Palacký University Olomouc
Faculty of Science
Department of Experimental physics

Doctoral Thesis

Study of spectra of top quarks produced in pairs in experiment ATLAS

Author	**Josef Pácalt**
Study programme:	Physics
Specialization:	Applied Physics
Study form:	Daily
Supervisor:	Mgr. Jiří Kvita, PhD
Date of submission:	January 2023

Prohlášení

Prohlašuji, že jsem předloženou disertační práci vypracoval samostatně pod vedením Mgr. Jiřího Kvity, Ph.D a že jsem použil zdrojů, které cituji a uvádím v seznamu použitých pramenů.

V Olomouci dne 23. 1. 2023

.....

Josef Pácalt

Bibliografická identifikace

Jméno a příjmení autora	Josef Pácalt
Název práce	Zkoumání spekter top kvarků produkovaných v párech na detektoru ATLAS
Typ práce	Disertační
Pracoviště	Katedra experimentální fyziky
Vedoucí práce	Mgr. Jiří Kvita, PhD
Rok obhajoby práce	2023
Abstrakt	Disertační práce je zaměřena na podrobné studium přínosu aplikace semi-boostovaných topologií při studiu procesu $pp \to Z'(y_0) \to t\bar{t}$, $t_1 \to W + b \to \nu + \ell + b$, $t_2 \to W + b \to b + q + \bar{q}$ za pomocí hypotetických částic Z' a y_0. Nárůst počtu události při začlenění těchto topologií do analýz je odhadnut mezi 20%–50%. V práci je zahrnuta studie síly signálního spektra hypotetické částice a vliv metody unfoldingu na sílu signálu. V obecné rovině lze tvrdit, že metoda unfoldingu oslabuje signál přibližně o 20%-30% s výjimkou úhlových proměnných, kde tento pokles není tolik patrný.
Klíčová slova	Top kvark, vektorová rezonance, skalární rezonance, differenciální účinný průřez, unfolding, signifikance signálu, semi-boostovaný režim
Počet stran	117
Počet příloh	2
Jazyk	Anglický

Bibliographical identification

Autor's first name and surname	Josef Pácalt
Title	Study of spectra of top quarks produced in pairs in experiment ATLAS
Type of thesis	Doctoral Thesis
Department	Department of Experimental Physics
Supervisor	Mgr. Jiří Kvita, PhD
The year of presentation	2023
Abstract	The dissertation work is focused on the detailed study of advantages of the application of semi-boosted topologies in studying the process $pp \to Z'(y_0) \to t\bar{t}$, $t_1 \to W + b \to \nu + \ell + b$, $t_2 \to W + b \to b + q + \bar{q}$ with usage of hypothetical particles Z' and y_0. The increase in the number of events by implementing of those topologies in analyses is evaluated to be between 20%–50%. The study of the signal strength of the hypothetical particle signal spectrum and influence of the unfolding method on the strength is also included in the work. The unfolding method weakens the signal strength by 20%-30%, with exception of the angular variables, where the decrease is not evident in such measure.
Keywords	Top quark, vector resonance, scalar resonance, differential cross-sections, unfolding, signal significance, semi-boosted regime
Number of pages	117
Number of appendices	2
Language	English

Poděkování

Rád bych touto cestou poděkoval všem, kteří mi pomohli dokončit tuto práci i doktorské studium. Velký dík patří všem kolegům z CERNu, kteří mne vedli a podporovali ať už v rámci kvalifikačního úkolu tak i v analýze top kvarků. Dále bych rád poděkoval svému školiteli Mgr. Jiřímu Kvitovi, Ph.D za jeho vedení, čas a důvěru ve zdárný konec studia.

Neméně velký dík patří i mým rodičům, kteří stáli za mnou a podporovali mne v časech těžších, které studium provázely.

Zvláštní poděkování si zaslouží Antonín "Štístko"Mládek za svůj neutuchající optimismus a zápal pro věc.

A děkuji všem mým přátelům a kamarádům, bez kterých by to občas nešlo.

Děkuji Vám všem.

Contents

Introduction

The advancements of knowledge in the experimental particle physics can be divided into two main streams, the astroparticle physics and the accelerator based physics. As the name of this work hints, it belongs to the second type of particle physics. Colliders helps physicists to reveal the yet to be know details about our universe, its history and its performance on sub-nuclear scales. Up to this time, the Standard Model (SM) theory proved to be precise enough with predictions of the existence of particles and their properties, although there are discrepancies in several cases leading to the theories Beyond the Standard Model (BSM). Several of the BSM theories deals with heavier versions of known bosons which could unravel possibilities of their existence in already measured spectra, hidden between uncertainties, or use them as a probe in simulated scenarios with a higher collision energy and beam luminosity.

This dissertation work aims to explore possibilities of adding almost unused topologies (semi-boosted) into reconstruction procedures for studies of top anti-top quark pair ($t\bar{t}$) spectra in the semi-leptonic decay channel (ℓ+jets). $t\bar{t}$ decay channels are named after products of decaying W bosons, which may decay to a lepton and a neutrino (leptonic decay), or to a of quark and anti-quark pair (hadronic decay), in other words in the semi-leptonic $t\bar{t}$ channel one W boson decays leptonically and the other one hadronically. Two topologies are commonly used at presence, resolved and boosted, but in between them lies a region, where the resolved topology transits to the boosted topology, while none of those has the number of events high enough. This region of energies is often omitted from analyses in both topologies.

To tune the process to the specific energy the aforementioned additional BSM heavy boson is used of a specific mass. Scanning over the area of interest is done by simulating the same process with different values of the heavy boson mass and full analysis for each topology is performed. The results are probed for the performance on the unfolding procedure and the loss of significance by this procedure is studied.

The dissertation work is divided into 6 chapters, where first chapter is dedicated to description of the CERN and its experiments, mainly the ATLAS experiment. Second chapter is a short introduction to the Standard Model, classification of events and motivation for analyses in the top quark field in particle physics. The third chapter presents energy corrections for the measured particle showers in the detector and measured objects definitions. The fourth chapter describes a private analysis sample generation, energy correction factors derivation, selection of events and objects and reconstruction of the top anti-top quark pairs. The fifth chapter is dedicated to the spectra shape comparison of chosen variables of reconstructed particles, spectra comparison in different topologies, resolution of reconstructed $t\bar{t}$ mass peak and migration of events between reconstruction levels in terms of topologies and comparison over different generated samples. The sixth and last chapter describes the performance of the simulated signal in the unfolding procedure for several variables and strength of the signal in the corrected spectrum.

Chapter 1

CERN, LHC and the ATLAS Experiment

The exploration and knowledge of particle physics is leaping forward with the advancement in digitization and also with industrial progress. The main problem of particle physics is the energy. Processes with heavier particles, to which top quark belongs, are visible only in collision where there is the energy high enough to create them. The methodology of studying particles is similar if not the same with first experiments ever studied in this field, the particles are sent to the fixed target and physicist then study products of the collision. The more advanced setup is performed by colliders, which enables the possibility to send particles against each other fo the collision. The increasing need for higher energy to uncover deeper knowledge lead to building larger colliders. The largest collider facility in the world has been build in the middle of Europe and there is no other laboratory which could compete in the energy value of the colliding particles in the time of writing this work.

The laboratory facility, history, experiments and notable breakthroughs in particle physics are described in this chapter. The first section serves as introduction to history and aims of the international laboratory. Large Hadron Collider(LHC) is described in the second section together with variables tied to physical layout of the collider for example framework coordinates. Third section covers overview of experiments currently operating at the LHC. The fourth section presents thorough description of the ATLAS experiment and definitions of variables tied to the construction layout of the detector.

1.1 CERN

Conseil Européen pour la Recherche Nucléaire (CERN) or European Council for Nuclear Research in English is a research organization with the first council meeting on 5 May 1952 [1]. This organization aimed to facilitate the nuclear research in Europe without any link to the military purposes. The acronym CERN was firstly used to name the organization committee, but then it remained unchanged for the description of the whole laboratory till this day.

The work on the site started on 17 May 1954 and the first accelerator, Synchrocyclotron (SC), was in full operational state on 16 August 1957 [2]. This accelerator contributed by the discovery of the rare decays of pions to electrons [3] one year later and its the energy limit was 600 MeV. The Proton-Synchrotron (PS) accelerator started to work on 24 November 1959 with maximum of energy for collisions of

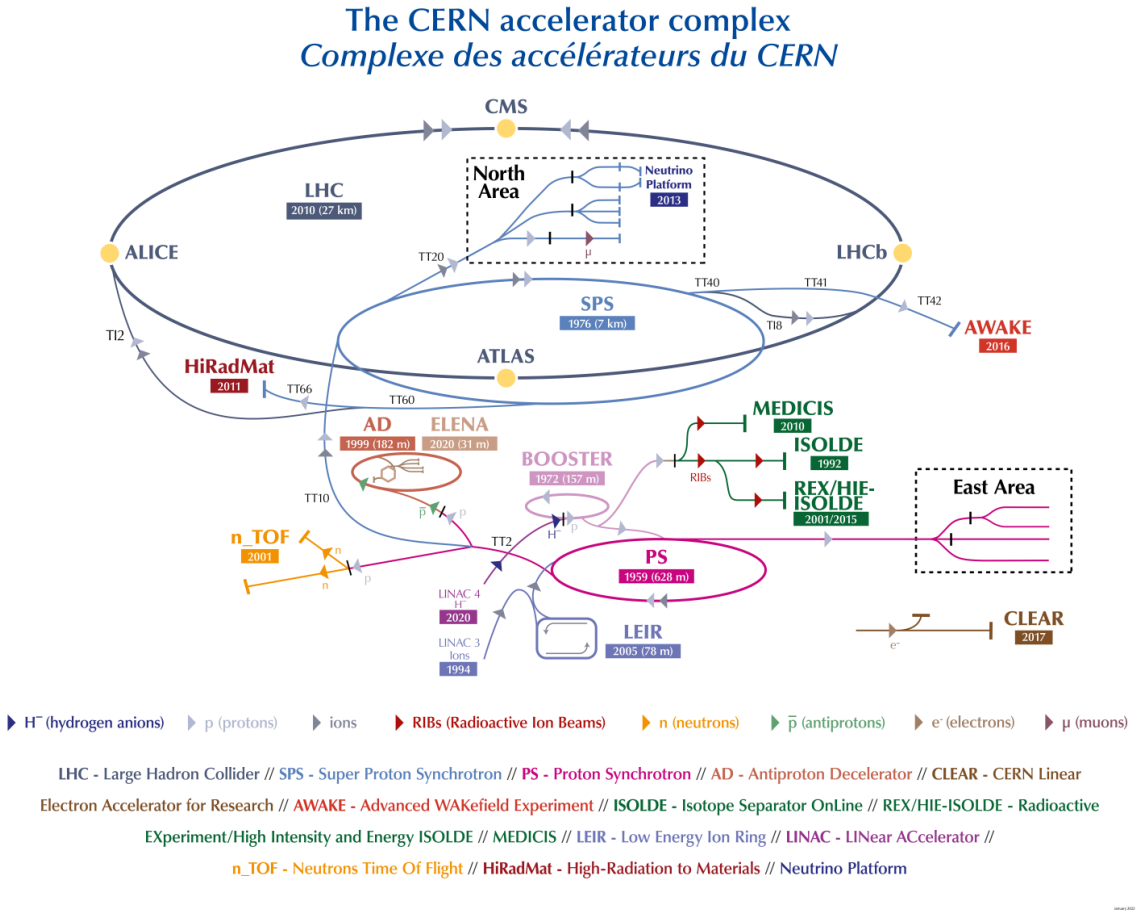


Figure 1.1: Overview of the accelerator complex at CERN. The figure was taken from[5].

28 GeV. The Proton-Synchrotron accelerator is still used as a pre-accelerator nowadays. The next step in achieving higher energies in accelerators at CERN site was Super Proton-Synchrotron (SPS), which started to operate on 3 May 1976 and is able to reach the energy up to 450 GeV. SPS is also used as pre-accelerator for the LHC. The SPS accelerator was the crucial instrument for the discovery of Z and W bosons in 1983 as a proton-antiproton collider, for which scientists Carlo Rubbia and Simon van der Meer shared the Nobel prize [4]. The Large Electron-Positron collider was the next generation in the history of accelerators at the CERN. The circumference of this accelerator is 27 kilometers and its energy limit started at 100 GeV but in the last years of its measurements was able to provide energy for collisions up to 208 GeV. The very important measurement at LEP confirmed that there are exactly tree families of light neutrinos. The LEP accelerator was shut down in November 2000, but the already present tunnel was used to build the LHC. The overview of accelerator complex is shown in Figure 1.1.

1.2 Large Hadron Collider

The Large Hadron Collider resides about 175 meters under the borderline between Switzerland and France. The LHC is the largest particle collider to this date with

circumference over 27 kilometers and four main intersection points, where the experiments resides. The accelerator begun its operation in the year 2005 at energy 500 GeV, which increased in 2010 to the energy of 3.5 TeV and after the upgrades the energy increased to the 6.5 TeV and after another upgrade the energy has increased to the 6.8 TeV [6, 7].

The largest achievement of the LHC is the discovery of the Higgs boson, the last particle of the Standard Model theory, in July 2012. Among its other successes belongs the observation of the quark-gluon plasma, the densest known form of matter, and observation of the bounded pentaquark states. The LHC measurements and experiments are important for probing theories beyond the Standard Model, for example the supersymetry.

1.3 Experiments at the LHC

There are four main experiments at the LHC, ATLAS (A Toroidal LHC Apparatus), CMS (Compact Muon Solenoid), ALICE (A Large Ion Collider Experiment) and LHCb (Large Hadron Collider beauty). The largest multipurpose experiments are ATLAS and CMS, both capable of measuring similar particle properties but with different sub-detectors, which work in tandem to cross-check the measured values of the other detector device. The ALICE experiment's main aim is to measure properties of collisions of the heavy ions. The experiment LHCb specializes in measurements of the charge and parity violation phenomena.

CERN supports research outside the LHC ring, for example LINAC (Linear Accelerator) for studying heavy ion collisions mainly lead ions, but also the oxygen, xenon and argon ions were studied.

1.4 The ATLAS Experiment

The ATLAS experiment is the largest one built at the LHC ring. The construction was finished in 2008 and first measurement was performed in September 2008. Over its 13-year operation history the detector was upgraded over several technical stops. The detector recorded in total the integrated luminosity of $140 fb^{-1}$, which helped to increase precision of unsubstantial amount of particle properties, ranging from top quark mass to exclusion limits for the heavier hypothetical particles in BSM theories.

The collaboration shields over 3000 physicist from 38 countries involving 183 institutions under one collaboration.

The detailed description of the ATLAS detector, detecting principles for different particles and other technicalities are described in the next section.

