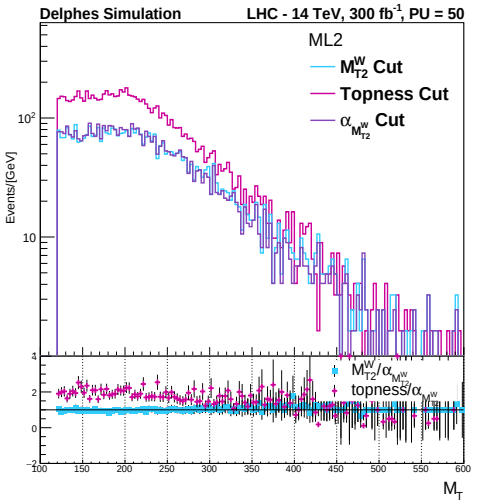
Figure 21 : The effect of the interchangeable variable selections on training variables, for signal $t\bar{t} + jets$ at 50PU, 300 fb^{-1} and 140PU, 3000 fb^{-1}.



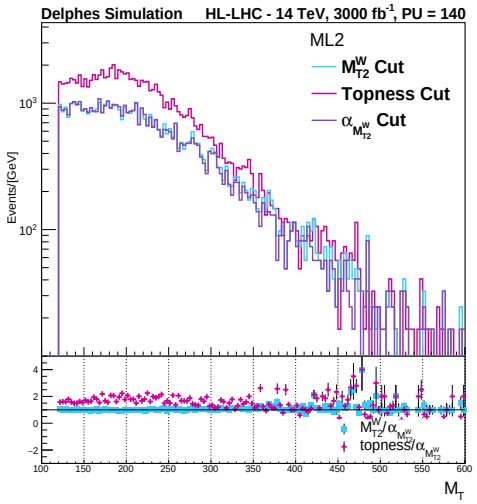
(a) H_T, LHC



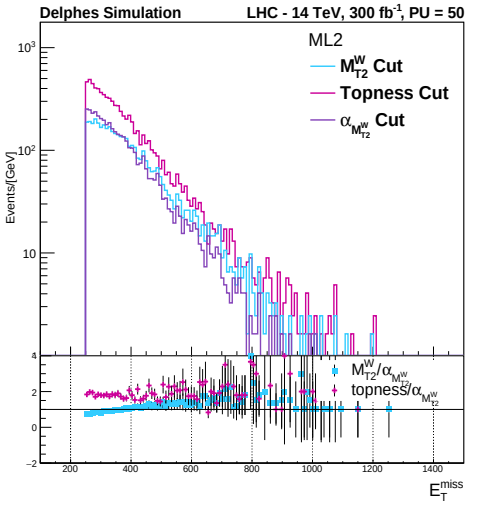
(b) H_T, HL-LHC



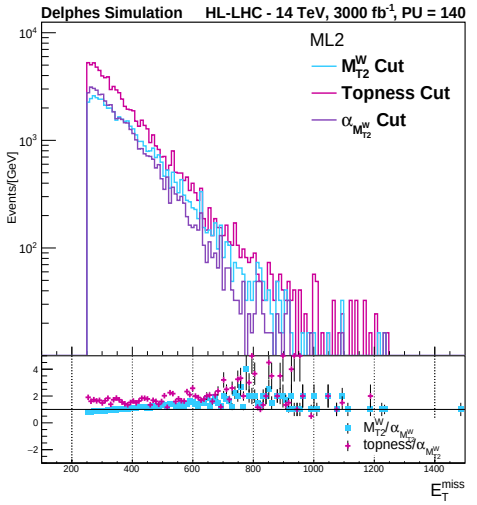