1.4.1 The ATLAS Detector

The ATLAS detector situated at the LHC at CERN is a multipurpose cylindrical detector for measurement of properties of particles, studies of high energy physics as well as signs of new physics. There are two beam pipes at the LHC in which two beams of protons are circulating in opposite directions. These beams are crossing in four interaction points where the detectors are build. Protons circulating in LHC are colliding in bunches for enhancing the chance of collision, each bunch containing approximately 10^{11} protons and colliding with frequency up to 40 MHz. There are

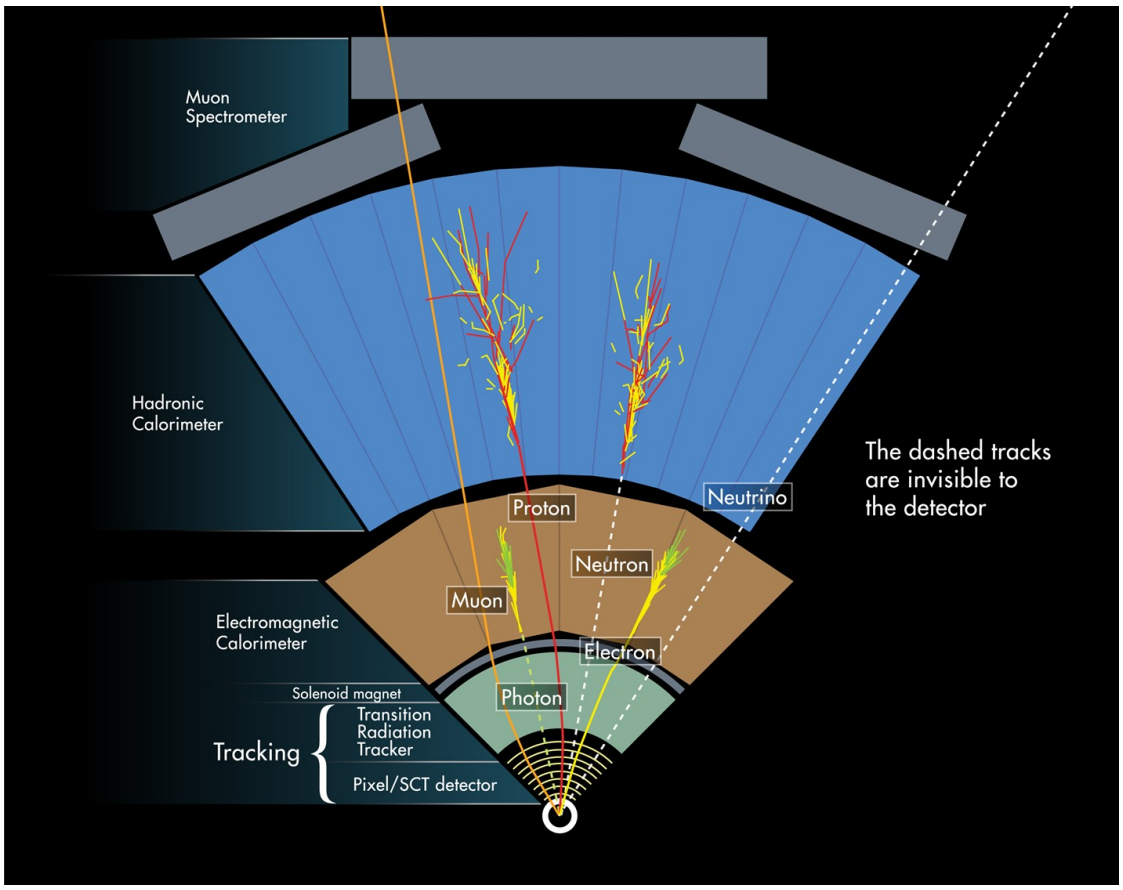


Figure 1.2: Different particle traces for the different sub-detectors of the ATLAS experiment. This figure was taken from [8].

multiple collisions of protons happening within each bunch crossing, a process called pile-up (PU), but usually only few of them can be classified as a "hard" process. The "hard" process is the creation of high-transverse momentum objects in proton-proton collision. These can be heavy quarks, leptons or bosons which are objects of interest for scientists and there is need of precise measurement of their properties or precise measurement of properties of their decay products.

The ATLAS uses right-hand coordinate system with x-axis pointing to the center of the ring, y-axis pointing up and z-axis pointing alongside the ring. This system of coordinates is also used for the description of events inside experiments. The experiments often use A-side and C-side notation, which is shorthand of clockwise and anticlockwise direction from the center of the experiment detector respectively.

For further details of the detection see Fig. 1.2 which illustrates the passage of individual particles through the sub-detectors of the ATLAS detector.

The details about the ATLAS detector are presented in this chapter. The actively measuring devices are described in the order of distance from the interaction point, thus the first section covers tracking detectors. The second section describes calorimeters. Finally the third section introduces muon chambers, the farthest layer of the detector from the interaction point.

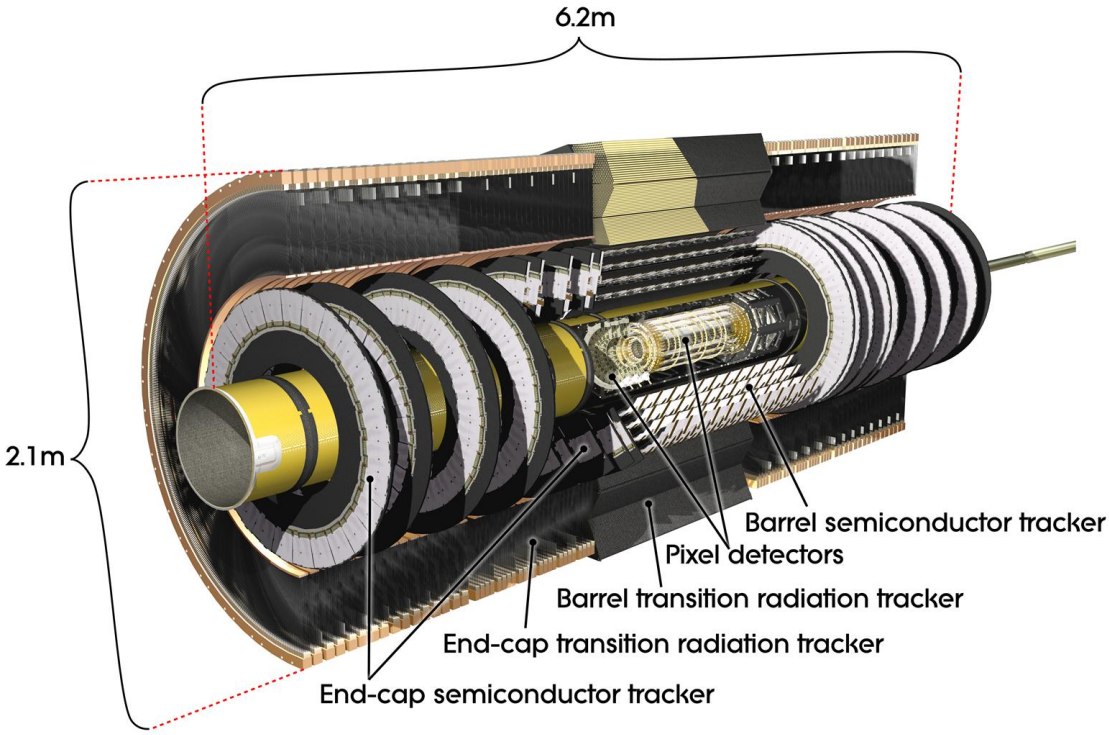


Figure 1.3: Overview of the inner detector of the ATLAS experiment. The figure was taken from [9].

1.4.2 Tracking Detectors

Tracking detectors are used for the reconstruction of charged particles trajectories in its volume and the particle track curvature, which is important for the identification of charge of the particle and for the vertex reconstruction algorithm, which is responsible for the calculation of the position of the interaction point. The information from the tracking detector is combined with the information from the calorimeter to identify the particle, for example to distinguish between a proton and a neutron. The ATLAS inner or tracking detectors are, in order from the interaction point, the pixel detector, the semiconductor tracker and transition radiation tracker (TRT). The overview of the inner detector is in Fig. 1.3.

First sub-detector from the interaction point is the pixel detector, which detects charged particles by using CMOS modul in each pixel. The principle is based on the generation of the charge by the ionizing particle in the semiconductor material, where the generated charge is drifting to the electrodes in the electric field and collected as a signal. Each of the modules has own read-out electronics for further processing and its small size allows fast charge collection. The pixel detector contains 92 000 000 pixels or channels in 1 736 inner tracker modules and in 288 end-cap modules. The inner pixel detector stands of three layers cylindricaly enclosing the interaction point with the resolution of the detector equal to 50×400 μm^2 and one innermost layer with resolution of $50\times$ 250 μm^2 and precision 10 μm [10]. The pixel detector is shown in Fig. 1.4

The detecting principle for the semiconductor tracker is based on the phenomenon, where the incoming particle creates the charge carriers in the material, electrons and

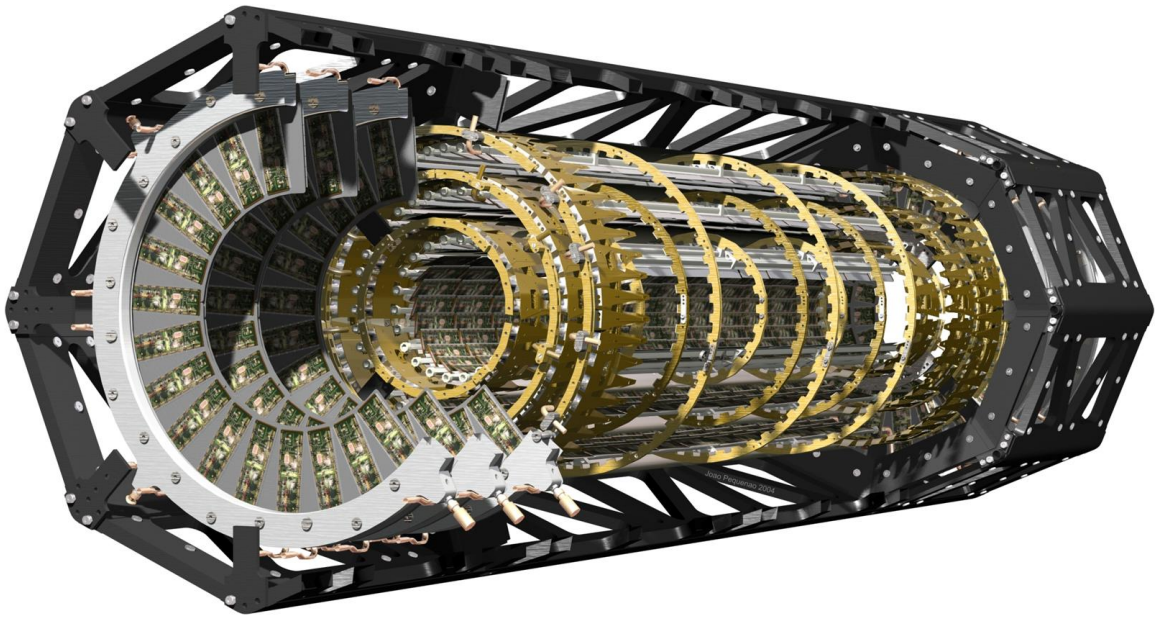


Figure 1.4: The pixel detector of the ATLAS experiment. The figure was taken from [11].

holes in our case, whose then traverse in the material to the corresponding voltage similar to the previous sub-detector. This is considered as a hit in the read-out channel for the reconstruction. The semiconductor tracker consists from four cylidrical layers in the central area of the detector and from nine end-cap disks on each side of the detector with over 6 000 000 channels in 6 398 barrel modules and in 9 505 end-cap modules. The charged particles traces are measured with precision up to 25 μm for the path in the transverse orientation to the beam pipe [12].

The transition radiation tracker is the outermost part of the inner detector and the detection principle based on the transition radiation phenomenon, where electromagnetic radiation is emitted by the charged particle crossing the boundary between two media. The main advantage of the detector is the ability to distinguish electrons from pions. It consists of 52 544 tubes each with length 144 cm in the barrel and 245 760 tubes in end-cap areas 37 cm long. All tubes have 4 mm in diameter and 30 μm thick gold-tungsten wire in the middle. The precision of the detector is 0.13 mm corresponding to the distance of the particle track to the wire [13].

1.4.3 Calorimeters

Calorimeters are detectors of energy of ionizing particles, usually placed after tracker detectors in the order from the interaction point. ATLAS sampling calorimeters are build from absorber, which is usually a kind of dense material suitable for the development of the particle shower interleaved with layers of active material which is used for measurement of the particle shower energy. The principle of measurement is to stop the incoming particle by collisions with the absorber material, which creates a particle shower and measure deposited energy of the shower in active medium layers. The downside of the measurement is that the original particle is destroyed in the process. Calorimeters are divided into two groups, electromagnetic and hadronic calorimeters,

where the electromagnetic calorimeter is device for measuring the energy of electrons, positrons and photons due to the energy loses in the calorimeter through the conversion of photons to electron-positron pairs and bremsstrahlung of electrons and positrons, which is known as the electromagnetic cascade, and hadronic calorimeters for measurements of energy of hadrons, mostly protons, neutrons and pions. Calorimeters have potential to detect almost all types of particles with exception for non-interacting neutrinos and muons, which ionization losses are usually too small to stop them.

The electromagnetic calorimeter of the ATLAS detector is a sampling calorimeter based on liquid argon active medium with a lead absorber. The lead absorber is folded in the accordion structure and liquid argon placed between the lead layers serves as the active medium cooled to the temperature of $-184°$C. The calorimeter encloses the tracking detectors in the barrel area with cylindrical layers parallel to the beam and two end-cap parts placed transversely to the beam. The barrel layers are 6.4 m long, 53 cm thick and contain 110 208 channels, whereas the discs in the end-cap area have 2,077 m in radius and are 0.632 m thick, with 63744 channels [14, 15].

The barrel part of the hadronic calorimeter lies radially further from the interaction point than the electromagnetic one and it is a sampling calorimeter built on the principle of alternating iron and scintillation tiles placed longitudinal to the beam. The calorimeter is divided in wedges, each covering the azimuthal angle of approximately 0.1 radian and 158,5 cm high, with summed number of channels equal to 9 836 [16]. The second part of the hadronic calorimeter are two wheels of end-cap detectors, with copper absorber and liquid argon as the active medium with thickness 0.8 m for the first wheel and 1.0 m for the second one both with radius 2.09 m, with summed number of read-out channels of 3648. The third part of the hadronic calorimeter also called the forward calorimeter is formed from tree discs with 0.455 m in diameter and with thickness of 0.45 m made of copper (first disc) and tungsten (second and third disc) absorber with inserted rods perpendicular to the beam with liquid argon active medium and number of read-out channels equal to 11 288 for both end-caps [15].

The layout of the ATLAS calorimeters is shown in Fig. 1.5.

1.4.4 Muon Chambers

Muon spectrometer is spatially the largest part of the ATLAS detector, its contribution cover roughly one half of whole detector. Muon spectrometer of the ATLAS detector consists of four different types of muon detecting systems, namely two muon trigger systems, thin gap chambers and resistive plate chambers and two precision measurement sub-detectors: monitored drift tubes and cathode strip chambers. Muon trigger systems are fast responding detecting devices used for triggering on events of interest with limited precision, whereas the precision measurement sub-detectors are measuring with higher precision but only events selected by triggers. The principle of muon spectrometer is similar for all four measuring sub-detectors, where ionizing particle interacts with gaseous mixture inside the detecting chamber creating anions and cations, which travel to the electrodes and read-out as the signal. The resolution of muon chambers is in average comparable with the size of the used chambers [18].

The resistive plate chambers are muon trigger sub-detector placed in the barrel area whereas the thin gap chambers are used in the end-cap area of the detector. The resistive plate chambers are made from two parallel resistive plates serving as electrodes with 2 mm gap filled with gaseous mixture of 94.7% $C_2H_2F_4$, 5% C_4H_{10} and 0.3% SF_6. The thin gap chambers are multiwire proportional ionizing detection chambers with

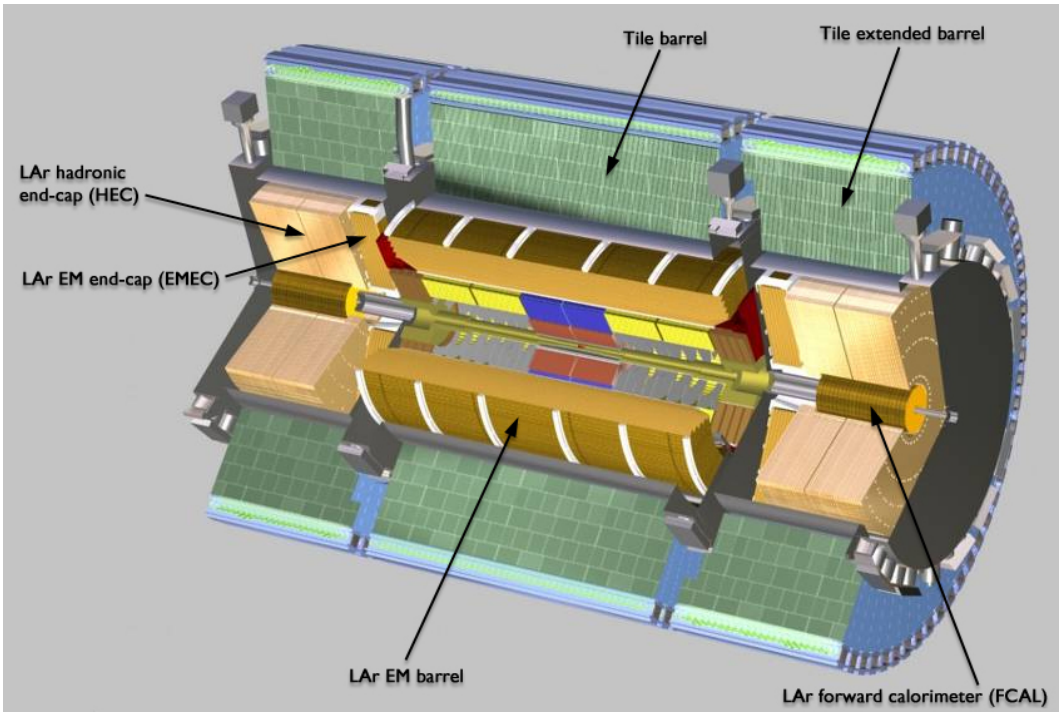


Figure 1.5: The layout of the electromagnetic and hadronic calorimeter of the ATLAS experiment. The figure was taken from [17].

distance between wires of 1.4 mm which serves as anode and distance of wire to cathode 1.8 mm and wire thickness of 50 μm. The used gas mixture is combination of CO_2 55% and 45% n-pentane [18].

The monitored drift tubes are placed in both, perpendicularly in barrel area and transversely in end-caps, occupying most of the space of the whole spectrometer. The aluminum tubes with 3 mm in diameter and from 0.9 to 6.2 m in length are placed in three or four layers filled with active gas mixture of 93% Ar and 7% CO_2. The collecting electrode is 50 μm W-Rn wire placed in the middle of the tube with average resolution of about 80 μm per wire. Finally, the cathode strip chambers are multiwire proportional ionizing detection chambers with cathode strip read-out. The distance between the wires is 2.54 mm, with gas mixture of 30% Ar, 50% CO_2 and 20% CF_4 with resolution of $50 - 300\mu$m in dependence on the counting rate [18].

The overview of the muon chamber of the ATLAS detector is shown in Fig. 1.6.

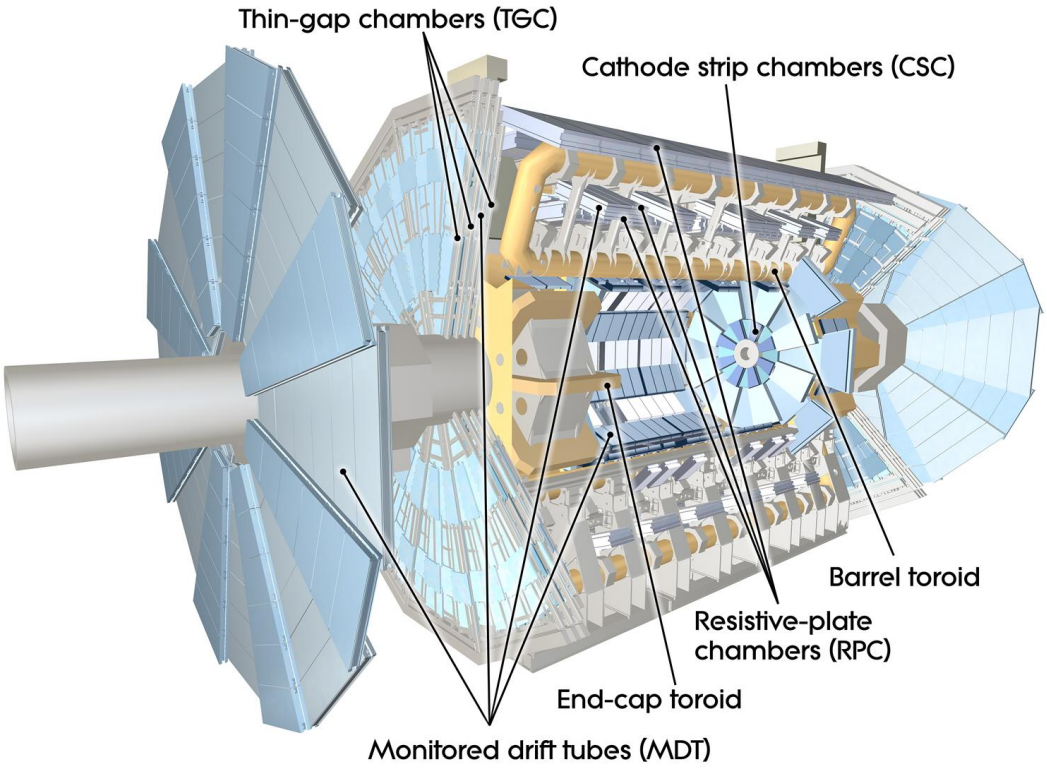


Figure 1.6: The layout and visualization of the muon spectrometer chambers of the ATLAS experiment. The figure was taken from [19].

Chapter 2

Standard Model and the Top Quark

The Standard Model (SM) is a theory of particle physics where the particles consists of three generation of quarks and leptons and interactions mediated by the corresponding gauge bosons. The last discovered particle is the Higgs boson, which in theory gives the other particles its mass by interaction with the Higgs field [20, 21, 22]. The overview of known elementary particles is shown in Fig. 2.1. The motivation for top quark measurement, its properties and production on LHC is presented in this chapter as well as the latest analyses results.

2.1 Standard Model

The Standard Model is successful explanation of the basic principles of the matter. The SM is based on interactions of strong (Quantum chromodynamic theory), weak and electromagnetic (Qantum Elecroweak theory) force on the sub-atomic scale mediated by gauge particles [23, 24, 25, 26, 27, 28] and using quarks and leptons as basic building blocks for particles, first described by Gell-Mann and Zweig in 1964 [29, 30] and later incorporated in the QCD theory, which postulates that SM hadrons are composed from quarks and anti-quarks. The history of the SM formulation is described for example in [31], which was confirmed in 1974 by the discovery of c-quark in the particle named as J/ψ [32, 33]. Even though the SM was describing matter on sub-atomic scales, there were missing undiscovered particles, τ lepton [34], bottom [34] and top quark [35, 36] and Higgs boson [37, 38]. The overview of basic particles in the Standard Model is shown in Fig. 2.1. Although, SM has been improved over the years, there are phenomenons, which the theory is unable to explain and thus opening the space for the theories Beyond the Standard Model (BSM) such as supersymmetry theory [39], string theory [40] or theory of composited Higgs boson [41] and more.

2.2 Top Quark

The top quark is the heaviest quark in the SM and its discovery was announced in 1995 by the CDF and D0 experiments at the Tevatron collider [35, 36]. The experiments at Tevatron were the first to measure the top quark properties. Later, when the LHC started to operate, ATLAS and CMS experiments joined the efforts [43]. The top quark mass combined results from aforementioned experiments is $M_{\mathrm{t}} = 172.69 \pm 0.30$ GeV [44], which is an important parameter in studies of the stability of the universe [45], thus the interest in increasing the precision of the measurement. The top quark, with

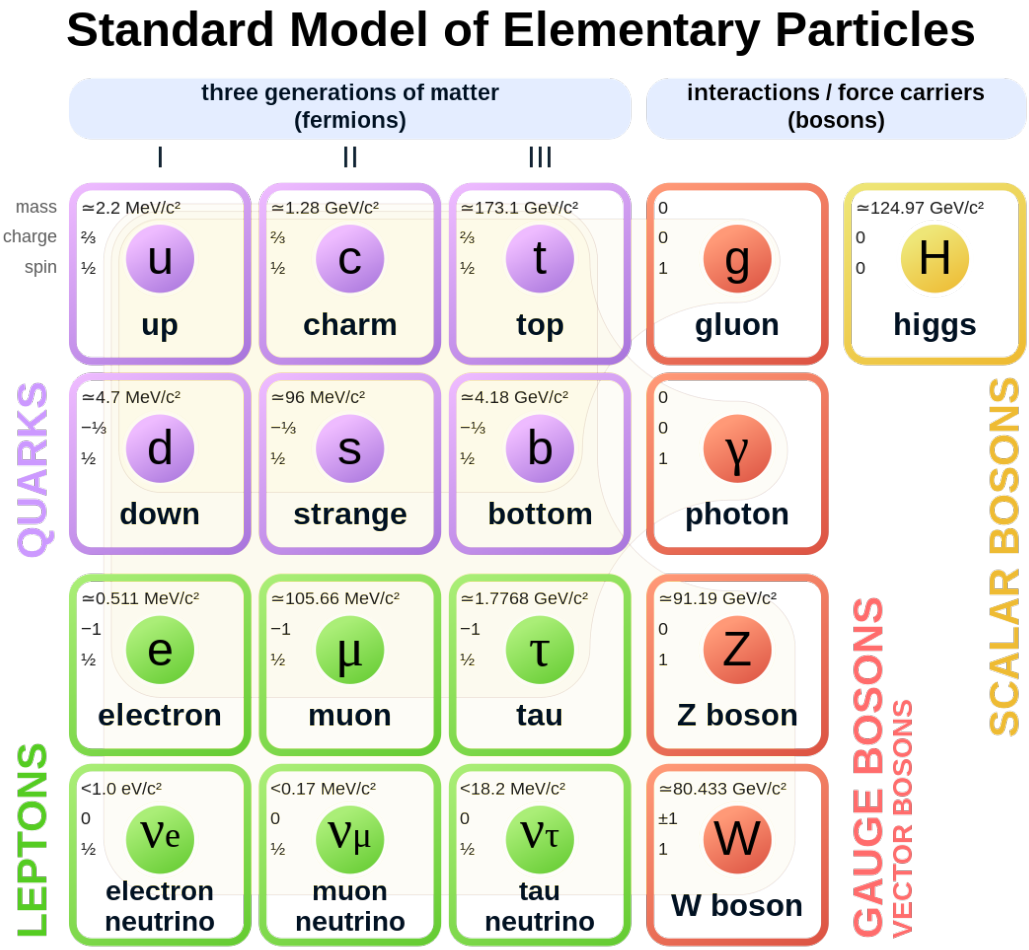


Figure 2.1: Overview of the Standard Model basic particles with additional information about their charge, mass and spin. The figure was taken from [42].

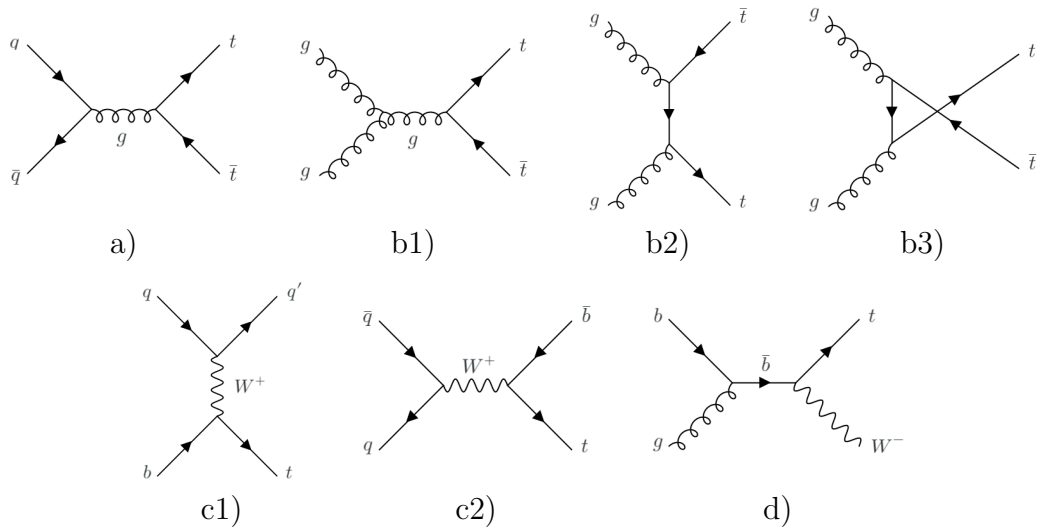


Figure 2.2: Feynman diagrams of top quark production on the LHC at CERN for $t\bar{t}$ pairs (top row) and for single top (bottom row). The $t\bar{t}$ production quark-antiquark annihilation is described by diagram a), the gluon-gluon production diagrams are shown in b1) for s-channel, in b2) for t-channel and in b3) for u-channel. The single top production is described in diagrams c1) for t-channel and c2) for s-channel, the associated single top production with W boson is shown in diagram d).

its large mass in comparison to other quarks and very short lifetime, decays before the hadronization can occur, thus the measurement of its properties can not be done directly. The large mass is responsible for largest coupling to the Higgs boson in comparison to the rest of the quarks, thus being important for measuring Higgs boson properties and for BSM theories.

2.2.1 Top Quark Production at the LHC

The top quark production at the LHC can be categorized to single top production and top anti-top quark pair $t\bar{t}$ production, but any combinations of those, for example top anti-top quark pair production with associated single top quark production or two top anti-top quark pairs production, also referred to as the four top quark production, is possible but with smaller probability. The $t\bar{t}$ pairs are generated either by the quark anti-quark annihilation or gluon-gluon fusion, see the corresponding Feynman diagrams in Fig. 2.2 a) and b1) to b3) for different production channels. Those production channels are describing the momentum conversion from initial state particles to the final state particles and are named after the Mandelstamm variables [46] s-channel, t-channel and u-channel. The $t\bar{t}$ production is dominated by the gluon-gluon fusion, which correspond to the 90% of the generated pairs, remaining 10% are events from the quark-antiquark annihilation.

The single top quark production, which is mediated through the W bosons and weak interaction, is shown in Fig. 2.2 c1), c2) and d) for main production channels and for associated production with W boson respectively. The dominating production channel is t-channel (Fig. 2.2 c1)), which corresponds to roughly 73% of the single top events, the s-channel (Fig. 2.2 c2)) adds only 3% of events. 24% single top events are produced via associated production with the W boson (Fig. 2.2 d)).

This thesis is aimed to study the top anti-top quark pair events on the LHC with the ATLAS detector, although the datasets measured at CERN laboratory are not directly used. The private analysis is based on experience from collaboration with ATLAS top quark working group, where I have contributed to the internal note, which resulted in publication [47], more details in section 2.4. The datasets used in the thesis are simulation of the processes ongoing on the LHC and the ATLAS detector.

2.3 Channels and Topologies

The production of the top anti-top quark pairs is separated into decay channels by the final products of the decay. The majority of top quarks and anti-top quarks, around 99%, decay to the W boson and bottom or anti-bottom quark and the decay happens before the top quark hadronizes or reaches the detection surface. Furthermore, the W boson decay either to the lepton and corresponding neutrino (leptonicaly decaying W boson) or quark anti-quark pair (hadronicaly decaying W boson). Similar notation is used to top quarks, in other words, leptonically decaying top quark has its daughter W boson decayed leptonically. The final decay products of W bosons determine the channel. There are two W bosons in each $t\bar{t}$ event which both decay either hadronically or leptonically leading to three combinations. If both W decayed hadronically, then we classify the event as an event in all-hadronic channel, if both W bosons decay leptonically, then the event belongs to the di-lepton channel. The semi-leptonic channel is defined by combination of one hadronically decaying W boson while the other W boson decay leptonically, which is the explored channel in this work.

The topology classification is needed for the analysis purposes after the process and channel is established. The topology is defined by the position of final states products in the detector, which is highly correlated to the momentum of the decayed particle. In the resolved topology all decay products are separated in the detector, no overlap of the jets is allowed. This topology often corresponds to decayed particles with lower momentum. The opposite boosted topology, where all product of the decay are reconstructed as one object, large jet, is corresponding to the decay of particles with large momentum. The semi-boosted topology lies in between those two, where some products are reconstructed as one object as in the boosted topology into the large jet, but rest of decay products are reconstructed separate as in the resolved topology.