(c) M_T, LHC



(d) M_T, HL-LHC

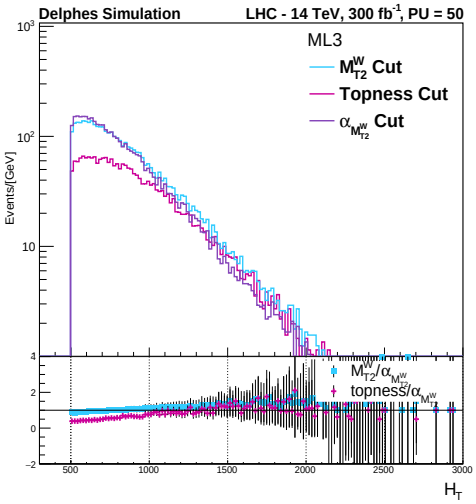


(e) E_T^{miss}, LHC

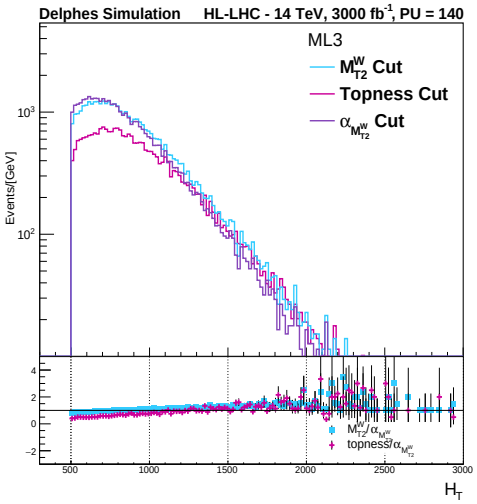


(f) E_T^{miss}, HL-LHC

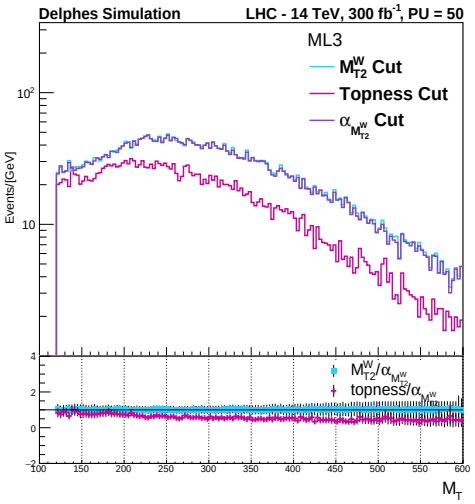
Figure 22 : The effect of the interchangeable variable selections on training variables, for signal ML2 at 50PU, 300 fb^{-1} and 140PU, 3000 fb^{-1}.