Two hypothetical particles, vector boson Z' and scalar boson y_0, are used for better exploration of this energy region. These particles are used as probing mechanism of reconstruction and it is possible to tune energy of the event by setting its mass. The studied process can be described as follows: $pp \to Z'(y_0) \to t\bar{t}$, $t_1 \to W + + b \to \nu + \ell + b$, $t_2 \to W + b \to b + q + \bar{q}$. Topologies mentioned above are different in final products observed in detector especially on the hadronicaly decaying side. The resulting jets, typical for resolved topology coming from hadronic decay, are not overlapping each other in the detector thus forming small jets, see Fig. 2.3 a) for the graphical description. This behavior is standard for events from the beginning of energetic spectrum of $M_{Z'}$. In contrary, the boosted topology exploits the overlap of products and part of the definition of this topology is large jet, which contains all final products of hadronicaly decaying top quark, see Fig. 2.3 d) for schematic visualization. It is obvious, that there is condition on sufficient energy of top quark for this topology to have the final products collimated enough to form large jet. In between of those topologies lies interval of energy spectrum, where the resolved events are diminishing, but the energy is not yet high enough to form boosted event. This transition region

between those two topologies is preferred by the semi-boosted and semi-boosted mixed topology, where two products of hadronicaly decaying top quark form large jet, but one product lies outside this large jet and forms single small jet, the difference between those two topologies is in the origin of the escaping product. The decaying schemes and final products of studied events for semi-boosted and semi-boosted mixed topology are shown on figure 2.3, b) and c) respectively.

2.4 Analyses in Top Quark Final States

The top quark physics field is rich on different analyses ranging from measuring top quark properties, for example top quark mass, spin correlations and polarization, forward-backward and charge asymmetry and more, to production cross sections of the single top quark and top anti-top quark pair.

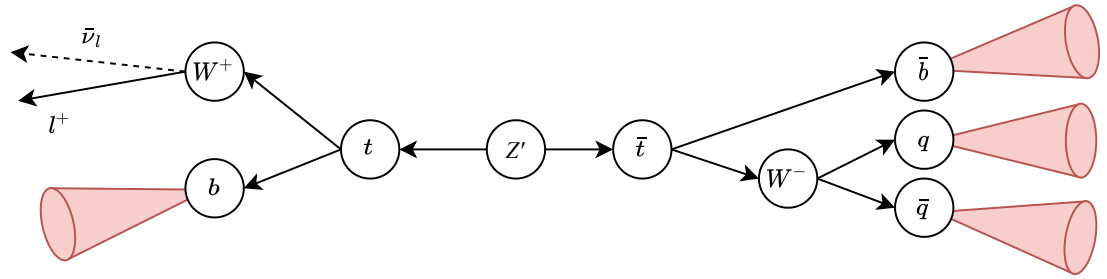
The top quark mass measurement is the most precisely measured property due to the measurement on all four main experiments on both at Tevatron and at LHC. The mass of the top quark is $M_{\mathrm{t}} = 172.69 \pm 0.30$ GeV according to the Particle Data Group [44]. This value is combination of measurements from Tevatron and LHC from different production channels [48, 49, 50, 51, 52, 53]. The top quark mass measurements are important for particle process generators as it is one of the parameters used for tuning the generated spectra to correspond to the measured data. The most precise measurements are from semi-leptonic channel as it has the best signal to background ratio and only one neutrino in the final state. One of the unique properties of the top quark decay is the spin correlation between top and anti-top quark born in pair, where the polarization is directly measured in angular distribution of their decay products. The results from LHC and Tevatron are listed here [54, 55, 56, 57, 58, 59, 60]

The single top quark production cross section has been first measured at Tevatron [61, 62], where the production is mainly through s-channel and t-channel whereas the main channel at LHC is the t-channel [63, 64, 65, 66, 67, 68, 69, 70]. The single top quark with associated W-boson production was measured at LHC only [71, 72, 73], but not at Tevatron due to the small cross section.

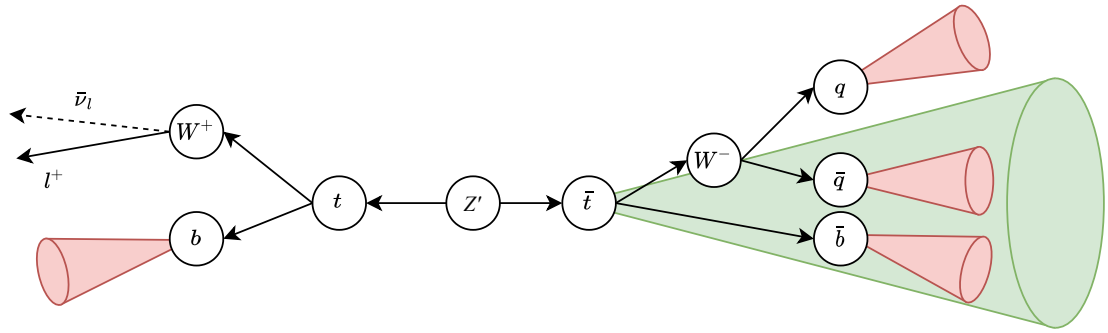
The top anti-top quark pair cross section has been measured by Tevatron [74, 75, 76] as well by LHC experiments namely ATLAS, CMS and LHCb [77, 78, 79, 80, 81, 82, 83, 84, 85]. The most precise measurements come from the di-lepton channel at Tevatron, exploiting the τ hadronic decay, as well at LHC, where the $e\mu$ combination has the best signal to noise ratio.

Thorough list of the analyses and their results can be found in [86]. The top quark physics future lies in precision measurements and its modeling as it is background for other analyses especially in Higgs boson analyses, but top quark maintain important role for the BSM theories as well.

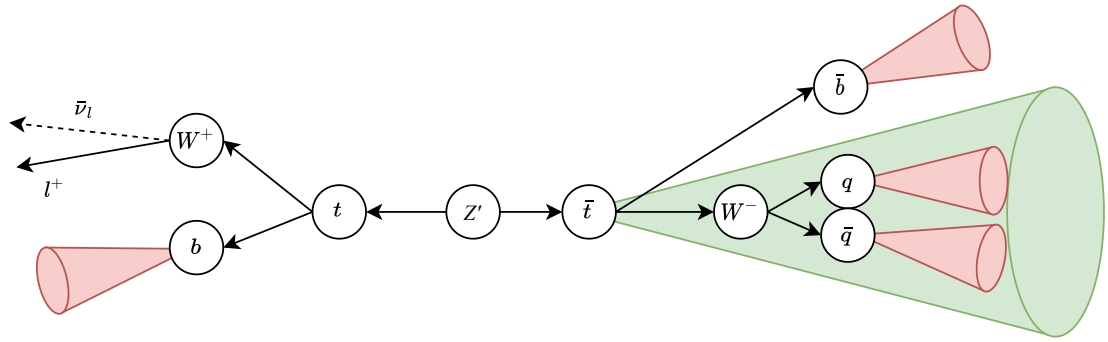
Results of measurement of the normalized differential cross section of top anti-top quark pair in [47] have been based on internal note, to which I have contributed. The analysis is performed on data from proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 8$ TeV and integrated luminosity of 20.3 fb^{-1}, recorded in 2012 with the ATLAS detector. Selected data are measured in the semi-leptonic channel, corrected for detector influences and compared with the simulation. The general agreement with the predicted values is apparent. My contribution to this analysis was implementation of several variables of interest to the analysis C++ code and optimization of binning for the unfolding procedure used for correction of the detector influences, for example the acceptance or the finite resolution of the detector.



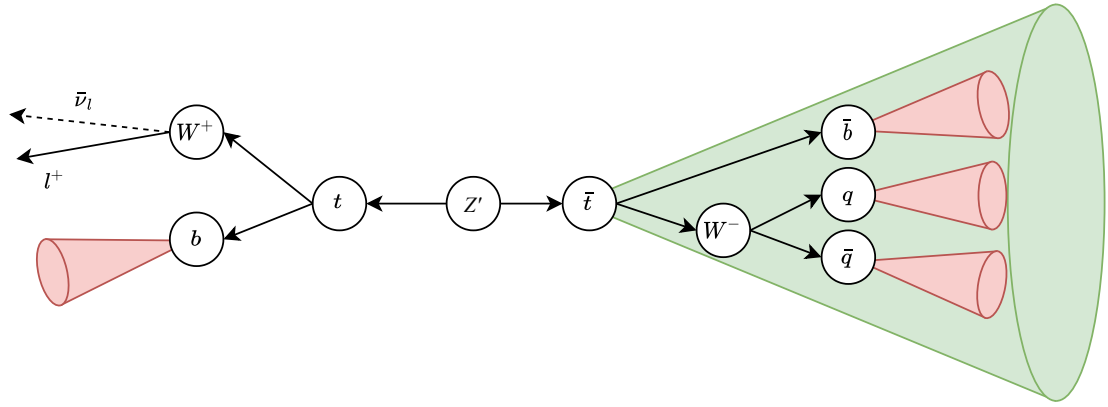
a)



b)



c)



d)

Figure 2.3: Schematics of an event $Z' \to t\bar{t}$ decay in resolved a), semi-boosted mixed b), semi-boosted c) and boosted d) topology.

Chapter 3

Objects Definition and Identification with the ATLAS Detector

The detection and identification of the particles and other objects used in the analyses is described in this chapter as well as the studied reconstruction levels. The identification is important to discriminate between final-state particles, originating from the studied process and other particles originating from hadronization and other background sources. There are three reconstruction levels describing the studied process.

The first, parton or generator (in simulation) level is corresponds to the theoretical point of view of the process and contains all the particles of the final-state before hadronization. The second is the particle level, which corresponds to the particles after the hadronization process and third, deals with particles and objects at the detector level with all imperfections of the detector.

The common practice is to compare theoretic prediction and measured spectra on the particle level. Parton shower and hadronization generators are used on the simulated hard process event to reach the particle level. On the other hand, detector corrections and unfolding procedure are used to reach the same level from the experimental point of view, which is described in detail in the chapter 6.

There are three categories of particles measured at the ATLAS detector on the particle level. First, stable particles, which are defined by the lifetime of the particle as $c\tau > 10$ mm in other words the particle is able to survive long enough to travel the distance from the interaction point to the detector, including hadrons. Second category are prompt leptons, which are generated final-state particles and not originating from hadronization. The prompt photons are in the third category, which are radiated in the studied process and are belong the final-state particles. Photons, produced by initial-state radiation are also included.

3.1 Leptons and Missing Transverse Energy

Leptons and neutrinos are crucial for reconstructing the W bosons, which may decay into a lepton and a neutrino pair. The detectors are able to directly detect leptons with exception of τ lepton, which decay before entering the detection areas of the detector although products of the decay are detected and there are method for its reconstruction. The detection of neutrinos is next to impossible and ATLAS detector is unable to detect them. However, there is a concept in the reconstruction

phase which replaces direct measurement of the neutrinos (at least their transverse momentum) by measurement of the imbalance in transverse momentum by summing four-vectors of all detected particles, which should equal to zero by definition due to the vanishing transverse momentum contribution of the incoming protons. This imbalance of momentum is caused by neutrinos escaping the detection and is equal to the negative sum of neutrino four-vectors from the studied final-state, called the missing transverse energy (MET), or momentum as the mass of neutrinos is negligible, and defined as follows

$$\overrightarrow{E}_{\mathrm{T}}^{\mathrm{miss}} = (-\sum_{\mathrm{i}} p_{\mathrm{x,i}}, -\sum_{\mathrm{i}} p_{\mathrm{y,i}}), \tag{3.1}$$

where $\overrightarrow{E}_{\mathrm{T}}^{\mathrm{miss}}$ is the missing transverse momentum vector, $\sum_{\mathrm{i}} p_{\mathrm{x,i}}$ is summed contribution of all detected particles momentum alongside x-axis and $\sum_{\mathrm{i}} p_{\mathrm{y,i}}$ is summed contribution of all detected particles momentum alongside y-axis. Both leptons and MET, are used in the reconstruction of the top anti-top quark pair events in the ATLAS analyses and in this work.

There are three definitions of leptons divided by the reconstruction steps. First are born-level leptons, which are particles before the final-state radiation, which corresponds to the Feynman diagrams without associated QED photon radiation of products of the process. Bare leptons are stable particles corresponding to the particle reconstruction level and may differ from the born leptons by the energy carried away by the emitted photons from the lepton. Dressed leptons are stable particles corresponding to the particle level as well, but the energy of photons in small cone around lepton are added back to the lepton to achieve better precision of the measurement. In other words the dressed leptons are prepared from bare leptons by correcting their energy for the photon radiation losses in hadronization process. Analysis described in this work is using bare leptons from the simulation and uses private dressing correction afterwards before the reconstruction of the event at the particle level, see section 4.3.1 for more information.

3.2 Jets

Most of the physics results performed at the ATLAS detector use jets as physics objects. The reconstructed objects hadronic final states known as jets, are products of dedicated jet reconstruction algorithms. There are three jet levels, as illustrated in Fig. 3.1. Parton level jet is constructed from constituents before the hadronization, *i. e.* quarks and gluons, next level is a particle level jet, which is built using stable particles, third level is the detector level, where the input for the jet algorithm are deposits of the energy in calorimeter.

3.2.1 Jet Reconstruction Algorithms

The reconstructed jet at any level can be described by its momentum and angular variables used to evaluate the direction of jet. The first variable of interest is the transverse momentum (p_{T}), which is the momentum of the particle in the transverse plane with respect to the direction of the incoming colliding protons. The sum of transverse momenta of all particles in event is expected to be zero, which results from the law of conservation of momentum. The higher the transverse momentum is, the

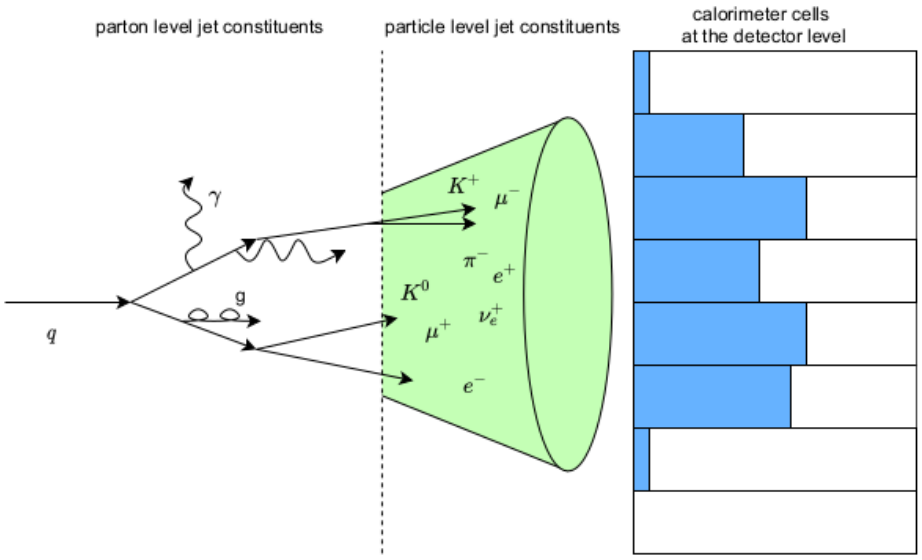


Figure 3.1: The visualization of the particle reconstruction levels from parton (particles before hadronization) through particle level (particles after the hadronization) to the detector level (measured particle tracks and deposits in calorimeters).

higher energy deposit in calorimeter is for a fixed direction of the jet. Two angular variables are used for the description of the direction of the jet. First variable is the azimuthal angle ϕ in the transverse plane in regard to the direction of incoming protons in the detector. The cylindrical ATLAS detector is symmetric in this variable which thus has no influence on the reconstruction of the jet. Second variable called pseudorapidity describes the elevation of the jet from the direction of the incoming protons and is defined by the following formula:

$$\eta = -\ln \tan \frac{\theta}{2}, \tag{3.2}$$

where θ is the angle or between the jet momentum direction and the beam axis, but the angle measured in laboratory frame may differ by the boost of the particle alongside the beam axis and is denoted as production angle θ^*. The jet direction lies in the transverse plane to the direction of the incoming colliding protons for $\eta = 0$. There are several jet reconstruction algorithms. These algorithms use iterative method for clustering jet constituents [87]. The condition for clustering is a metric defined as follows:

$$d_{\mathrm{i,j}} = \min(p^{\mathrm{q}}_{\mathrm{T,i}}, p^{\mathrm{q}}_{\mathrm{T,j}}) \frac{\Delta y^2_{\mathrm{i,j}} + \Delta \phi^2_{\mathrm{i,j}}}{R^2}, \tag{3.3}$$

where $p_{\mathrm{T,i}}$ and $p_{\mathrm{T,j}}$ are transverse momenta of examined constituents i and j, q is the parameter for different clustering algorithms and R is the distance parameter, roughly corresponding to the radius the resulting jet area in the calorimeter and y is pseudorapidity defined in high energy physics as

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}, \tag{3.4}$$

where E is energy and p_z is longitudinal momentum of the studied object, corresponding to direction alongside the LHC ring and denoted as the z-component of the momentum. Pseudorapidity converges to the rapidity in approximation $m \ll p$, where m is mass and p is momentum of the studied object. The algorithm then sums the four-momenta of the two constituents with smallest $d_{i,j}$ and removes them from the list of objects for reconstruction and adds the resulting constituent into the list and repeats the procedure.