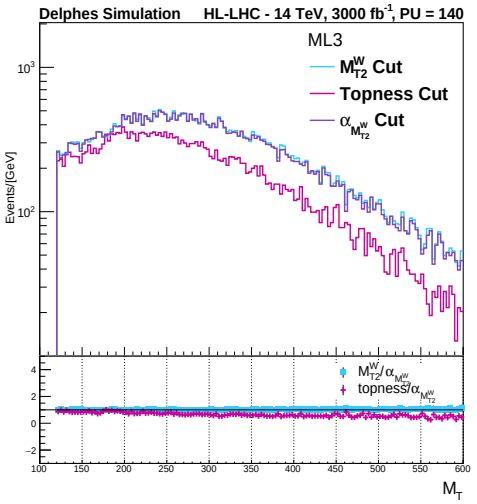
(a) H_T, LHC



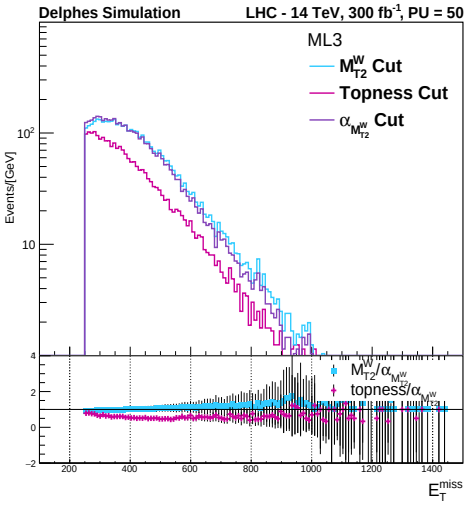
(b) H_T, HL-LHC



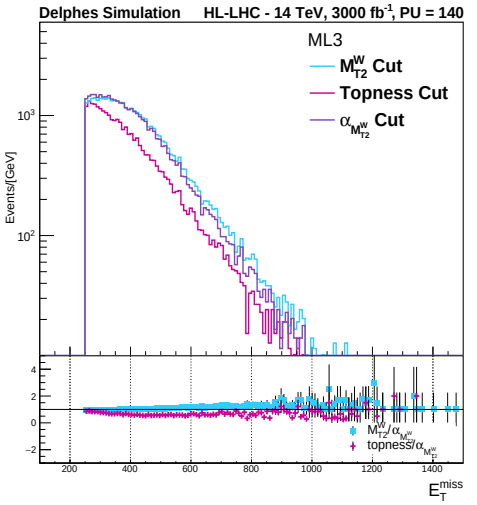
(c) M_T, LHC



(d) M_T, HL-LHC

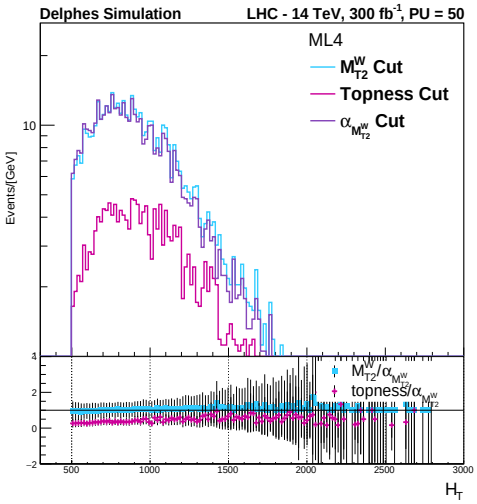


(e) E_T^{miss}, LHC

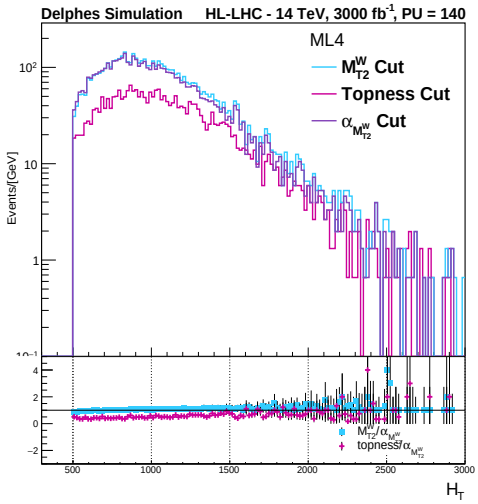


(f) E_T^{miss}, HL-LHC

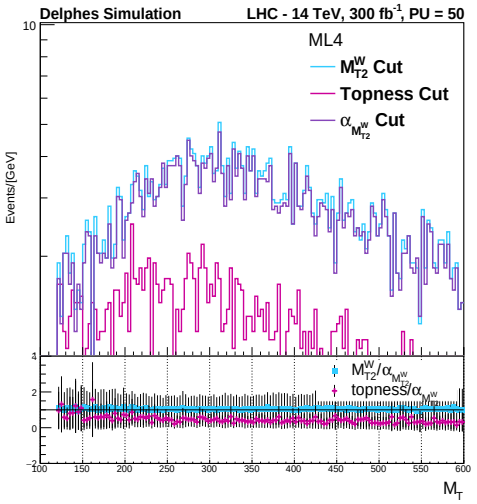
Figure 23 : The effect of the interchangeable variable selections on training variables, for signal ML3 at 50PU, 300 fb^{-1} and 140PU, 3000 fb^{-1}.