We recognize three basic algorithms in dependence on the parameter q in the exponent of the momentum. If the exponent is set to one, the algorithm is known as Kt algorithm. This algorithm reconstructs jets by clustering together softer particles first. This leads to a less compact shape of the jet. Second algorithm, known as the AntiKt, has the parameter q in the exponent set to minus one. The AntiKt reconstruction algorithm is generally preferred algorithm used in physics analyses thus there is need for the correction factors derivation for this algorithm. The reconstruction algorithm clusters the softer constituent first with constituents with higher transverse momentum instead of clustering them among themselves. The third algorithm is called Cambridge/Aachen algorithm, where the parameter q is set to zero.

3.2.2 Jet Correction Chain and Jet Energy Scale

The energy of a jet registered by the detector is not measured precisely, there are several influences due to the detector and the nature of collisions. First, there are pile-up events and particles, such as products of non-"hard" process, which form additional to the energy in the cells of the calorimeter. There is always a fraction of soft jet constituents coming from final or initial state radiation. These constituents are radiated by initial colliding partons, called initial state radiation, or by partons produced in the final state. There is a correction applied to take into account this influence and called the pile-up subtraction, lowering the deposited energy in cells of the calorimeter by an offset term, ρ in the correction formula 3.5, which is evaluated on event-by-event basis in dependence on η and the jet area A. Second, two residual corrections have to be applied as there is still dependence of the previous correction on jet transverse momentum and position in calorimeter. The first one, known as a μ-term, corrects for energy of products of particles coming from previous "hard" process, depending on average number of proton-proton interactions per bunch crossing $\langle\mu\rangle$ and is described by β in 3.5. The second is the residual pile-up correction, when there are more than one "hard" process, depending on number of reconstructed primary vertices and is represented by parameter α in 3.5, a so-called NPV-term. Reconstructed primary vertices are reconstructed positions where the proton-proton collision have occurred. In summary these are called pile-up and residual pile-up corrections and are often rewritten as

$$p_{\mathrm{T}}^{\mathrm{PU-corr}} = p_{\mathrm{T}}^{\mathrm{det}} - \rho \times A - \alpha \times (N_{\mathrm{PV}-1}) - \beta \times \langle\mu\rangle\,, \tag{3.5}$$

where $p_{\mathrm{T}}^{\mathrm{PU-corr}}$ is the transverse momentum of the jet corrected for the pile-up influences.

The dependence of the terms (on p_T and η) is evaluated with the other term fixed, *e.g.* the NPV term is fixed when evaluating dependence of the correction on μ.

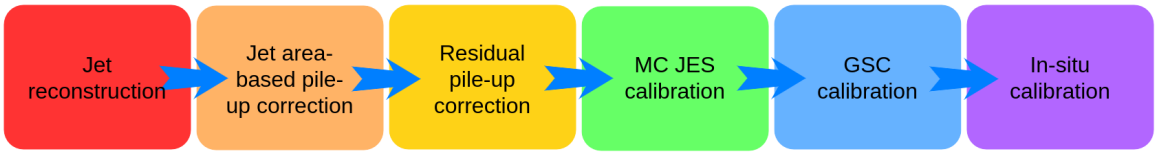


Figure 3.2: The flowchart of corrections applied to the jets.

Next, the most important correction, is the jet energy scale correction, which describes the energy response of the detector as function of the angle in the detector and energy of the jet. The response of the detector is defined as the average of the ratio of the jet energies on detector level over the particle level.

The fourth step in jet calibration is the Global Sequential Calibration, which takes into account the flavor of the original parton from which the jet was formed. This step was added due to high response difference between gluon and quark initiated jets. This difference formed the largest addition to the systematic uncertainties before the Global Sequential Calibration was introduced. The last step of correcting is the in situ correction, which uses not only simulation but measured data as well. This step uses the balance of energy of physical objects in the transverse plane of the detector. The summed energy from all products of a collision in the transverse plane should be equal to zero. The ratio between the Monte Carlo simulated samples balance and data balance gives the correction value, which is then applied to the data. The physical objects for the in situ techniques are usually photons or Z bosons produced in association with jets [88]. This article describes the second and third step and evaluation of correction factors values. The workflow chart is described in Fig. 3.2

After the application of the residual pile-up correction the JES response step follows. The JES correction factors are evaluated by a specialized package, which I helped to operate during the qualification task for the ATLAS collaboration. Those factors provide information about the ratio between the detector and the particle level energy. This ratio describes ineffectiveness of the calorimeter in detection of the jet energy. Different jets are used for this derivation including jets with lower transverse momentum thresholds for better evaluation of correction factors around values where the threshold values of jet transverse momentum for physics analyses lie. Correction factors for the ATLAS detector data are derived for each jet collection, namely electromagnetic scale (EM scale) and particle flow (PFlow) for small-R jets and local cluster weighted (LCW) scale for large-R jets [87]. The response of jets is usually lower for lower energy jets and then the behavior is expected to rise, but not to exceed unity on average. The method to calculate the response uses information from Monte Carlo simulation and forms the ratio between values of the jet energy measured by the simulated detector, introduced as detector level value (E_{det}), and energy of the particle level jet, often referred as particle level value (E_{ptcl}). Matching of jets for the response calculation is done in following steps. First, the jet at the detector level is chosen, then there are chosen matching candidates on particle level complying the distance parameter cut (R_{cut}) defined as follows

$$R_{\mathrm{cut}} > \Delta R = \sqrt{\Delta y^2_{\mathrm{det,ptcl}} + \Delta\phi^2_{\mathrm{det,ptcl}}}, \tag{3.6}$$

where variables with “det” and “ptcl” subscript corresponds to detector level and

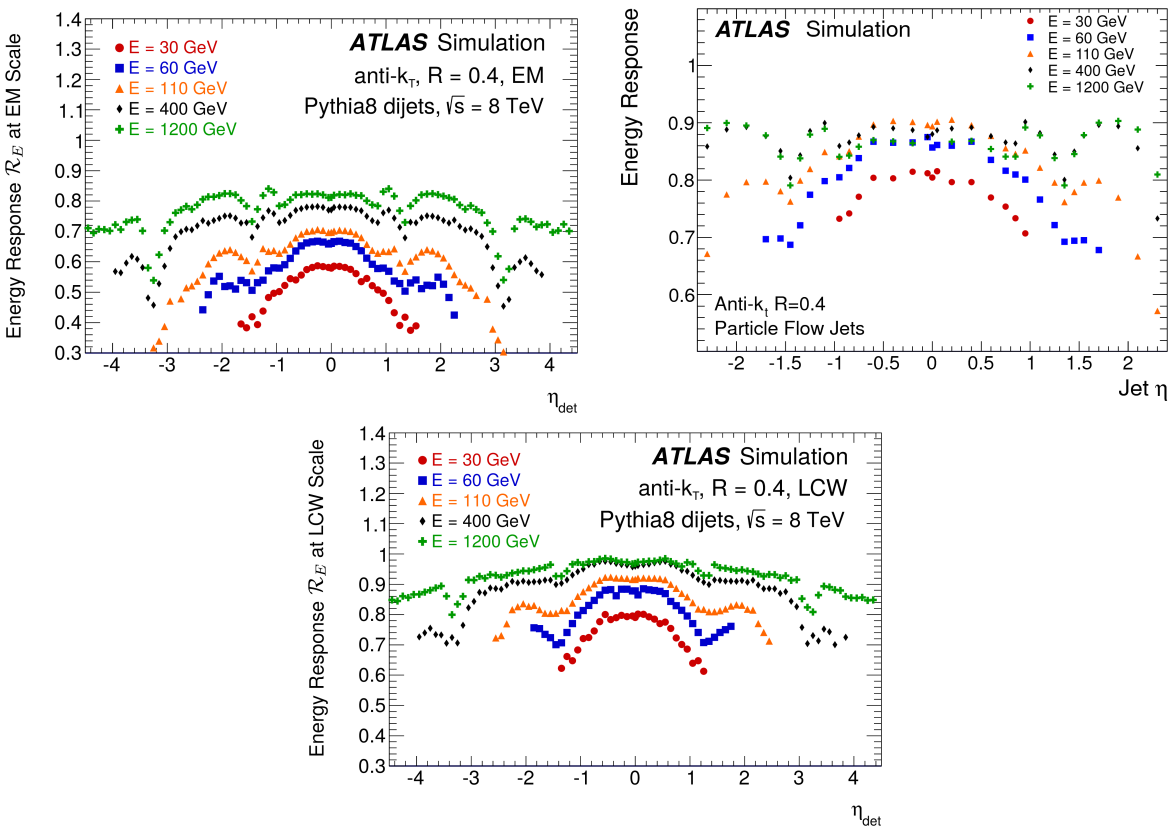


Figure 3.3: The response $\mathcal{R}_E$ of detector level jets in dependence on the energy and η of the jet for the EM scale (top left), PFlow (top right) and LCW scale (bottom) jet collections. The figure was taken from [89, 90].

particle level, respectively. The particle level jet candidate with the smallest distance parameter is chosen as a matching jet. The jet response is defined as

$$\mathcal{R}_E(\eta,p_{\mathrm{T}}) = \left\langle \frac{E^{\mathrm{jet}}_{\mathrm{det}}}{E^{\mathrm{jet}}_{\mathrm{ptcl}}} \right\rangle_{\mathrm{events}} . \tag{3.7}$$

The inverse value of response $\mathcal{R}_E(\eta,p_{\mathrm{T}})$ is the JES correction for given transverse momentum in a specified area in the detector. The algorithm prepares histograms of the response values corresponding to the energy interval specified by user and each histogram is fitted by a Gaussian function as there is expectation that the responses are normally distributed. The mean of the fit is plotted against the reconstructed energy and fitted by a polynomial logarithmic function. This function is used to interpolate response for any given energy. The response depends on energy and position in the detector described by the variable η and its shown in Fig. 3.3 for EM scale (top left), PFlow (top right) and LCW scale (bottom).

The LCW scale jet collection has better response of the detector as mentioned before. The η range of the PFlow jet collection is shorter as the response is about the same from $|\eta| > 2.4$ with the EM scale collection when there is no longer available the information from tracking detectors for PFlow jet collection. The difference between collections in forward regions may be result of few more aspects, as influence of the lower threshold of the jets incoming to the calibration, which was different for each of those collections as well there may be misbehavior in reconstruction of different

collections which is yet to be understood. The drops in response for $|\eta|$ values around 1.3 and 3.2 are present in all jet collections and are caused by the construction of the detector due to the embedding of parts of detector in these rapidity regions. Detailed information can be found in [91, 89, 90].

As a part of the qualification task I was responsible for the JES factors derivation for the aforementioned jet collections for run 2 and prepared the figures to the internal note on which the following publication [92] was based.

3.2.3 *b*-tagging of Jets

The presence of the *b*-hadrons in the jet is used for the identification of jets originating from the *b*-quark, which is widely used in analyses and may be used in the new physics studies. Top quark analyses are not an exception as the main decay products of top quark is to the bottom quark and W boson. Techniques for determination of the originating particle helps in the reconstruction processes in the analyses and there are three main approaches to the problem.

One approach, based on impact parameter algorithm, is through association of tracks to the position of the primary vertex. The impact parameter is the longitudinal (z_0) or transversal distance (d_0), with respect to the beam, of the track to the primary vertex. The lifetime of the B-hadron, which is around 1.5 ps and its subsequent decay to the secondary vertex, which is displaced from the proton-proton collision, is exploited. The tracks not associated to the primary vertex are used for the probability density function calculation of the original particle flavor by comparison with probability density function of reference histograms from the Monte Carlo simulation.

The second vertex finding algorithm is based on the discrimination of the flavor of jets originating from *b*-quark or B-hadron using the lifetime of aforementioned B-hadron as in previous case. Reconstruction of the secondary vertex, which is displaced from the primary vertex, combines information from tracking detector as well from the calorimeter.

The third approach called topological multi-vertex finding algorithm exploits the topological structure of the weak decay inside the jet and tries to reconstruct the whole decay process from the B-hadron approximation through its path and vertex position estimate.

The more sophisticated methods are relying on the machine learning, namely boosted decision tree and deep feed-forward neural networks, both using results from previously mentioned methods as the input for training.

The efficiency of the *b*-tagging, which corresponds to the probability of a successful jet *b*-tagging, is very important for the analyses, most of the *b*-taggers operate between 60 - 85% operation point, tested on sample with top anti-top quark pair events [93].

Chapter 4

Private Analysis

A private analysis is described in this chapter. It aims to study the process where top anti-top quark pair decays in the ℓ + jets channel and its performance in different topologies with focus on the semi-boosted one. The motivation for this study is the enhancement of the measurable events by adding the semi-boosted and semi-boosted mixed topology into the analyses in the top anti-top quark final states, which would be otherwise discarded. The comparison of the performance of the unfolding procedure between topologies is present with its evaluation of influence on the signal strength of the presence of the simulated hypothetical particle. The hypothetical particles Z' and y_0 are used as a probe, where the different values of mass of the hypothetical particle are used to tune in the desired energy region for topology studies. Its reconstructed mass peak is used as the tool for resolution studies. This chapter contains description of preparation of the simulated samples, the simulation of the ATLAS detector, private JES correction derivation and the event selection followed by the reconstruction description in all four studied topologies.

The analysis framework consists of three main parts: The preparation of the samples, the selection and reconstruction of the selected events and the unfolding procedure. The samples are prepared with the MadGraph5 package and the simulation of the ATLAS-like detector is performed with the DELPHES package, see section 4.1 for more information. The framework has been implemented with a private algorithm for JES factors derivation and application. Control plots from this procedure are presented for small and large jets in section 4.2. Jets are enter the events selection after the application of this JES correction and tagging on the original particle is performed on the large jets enhance the performance of the event selection, see section 4.3.2 for further information. The event selection is described in section 4.3 and the reconstruction of the $t\bar{t}$ kinematics in final section of this chapter 4.4.

4.1 Generated Samples

Samples were simulated by the MadGraph5 2.6.4 software package [94] with specialized model package (`Abelian_Higgs_Model_UFO`) for simulation of a hypothetical vector particle Z' and scalar particle y_0 [95, 96, 97]. Samples were prepared for different values of mass of this hypothetical particle, ranging in interval of 500 – 1000 GeV.

All samples hold information from the both detector and particle levels. The detector level corresponds to measured data with imperfections as coming from a close-to-real detector. Particle level describes particles just before entering the detector.

Events are simulated to correspond to proton-proton future collisions at the Large

Table 4.1: The cross-section, processes details and the generated number of events for the samples generated by the MadGraph5 package; c.c. stands for the charge conjugation and ℓ for the electron or muon.

Sample	Cross-section[pb]	Generated process	Events
Wbb+jets	153.43	$pp \to W^+ + j, W^+ \to l^+ + \nu_l$+c.c.	655 855
$WWbb$+jets	180.31	$pp \to W^+W^-b\bar{b}, W^+ \to l^+ + \nu_l, W^- \to jj$+c.c.	1 000 000
Z'(M = 1000 GeV)	0.000061	$pp \to Z' \to t\bar{t}, t \to bjj, \bar{t} \to \bar{b}l^-\bar{\nu}_l$+c.c.	500 000
Z'(M = 900 GeV)	0.000083	$pp \to Z' \to t\bar{t}, t \to bjj, \bar{t} \to \bar{b}l^-\bar{\nu}_l$+c.c.	500 000
Z'(M = 800 GeV)	0.00012	$pp \to Z' \to t\bar{t}, t \to bjj, \bar{t} \to \bar{b}l^-\bar{\nu}_l$+c.c.	500 000
Z'(M = 700 GeV)	0.00016	$pp \to Z' \to t\bar{t}, t \to bjj, \bar{t} \to \bar{b}l^-\bar{\nu}_l$+c.c.	500 000
Z'(M = 600 GeV)	0.00023	$pp \to Z' \to t\bar{t}, t \to bjj, \bar{t} \to \bar{b}l^-\bar{\nu}_l$+c.c.	500 000
Z'(M = 500 GeV)	0.00033	$pp \to Z' \to t\bar{t}, t \to bjj, \bar{t} \to \bar{b}l^-\bar{\nu}_l$+c.c.	500 000
y_0(M = 1000 GeV)	0.031	$pp \to y_0 \to t\bar{t}, t \to bjj, \bar{t} \to \bar{b}\ell^-\bar{\nu}_\ell$+c.c.	500 000
y_0(M = 900 GeV)	0.053	$pp \to y_0 \to t\bar{t}, t \to bjj, \bar{t} \to \bar{b}\ell^-\bar{\nu}_\ell$+c.c.	500 000
y_0(M = 800 GeV)	0.091	$pp \to y_0 \to t\bar{t}, t \to bjj, \bar{t} \to \bar{b}\ell^-\bar{\nu}_\ell$+c.c.	500 000
y_0(M = 700 GeV)	0.16	$pp \to y_0 \to t\bar{t}, t \to bjj, \bar{t} \to \bar{b}\ell^-\bar{\nu}_\ell$+c.c.	500 000
y_0(M = 600 GeV)	0.27	$pp \to y_0 \to t\bar{t}, t \to bjj, \bar{t} \to \bar{b}\ell^-\bar{\nu}_\ell$+c.c.	500 000
y_0(M = 500 GeV)	0.41	$pp \to y_0 \to t\bar{t}, t \to bjj, \bar{t} \to \bar{b}\ell^-\bar{\nu}_\ell$+c.c.	500 000
$t\bar{t}$+jets	178.6	$pp \to t\bar{t}, t \to bjj, \bar{t} \to \bar{b}\ell^-\bar{\nu}_\ell$+c.c.	2 934 961

Hadron Collider (LHC) at CERN with the center of mass energy $\sqrt{s} = 14$ TeV. Jets are the experimental signatures of hadronic final states of quarks and gluons, which form particle showers entering the detector and leaving their energy there. Jet constituents are clustered energy deposits in calorimeters or stable particles at the particle level. Small jets, with distance parameter $R = 0.4$, in all signal and background samples were generated with a cut on the transverse momentum on generator level of $p_{\mathrm{T,SJ}} > 20$ GeV to increase amount of events of interest in the detector simulation. Samples containing the hypothetical particle Z', Wbb+jets background sample and a $t\bar{t}$ sample were all generated at the next-to-leading order (NLO) of QCD, while samples with the hypothetical particle y_0 and $WWbb$+jets background sample were generated at the leading order (LO) only. The y_0 was simulated at the leading order due to the triangle loop process for which the NLO level was unavailable in the MadGraph5 simulation and the cause for the $WWbb$+jets background sample is a small contribution to the events reconstructed by the analysis framework.

The luminosity of the sample is calculated as follows:

$$L = \frac{N}{\sigma}, \tag{4.1}$$

where L is the luminosity of the sample, N is the number of generated events and σ is cross-section of the sample.

Spectra from different samples may be weighted to the same luminosity for comparison purposes or for the admixture for the unfolding procedure. The weight is defined as $L_{\mathrm{ref}} = wL_{\mathrm{w}}$ where L_{ref} is the luminosity of the reference sample, w is the weight for spectra scaling and L_{w} is the luminosity of the weighted spectrum sample.

4.1.1 Detector Simulation

The detector simulation is facilitated by the DELPHES simulation package version 3.4.1 [98], which emulates detector effects. Both the large and small jets, were reconstructed by FastJet algorithm [99] using jet distance parameter $R = 1$ and $R = 0.4$, respectively. The prepared samples were used as input for a private framework based on ROOT 6.16 [100], where the results were analyzed and saved to output files from which the plots were drawn afterwards. The DELPHES simulation consist of evaluating tracks of the particles in the magnetic field, hadronic and electromagnetic calorimeter response tuned to correspond to the ATLAS detector as well as the muon identification in spectrometers. Deposits in the calorimeter are consistent with measurement of jets, electrons and are used for the calculation of missing energy. The jet b-tagging is also present in the package for different working points corresponding to the ones used in the analyses. The response of calorimeters is used to calculate jet energy scale correction function, which is slightly worse than the JES functions used in the private analysis for small jets especially in high η regions. The DELPHES-provided JES should not be used in case of large jets, where are the large discrepancies in closure tests.

The ATLAS detector template card for DELPHES configuration was used, but with adjusted setting to cover additional information, namely B-hadron information, Z' particle information, which was used for the verification of the framework, and photon information. The b-tagging working point was set to correspond to the 80% working point, whereas the DELPHES-provided JES correction was performed by following formula:

$$\mathcal{JES} = \sqrt{1 + \frac{a - b * |\eta_{\mathrm{J}}|^2}{p_{\mathrm{T,J}}}}, \tag{4.2}$$

where $\mathcal{JES}$ is the jet energy scale correction factor for the jet, parameter a is set to 3 GeV, parameter b value is 0.2 GeV, η_{J} is the pseudorapidity of the jet and $p_{\mathrm{T,J}}$ is the transverse momentum of the jet in GeV.

4.2 Private Jet Energy Scale

The DELPHES-provided JES correction for large jets is not used at all, but the DELPHES JES (JetJES collection) for small jets is, but the ratio of the small jet energy on the detector level over the energy of the matched particle jet plotted against the small jet η on the detector level is showing large dependence on η as it is shown in Fig. 4.3 left, thus, the small jets are corrected for this phenomenon in same manner as large jets by private additional JES correction factors.

The jet energy scale correction used in the analysis framework has been derived from the $t\bar{t}$ sample specially generated for this purpose with lower cut on the minimal transverse momentum of jets (15 GeV) to enhance performance of JES correction for jets with lower transverse energy, namely around 25 GeV. The JES factor derivation is performed in a similar manner to the ATLAS working group method and amplified both large and small DELPHES jets.

The first step of the JES factors derivation is matching the jets at the detector level with jets on particle level. Matching is performed by choosing the jet with largest transverse momentum at the detector level and then finding the jet from particle collection, which is angularly closest to the detector level jet in terms of ΔR defined in

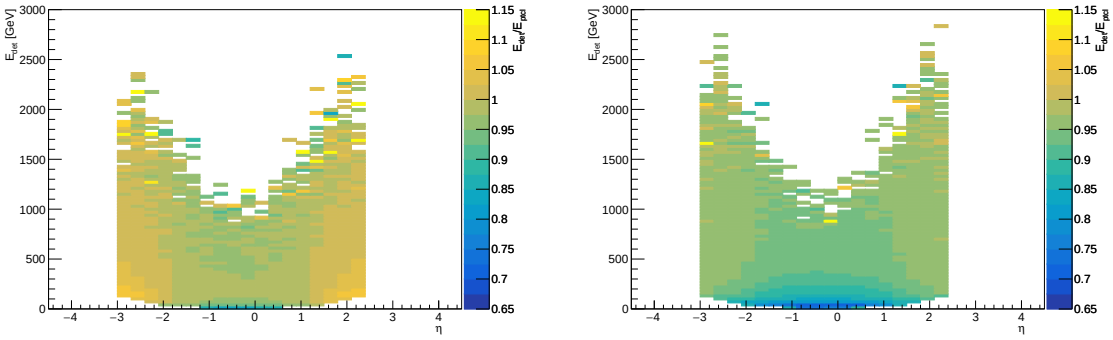


Figure 4.1: The visualization of the response of the detector for jets vs. η and energy on the detector level for small jets (left) and large jets (right).

equation 3.6. Matched jets are separated according to their η value. The η interval in which JES factors are derived is set to 0.3, which is coarser than at ATLAS but sufficient for the private analysis purposes. The response of the detector is evaluated for each of the η interval, but it is not only depending on the η but on the energy of the jet as well. The response for each combination of energy and η is forming one dimensional plot, which is fitted by the Gaussian curve to obtain the mean value corresponding to the JES factor for the given energy bin. The response factors are then fitted by an exponential function to obtain the response function which allows to correct any jet energy in the studied range. Those response functions are stored in separate calibration file and each jet used in the private analysis is corrected. The example of the fitted response of the detector function for the central region and region with higher η is shown Fig. 4.2 for small jets and for large jets. The general visualization of the response of the detector correction is shown in Fig. 4.1 (left) for small jets and (right) for large jets. The general trend is that the jets in the central area of ATLAS detector are better corrected at the lower energies whereas the jets in the more forward region are better corrected at the larger energies.

The closure test, the ratio of corrected small jet energy at the detector level over the energy of the matched small jet at the particle level plotted against the small jet η at the detector level is shown in Fig. 4.3 (right). For the large jets, the same is shown in Fig. 4.4.

Closure tests show that the deviation from unity from ratio between the corrected detector level small jet and the generated particle level matched small jets are under 5% value in both, transverse momentum ($p_{\mathrm{T,J}}$) and η closures, whereas the deviation for large jets is over 5% but only for the transverse momentum values which are not used in private analysis. Overall performance of JES correction of detector level jets, small and large, is in fair agreement with the particle level generated matched jets, respectively, used in the private analysis.

4.3 Event Selection

Events considered in the analysis are reconstructed at two levels; once with the DELPHES ATLAS detector simulation, forming detector level spectra, and at the particle level. The event selection and the requirements differ slightly for the recon-

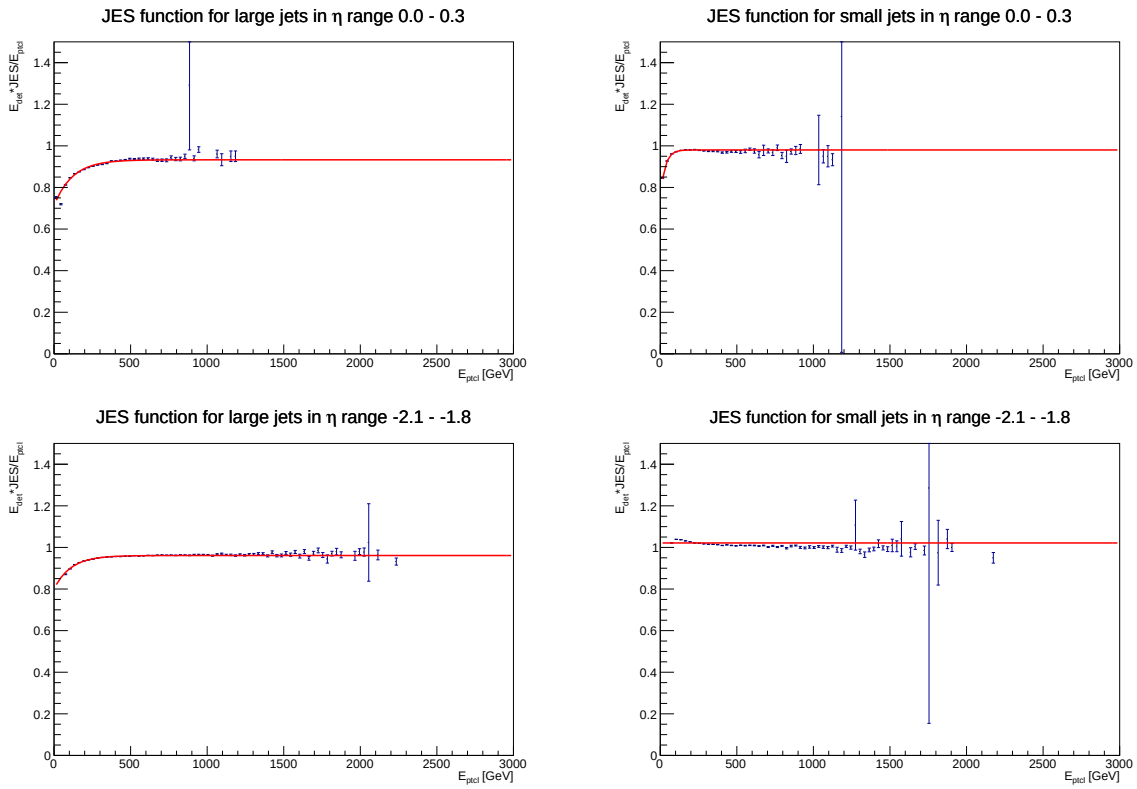


Figure 4.2: Derived factors fitted by the response function in the central ATLAS region η bin (top row) and more forward region η bin (bottom row) for large (left) and small (right) jets.

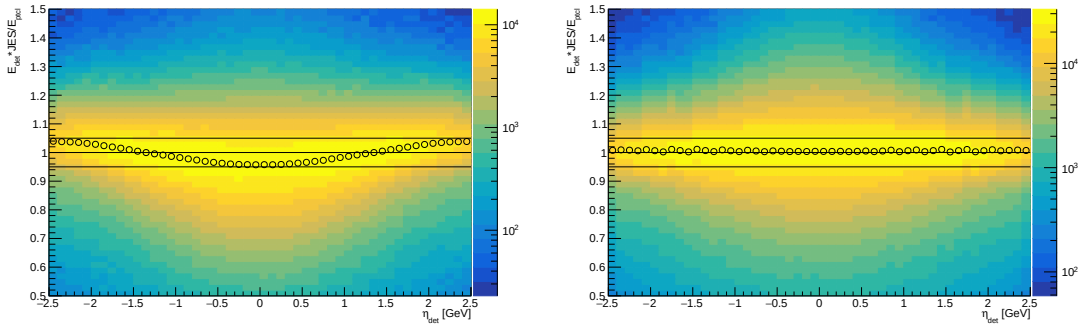


Figure 4.3: Closure test,*i. e.* the ratio of energy of JES corrected jet at the detector level over energy of matched jet at the particle level, of small jet energy corrected for the JES in dependence on η by the DELPHES package only (left) and by the DELPHES-provided JES with the application of the private JES (right), tested on the $t\bar{t}$ sample. The horizontal lines denote the 5% difference between the corrected detector level jet and the generated particle level matched jet.