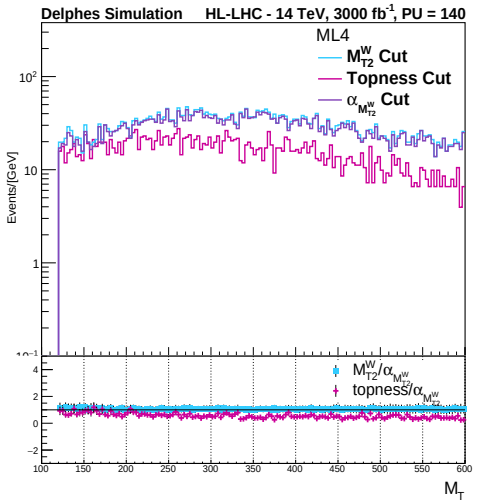
(a) H_T, LHC



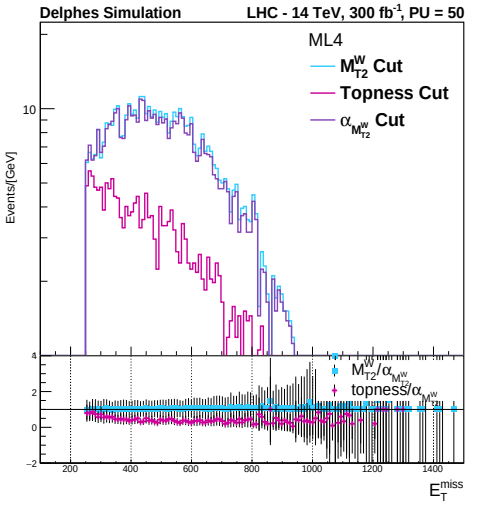
(b) H_T, HL-LHC



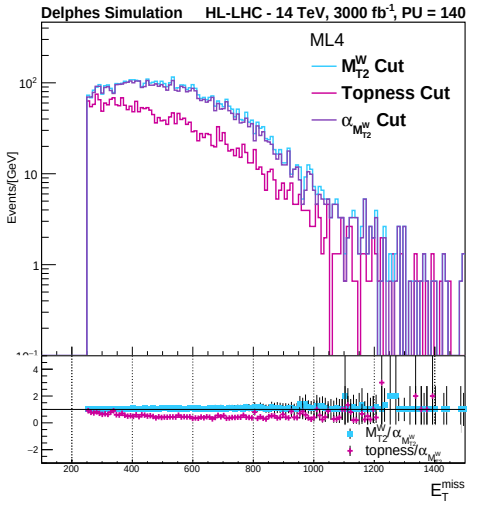
(c) M_T, LHC



(d) M_T, HL-LHC



(e) E_T^{miss}, LHC



(f) E_T^{miss}, HL-LHC

Figure 24 : The effect of the interchangeable variable selections on training variables, for signal ML4 at 50PU, 300 fb^{-1} and 140PU, 3000 fb^{-1}.

Table 4 : Summary of the number of signal and background events left after baseline and signal selections, for 50PU, 300 fb^{-1}.

Cutflow	**VV+jets**	**$t\bar{t}$+jets**	**V+jets**	**Top+jets**	**Total SM**	**ML1**	**ML2**	**ML3**	**ML4**	**ML5**
#Events	1108220000	2161240000	168426000000	620862000	172316322000	2430540	977610	114502	13171	11197
Lepton(e,μ)	105834000	453040000	5634380000	76718300	6269972300	554686	224433	28356	3121	2668
nJets ≥ 4	2982040	110592000	52761300	1903550	168238890	337819	141234	21038	2068	2038
bjets ≥ 2	1068210	99919900	15005600	1697330	117691040	277296	116786	18824	1832	1910
$E_T^{miss}>120$	38432	1409080	308651	23559	1779722	55089	29215	10259	944	1602
$M_T>120$	3401	177115	7222	906	188644	27546	15947	8030	790	1413
$H_T>500$	2628	132101	6183	821	141733	24758	14054	6642	744	1363

Table 5 : Summary of the number of signal and background events left after interchangeable kinematic selections, for 50PU, 300 fb^{-1}. These selections are applied one at a time after baseline and signal selections.

Cutflow	**VV+jets**	**$t\bar{t}$+jets**	**V+jets**	**Top+jets**	**Total SM**	**ML1**	**ML2**	**ML3**	**ML4**	**ML5**
$\alpha_{M_{T2}^W} > 0.20$	1891	40444	4708	266	47309	6542	4130	3710	474	903
$M_{T2}^W > 170$	1925	35048	4857	274	42104	6562	4125	3765	496	947
$Topness < 7$	526	90980	644	518	92668	14216	7621	2220	197	344

Table 6 : Summary of the number of signal and background events left after baseline and signal selections, for 50PU, 300 fb^{-1}.