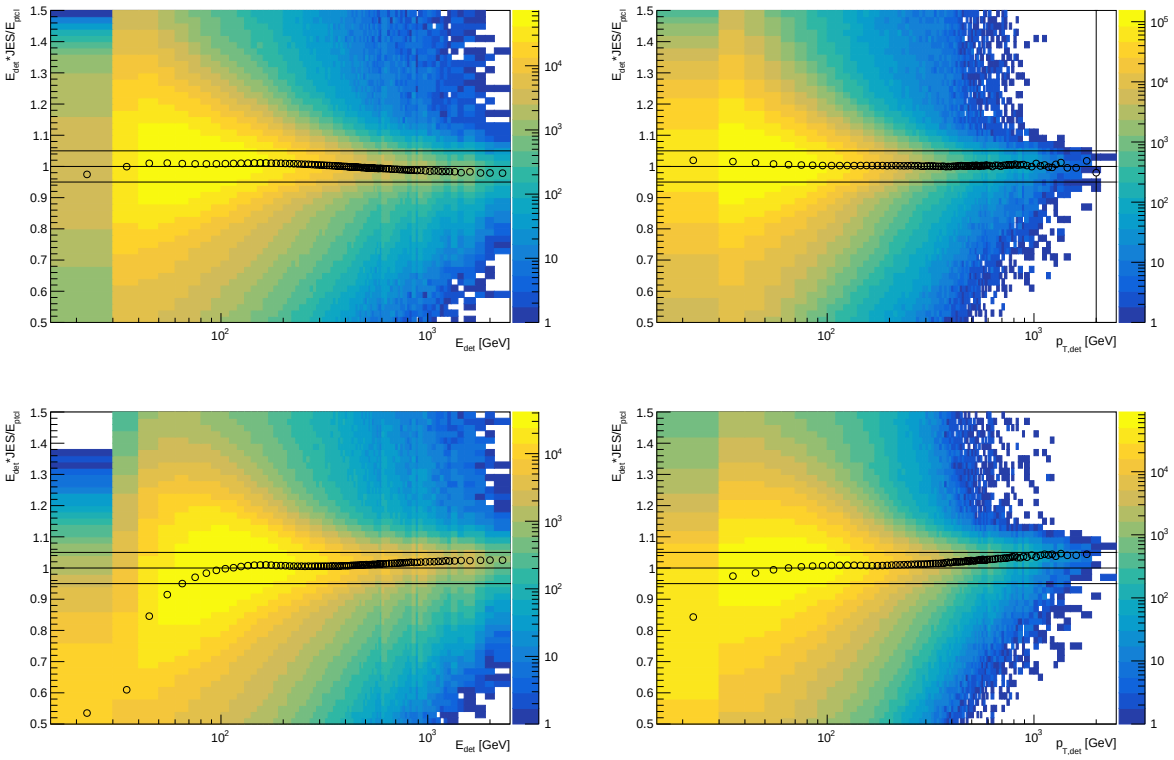


Figure 4.4: Closure test, *i. e.* the ratio of energy of JES corrected jet at the detector level over energy of matched jet at the particle level, by private JES factors in dependence on the energy (left) and the transverse momentum (right) of the small jets (top row) and large jets (bottom row) at the detector level, tested on the $t\bar{t}$ sample. The horizontal lines denote the 5% difference between the corrected detector level jet and the generated particle level matched jet.

struction level and for the boosted, semi-boosted, semi-boosted mixed and resolved topologies and are described below.

4.3.1 Missing Transverse Energy and Lepton Selection

The missing transverse energy ($E_{\mathrm{T,miss}}$) used in the private analysis is defined in same manner as in equation 3.7. The magnitude of the missing transverse energy is required to be $E_{\mathrm{T,miss}} > 25$ GeV for all topologies as well as for both the detector and the particle levels. This ensures that only events in which neutrinos carry away a considerable amount of energy from the decay are chosen for the analysis. This is a standard requirement for the missing energy in most of top quark analyses in channels including leptons.

The second condition in the selection order is the requirement on the lepton (muon or electron) transverse momentum, namely $p_\ell > 25$ GeV. Similar value is also commonly used in top quark analyses. Tau leptons are not directly considered in the simulation as they decay before they enter the detector, but their leptonic decay products (ℓ, μ) are. In case there are more electrons or muons fulfilling this requirement, only the electron or muon with the highest transverse momentum is taken into account. These requirements are the same for all topologies. Events are separated to an electron or a muon channel by the flavor of the selected lepton. The lepton may radiate low energy photons which are highly collimated. The separation of those radiated photons and the lepton is below the resolution of the detector and the photon energy is added by construction at the detector level. The lepton dressing procedure is performed at the particle level reconstruction to correct for this phenomenon. Photon four-vectors, fulfilling the condition of the angular separation threshold $\Delta R_{\gamma,\ell} = \sqrt{\Delta\eta^2_{\gamma,\ell} + \Delta\phi^2_{\gamma,\ell}} < 0.1$, are added to the lepton four-vector.

4.3.2 Large Jet Selection

The large jet four-vector is the result of the reconstruction with the anti-k_t algorithm with a distance parameter $R = 1$. A private jet energy scale correction as described in section 4.2 is applied to the detector level large jet before the selection, derived on the $t\bar{t}$ sample. The magnitude of the jet energy scale correction is about 5% depending on η and p_{T}. The transverse momentum of large jets is required to be $p_{\mathrm{T,LJ}} > 100$ GeV. This condition helps to reduce the number of events with jets not coming from top quark decays. Furthermore, all large jets are considered in the pseudorapidity range $|\eta| < 2.5$. This constraint ensures in practice better jet identification as the forward region of the detector is not instrumented for tracking and has a worse energy resolution. The isolation criterion of jets from the lepton ensures that the selected lepton is not contained within the large jet by following the requirement of $\Delta R_{\mathrm{LJ},\ell} = \sqrt{\Delta\eta^2_{\mathrm{LJ},\ell} + \Delta\phi^2_{\mathrm{LJ},\ell}} > 1$. All these requirements are applied to all three topologies[1] and both the detector and the particle levels. Each large jet is then probed for the top quark and W boson tagging, first for the hypothesis as coming from the top quark decay, then, in the semi-boosted topology, as coming from the W boson decay and in case none of the tagging was successful, the event is then considered as a candidate for the semi-boosted mixed or the resolved topology.

[1] There is no large jet in the resolved topology.

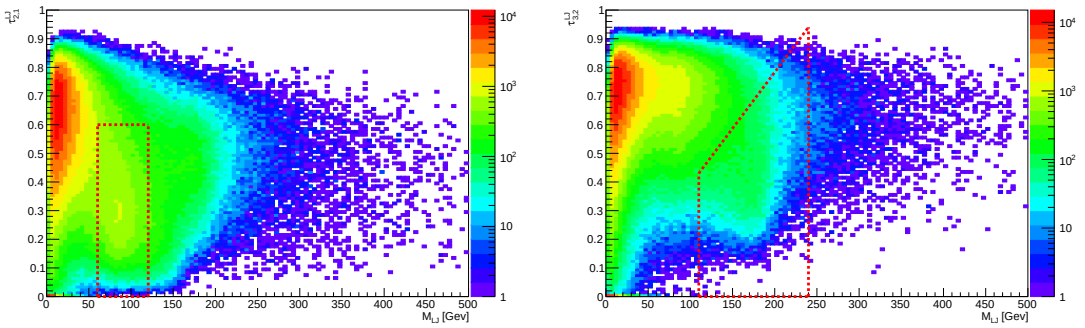


Figure 4.5: Distribution of the jet substructure variable $\tau_{2,1}$ (left) and $\tau_{3,2}$ (right) in the dependence on the large jet mass (M_{LJ}) for the sample with $M_{\mathrm{y_0}} = 700$ GeV at the detector level. The large jets in the red dotted area are selected for the reconstruction of the boosted W boson and top quark. The mass window is set around the expected mass values equal to the W boson and top quark mass, respectively.

Tagging for the boosted topology is based on the constraint on the mass of the large jet 110 GeV$< M_{\mathrm{LJ}} <$ 240 GeV and a constraint combining the large jet mass and a jet substructure variable $\tau_{3,2}$ [101], describing the possibility of finding three small jets inside the studied large jets rather than two small jets, as $M_{\mathrm{LJ}}/\tau_{3,2} > 256$ GeV. The value for the second constraint was added to avoid background jets, *e.g.* a large jet from the W boson. The selection is depicted in Fig. 4.5 (right) by the area inside the red dotted lines.

Tagging of large jets for the semi-boosted topology is based on the large jet mass window 60 GeV$< M_{\mathrm{LJ}} <$120 GeV and the jet substructure variable $\tau_{2,1}$, describing the possibility of finding two small jets inside the studied large jets rather than one small jet, $\tau_{2,1} < 0.6$. The selection is shown in Fig. 4.5 (left).

In order to evaluate the tagging efficiencies, a comparison to the generator level information is needed to define a truth jet label as top, W or light otherwise. The angularly closest large-R jet to the direction of the original top quark was used as a probe for the truth top tagging efficiency $\varepsilon_{\mathrm{top}}$ which is defined as follows

$$\varepsilon_{\mathrm{top}} = \frac{N_{\mathrm{jet,top}}^{\mathrm{match\&tag}}}{N_{\mathrm{jet,top}}^{\mathrm{match}}}\,, \tag{4.3}$$

where $N_{\mathrm{jet,top}}^{\mathrm{match\&tag}}$ is the number of jets matched to the original top quark and top-tagged by the tagging technique; and $N_{\mathrm{jet,top}}^{\mathrm{match}}$ is number of all jets originating from the top quark. The truth W boson tagging efficiency is defined in a similar manner.

The mistag (fake) efficiency was also evaluated, which describes the false positivity of the tagger on jets not originating from the top quark or the W boson. It is defined by the following formula in case of the top tagging

$$\varepsilon_{\mathrm{mis,top}} = \frac{N_{\mathrm{jet,top}}^{\neg\mathrm{match\&tag}}}{N_{\mathrm{jet,top}}^{\neg\mathrm{match}}}\,, \tag{4.4}$$

where $\varepsilon_{\mathrm{mis,top}}$ is the mistag efficiency of the top quark tagger, $N_{\mathrm{jet,top}}^{\neg\mathrm{match\&tag}}$ is the number of large jets which are not matched to the generator level top quark but are top-tagged by the tagger; and $N_{\mathrm{jet,top}}^{\neg\mathrm{match}}$ is the number of all large jets not matched to the generator

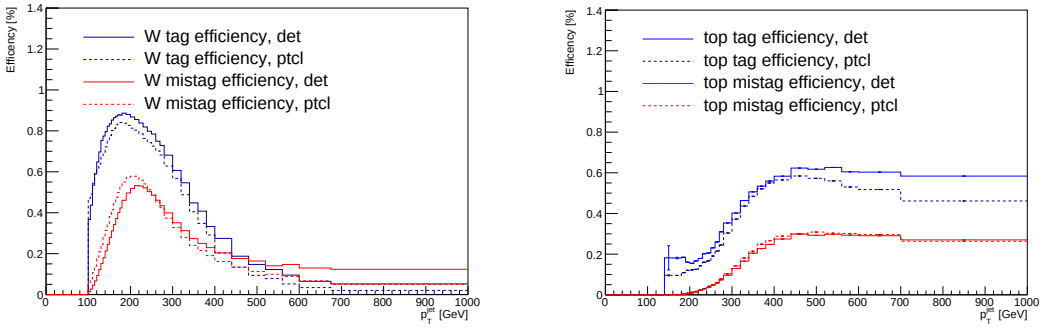


Figure 4.6: Tagging and mistag efficiencies for the W boson (left) and for the top quark (right) in dependence on the transverse momentum of the large jet studied on the $t\bar{t}$ sample at both particle and detector levels.

level top quark. The W mistag efficiency is calculated in a similar way for the both detector and particle levels.

Both tag and mistag efficiencies are shown in Fig 4.6 for the W (left) and top (right) tagger, studied at both detector and particle levels in dependence on the transverse momentum of the large jet. The sample for the mistag efficiency was the production of $2j2b$ events and contained 32.5M events, with events generated in exclusive p_{T} ranges of the leading and sub-leading jets in order to populate the phase space of higher transverse momenta.

The case of the semi-boosted mixed topology is also considered for tagging, but unlike in the semi-boosted topology, where the large jet can be tagged as coming from the W boson, there is no expected peak in the large jet mass spectrum in the semi-boosted mixed topology, and it cannot be tagged based on its mass. Thus, there is another requirement for the large jet to contain a small jet originating from the b quark decay by requiring the condition $\Delta R_{\mathrm{LJ,SJ}} = \sqrt{\Delta\eta^2_{\mathrm{LJ,SJ}} + \Delta\phi^2_{\mathrm{LJ,SJ}}} < 0.5$ at the detector level, see Section 4.3.3 for further information about small jets. A similar condition is set at the particle level for the large jet to contain a B-hadron within $\Delta R_{\mathrm{LJ,B-had}} < 0.5$.

4.3.3 Small Jet Selection

The general requirements for the small jet selection are on the transverse momentum $p_T > 25$ GeV and the isolation from the selected lepton $\Delta R_{\mathrm{SJ},\ell} > 0.5$. The private jet energy scale correction is applied to detector level small jets before the selection, which was derived on the $t\bar{t}$ sample on top of the DELPHES default jet energy scale. The magnitude of this residual jet energy scale correction is about 2%. The b-tagging of the small jets is done by the DELPHES simulation at the detector level and by the requirement of containing a B-hadron at the particle level by fulfilling the requirement of $\Delta R_{\mathrm{SJ,B-had}} < 0.4$.

The small jet for the reconstruction of the leptonically decaying top quark has to fulfill the angular condition $\Delta R_{\mathrm{SJ},\ell} < 2$, which ensures that it lies in the vicinity of the selected lepton, and must be b-tagged. The large jet isolation condition $\Delta R_{\mathrm{SJ,LJ}} > 1.5$ applies to all topologies with the exception of the resolved topology, where there is no large jet. Such a selected small jet is then removed from the jet collection and

from further consideration. This is the only selected small jet in case of the boosted topology.

For the reconstruction of the hadronically decaying top quark in the semi-boosted topology a small b-tagged jet is required in the vicinity to the selected large jet $1 < \Delta R_{\mathrm{SJ,LJ}} < 1.5$. Thus a partial overlap between the selected large jet and the considered small jet is allowed, *i.e.* the selected b-tagged small jet can be partially contained in the selected large jet.

The conditions for the semi-boosted mixed topology are similar to the conditions for the semi-boosted topology. The vicinity condition to the large jet remains unchanged but the small jet is required not to be b-tagged. This condition together with the b-tag of the selected large jet ensures the correct selection for this topology.

The resolved topology selection is tried as the last option before the event is discarded. The reconstruction of the hadronically decaying top quark in the resolved topology requires three small jets, one of them b-tagged. The algorithm first takes two small non-b-tagged jets with the highest transverse momentum and tests their invariant mass $M_{\mathrm{SJ,SJ}} < 120$ GeV to avoid dijets not corresponding to the mass of the W boson. Then it adds the four-vector of the remaining b-tagged jet[2]. If such three jets are found, the event is accepted.

4.4 Top Quark Pair Reconstruction

The top quark pair reconstruction is independent on the event on the complementary reconstruction level, thus the different reconstructed topology event on the detector and the particle level is not forbidden. Another parallelization in reconstruction is in the lepton selection, where the event with muon and electron is reconstructed independently if both of them fulfill the selection criteria. Parallelization in the lepton selection and independence on the complementary event is used to increase the number of reconstructed events and efficiency of the algorithm. The topology matching between the detector and particle level is only performed for the spectra undergoing the unfolding procedure.

4.4.1 Reconstruction of Leptonically Decaying Top Quark

Reconstruction of leptonically decaying part of event is same for all topologies. First, the condition on missing energy is tested, then the presence of the selected electron or muon is verified. The object reconstructed from four-momenta of the chosen lepton and missing energy, considered as the energy of undetected neutrino, is corresponding with W boson, but due to uncertainty coming from not knowing the longitudinal component of neutrino momentum there are up to two possible values of the W boson longitudinal momentum component. The ambiguity is caused by the calculation from condition $M_{\ell,\nu} = M_{\mathrm{W}}$, where$M_{\ell,\nu}$ is combined invariant mass of the reconstructed W boson from lepton and missing energy, which leads to the quadratic equation. The more central value of the longitudinal component of reconstructed neutrino momentum is usually selected in ATLAS collaboration analyses. The last object needed for the leptonically decaying top quark reconstruction is the selected b-tagged small jet. The leptonically decaying top quark four-momentum is then obtained as the combination of four-momenta of the reconstructed W boson object and the selected

[2]One b-tagged jet is used in the reconstruction of the leptonically decaying top quark.

small jet. In case that the algorithm was able to reconstruct the leptonically decaying top quark with selected electron and muon simultaneously, the top quark candidate reconstructed from lepton with higher transverse momentum is selected. Reconstruction of the leptonically decaying top quark is identical for the detector and particle level, with exception of the small jet b-tagging.