Cutflow	**VV+jets**	**$t\bar{t}$+jets**	**V+jets**	**Top+jets**	**Total SM**	**ML1**	**ML2**	**ML3**	**ML4**	**ML5**
#Events	11082100000	21569500000	1684020000000	6206650000	1722878250000	24305400	9776100	1145020	131716	111971
Lepton(e,μ)	1042020000	4600470000	55007300000	771728000	61421518000	5680390	2301950	290616	31899	27442
nJets ≥ 4	51045400	1191630000	1441450000	45384500	2729509900	4248170	1753300	248609	25672	24364
bjets ≥ 2	15617400	1036150000	336086000	36346400	1424199800	3032490	1271970	209584	21128	22281
$E_T^{miss}>120$	623243	23709200	6115080	597453	31044976	629957	327866	115831	11108	18811
$M_T>120$	101264	5044740	760005	108045	6014054	316882	176312	88901	9154	16372
$H_T>500$	66216	2748100	389177	54177	3257670	290329	159587	79440	8794	16028

Table 7 : Summary of the number of signal and background events left after interchangeable kinematic selections, for 140PU, 3000 fb^{-1}. These selections are applied one at a time after baseline and signal selections.

Cutflow	**VV+jets**	**$t\bar{t}$+jets**	**V+jets**	**Top+jets**	**Total SM**	**ML1**	**ML2**	**ML3**	**ML4**	**ML5**
$\alpha_{M_{T2}^W} > 0.20$	45140	952408	264792	18684	1281024	81585	47968	38999	4999	9716
$M_{T2}^W > 170$	43490	794577	248822	15007	1101896	82111	48823	39775	5311	10350
$Topness < 7$	15171	1752680	92339	33797	1893987	149418	80555	26561	2580	4330

Appendix A.8

Pileup Simulation with Delphes 3.0.10

Pileup cards are obtained from

```
http://cmssw.cvs.cern.ch/cgi-bin/cmssw.cgi/UserCode/
spadhi/Snowmass/Cards/.
```

Three cards were used for NoPU, 50PU and 140PU scenarios.

```
delphes_card_Snowmass_NoPileUp.tcl
delphes_card_Snowmass_50PileUp.tcl
delphes_card_Snowmass_140PileUp.tcl
```

Pileup is read from the binary file MinBias.pileup that is saved in Delphes-3.0.10 environment. For 50PU, pileup and vertex is set with:

```
# pre-generated minbias input file
set PileUpFile MinBias.pileup
# average expected pile up
set MeanPileUp 50
# spread in the beam direction in m (assumes gaussian)
set ZVertexSpread 0.05
```

Resolution is set in meters with

```
 # assume perfect pile-up subtraction for tracks with |z| > fZVerte
# Z vertex resolution in m
set ZVertexResolution 0.0001
```

For 140PU, pileup and vertex is set with:

```
# pre-generated minbias input file
set PileUpFile MinBias.pileup
# average expected pile up
set MeanPileUp 140
# spread in the beam direction in m (assumes gaussian)
set ZVertexSpread 0.05
```

Resolution is set in meters with

```
 # assume perfect pile-up subtraction for tracks with |z| > fZVerte
# Z vertex resolution in m
set ZVertexResolution 0.0001
```

CURRICULUM VITAE

Name Surname: Ceren Guzelgun

Place and Date of Birth: KONAK, 24.08.1992

E-Mail: ceren.guzelgun@cern.ch

EDUCATION:

- **B.Sc.:** 2016, Istanbul Technical University, Science and Letters, Engineering Physics
- **M.Sc.:** 2018, Istanbul Technical University, Institute of Science and Technology, Engineering Physics

PROFESSIONAL EXPERIENCE AND REWARDS:

- 2018 Visiting Researcher at Deutsches Elektronen-Synchrotron
- 2017 Summer Student at Deutsches Elektronen-Synchrotron

OTHER PUBLICATIONS, PRESENTATIONS AND PATENTS:

- C. Guzelgun, (2017, August) Search for Long Lived Charginos, voted among the students and supervisors as one of the 3 talks that would represent CMS at the end of summer talks in DESY, Hamburg, GERMANY.
- C. Guzelgun, (2016, August 19-22). Visualising Geodesics and Quantum Field Thoery Solutions for Black Holes Poster presented at the biennial meeting of the Hellenic Society on Relativity, Gravitation and Cosmology, ”Recent Developments in Gravity”, Mykonos, GREECE.