4.4.2 Reconstruction of Hadronically Decaying Top Quark.

The reconstruction of the hadronically decaying top quark is different for each of the topologies. The algorithm first tries to find the boosted topology top quark pair candidate, in case of failure it attempts to find the semi-boosted topology pair and in case of subsequent failure it tries the semi-boosted mixed topology selection. The resolved topology top quark pair reconstruction takes place only if all previous topology conditions fail. The studied event is discarded only in case in which algorithm was unsuccessful in reconstruction of event in all topologies.

The reconstruction of the hadronically decaying top quark in the boosted topology is based on selection of top tagged large jet, as mentioned in previous section, which is considered as the top quark candidate, because it should contain most of the decay products corresponding to the original particle.

The reconstruction of the hadronically decaying top quark in the semi-boosted topology is similar but the selected tagged large jet corresponds to the W boson instead to the top quark as in the previous case. Thus, the small jet tagged by the reconstruction algorithm as containing products of B-hadron decay is needed to complete the four-momentum of the reconstructed top quark.

The reconstruction of the hadronically decaying top quark in the semi-boosted mixed topology is performed by summing one large b-tagged jet and one non-b-tagged small jet four-vectors. The invariant mass of such a large jet does not produce a peak corresponding to any particle but in the combination with the selected small jet the resulting invariant mass peak should correspond to the one of the top quark as shown in Fig. 5.16.

The reconstructed hadronically decaying top quark four-momentum in the resolved topology is the sum of four-momenta of three selected small jets as described in section 4.3.3. As there is no selected large jet but only small jets, this topology is the most prone one to the miscombination of jets, which is one of the reasons of slightly worse resolution in the reconstruction of the top quark pair invariant mass.

Finally, the reconstructed top anti-top quark pair four-momentum is obtained as a combination of four-momenta of the reconstructed hadronically and leptonically decaying top quark.

Chapter 5

Results

Results from selection and reconstruction are presented in this chapter following step-by-step the aforementioned procedures. This involves basic characteristics of the signal samples, control plots of variables of different particles used in the reconstruction and comparison between the two signal samples with $M_{Z'} = 700$ GeV and $M_{y_0} = 700$ GeV. The signal sample with the hypothetical particle y_0 with decay width of 10% of its invariant mass is chosen for the results presentation.

5.1 Shape Comparison

Several basic plots and spectra are presented in this subsection ranging from the multiplicity of the jets to the overview of spectra for different variables in the reconstruction process. Multiplicities of all small jets in the generated signal samples with $M_{Z'} = 700$ GeV and $M_{y_0} = 700$ GeV are shown in Fig. 5.1 (left) while the multiplicity of the large jets is shown in Fig. 5.1 (right) for both detector and particle levels. Similarly, the multiplicity of the b-tagged small jets and large jets is presented in Fig. 5.2 respectively. The jet multiplicity and b-tagged jet multiplicity shows similar trend for small and large jets, where the multiplicity at the particle level is larger than multiplicity at the detector level. Different topologies need different number of small and large jets, some of them b-tagged. Multiplicity ranges from two up to ten small jets in signal samples and confirms that there are enough jets for the reconstruction. The situation with large jets is different as there is smaller amount of large jets in the event, but only one chosen large jet is needed for the event reconstruction if any. Overall, jet multiplicity and the multiplicity of b-tagged jets is similar comparing the y_0 and the Z' signal samples.

5.2 Spectra Shape Comparison

The shape spectra comparison is performed for both detector and particle level for the signal samples with $M_{Z'} = 700$ GeV and $M_{y_0} = 700$ GeV before and after the basic selection for all spectra in this section.

Small jets or small-R jets are crucial for reconstruction in all studied cases. There are basic conditions same for all topologies as mentioned in previous section, for example cut on the transverse momentum. Spectra of the transverse momentum of small jets is shown in Fig. 5.3. The amount of small jet with higher transverse momentum spectrum drops, which is expected and corresponds to the standard model theory.

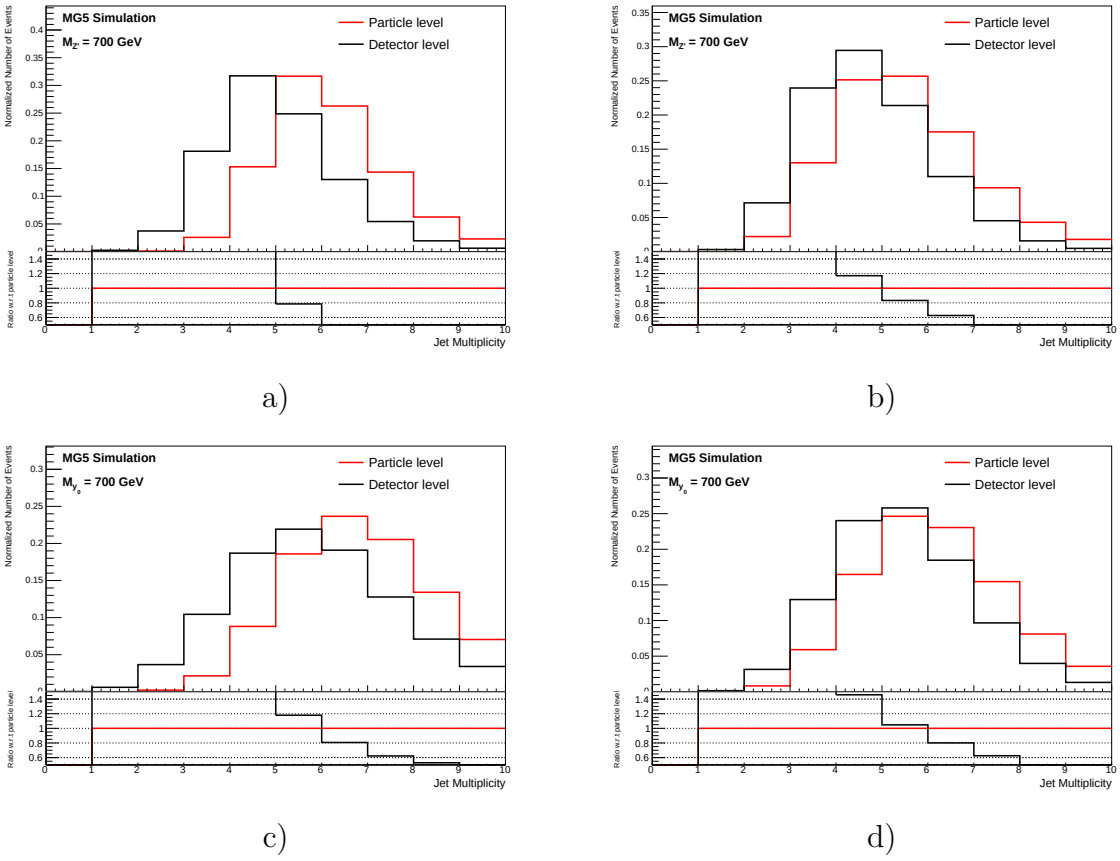


Figure 5.1: Normalized distribution of the number of small jets in the event (a), c)) and number of large jets in the event (b), d)) for both detector and particle levels for the sample with $M_{Z'} = 700$ GeV (top row) and for the sample with $M_{y_0} = 700$ GeV (bottom row).

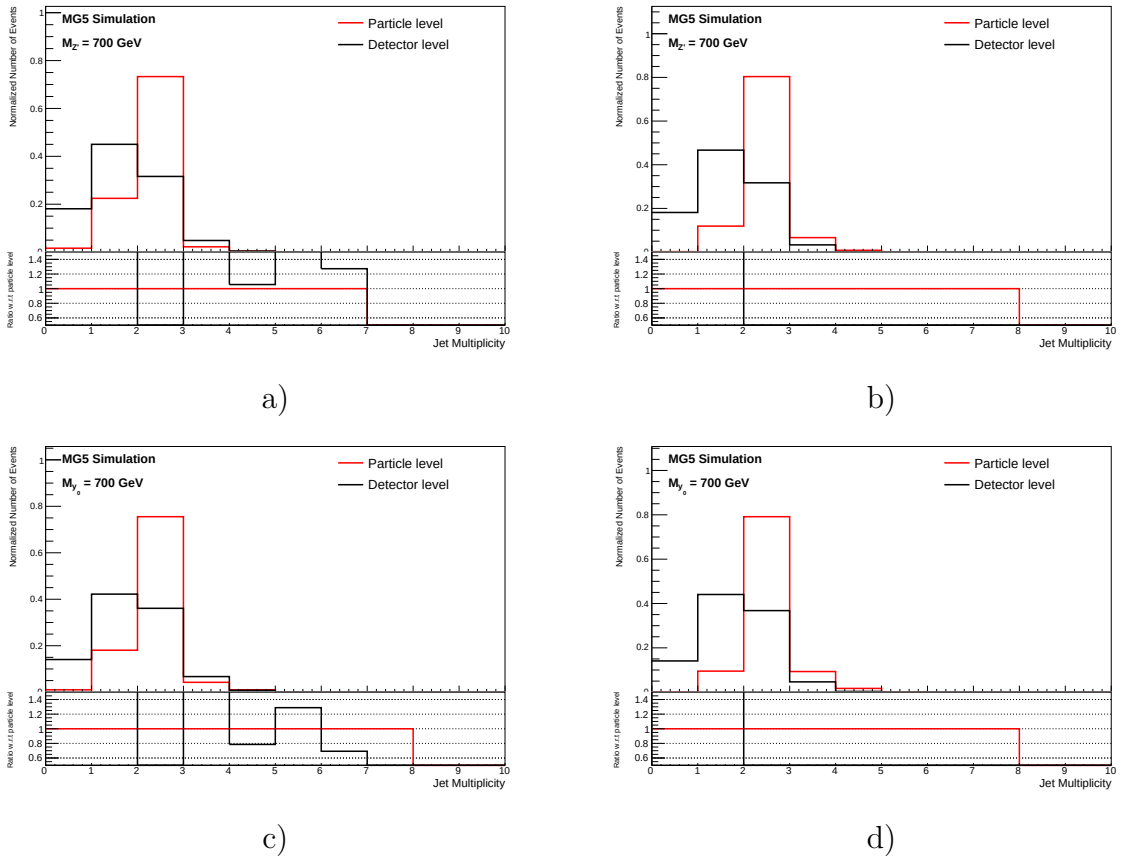


Figure 5.2: Normalized distribution of the number of b-tagged small jets in the event (a), c)) and number of b-tagged large jets in the event (b), d)) for both detector and particle levels for the sample with $M_{Z'} = 700$ GeV (top row) and for the sample with $M_{y_0} = 700$ GeV (bottom row).

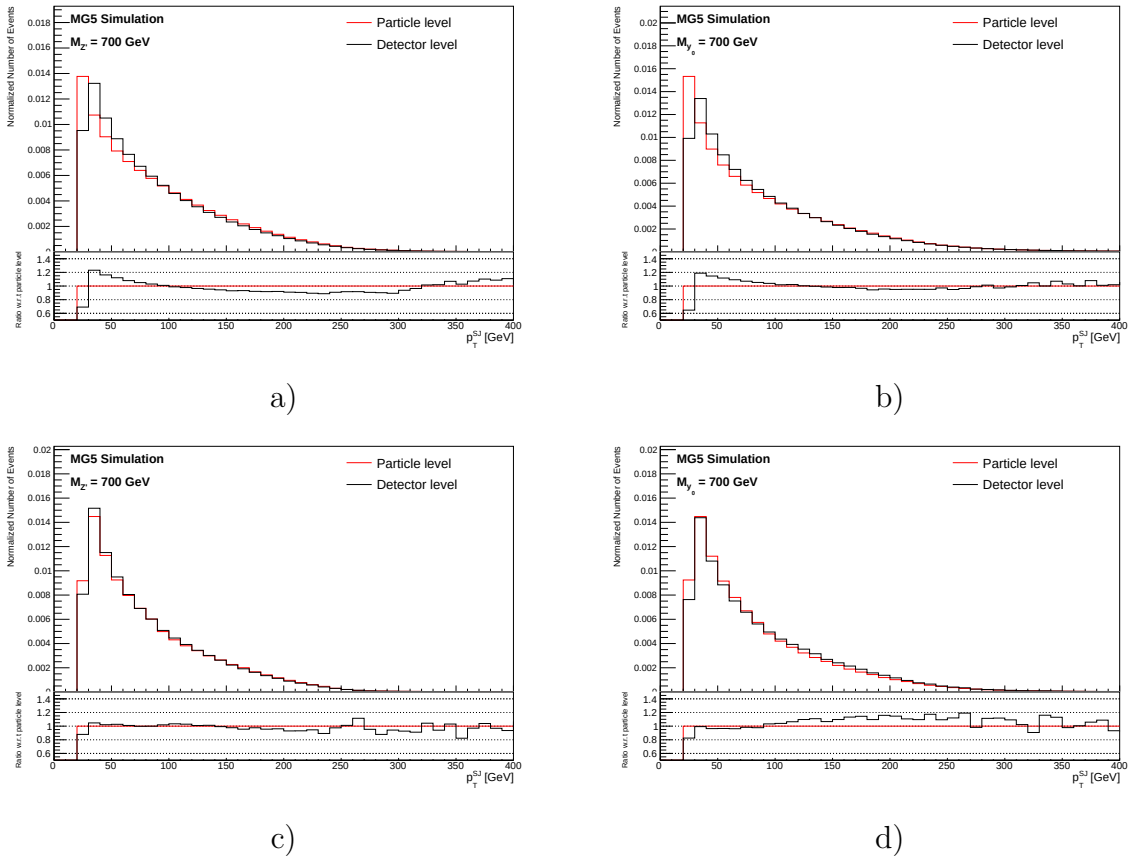


Figure 5.3: Normalized spectrum of the transverse momentum of all small jets in all events (p_{T}^{SJ}) before (a), b)) and after (c), d)) the application of the selection conditions for both detector and particle levels for the sample with $M_{\mathrm{Z}'} = 700$ GeV (a), c)) and for the sample with $M_{\mathrm{y_0}} = 700$ GeV (b), d)). Lower pads show the ratio between detector and particle level.

The combined spectrum of small jet transverse momenta from all samples is shown in Fig 5.4 all spectra are weighted to the luminosity of the $t\bar{t}$ sample ($L_{\mathrm{t\bar{t}}} = 16.44 fm^{-1}$) for the comparison, see 4.1 for the details. The black dotted spectrum denoted as pseudodata corresponds to the same combined and weighted spectrum from the statistically independent sample. The bottom part shows the ratio between the combined spectrum and pseudodata and ratio of combined sample with contribution from the hypothetical particle over the combined spectrum without the signal contribution, thus showing the strength of the signal.

Large jets or large-R jets are needed for the event reconstruction and the basic selection conditions are on the transverse momentum and the centrality of the large jets. The large jet transverse momentum spectrum is shown in Fig. 5.5. The distribution of the large jets with higher energies is similar to the case of small jets with exception of plateau in higher transverse momentum, which is more visible for the Z' sample. The reason behind is that large jets cover larger area than small jets and may contain one or more small jets contributing to the transverse momentum.

The combined spectrum of large jet transverse momenta from all samples is shown in Fig 5.6 all spectra are weighted to the luminosity of the $t\bar{t}$ sample see 4.1 for details. The black dotted spectrum denoted as pseudodata corresponds to the same combined

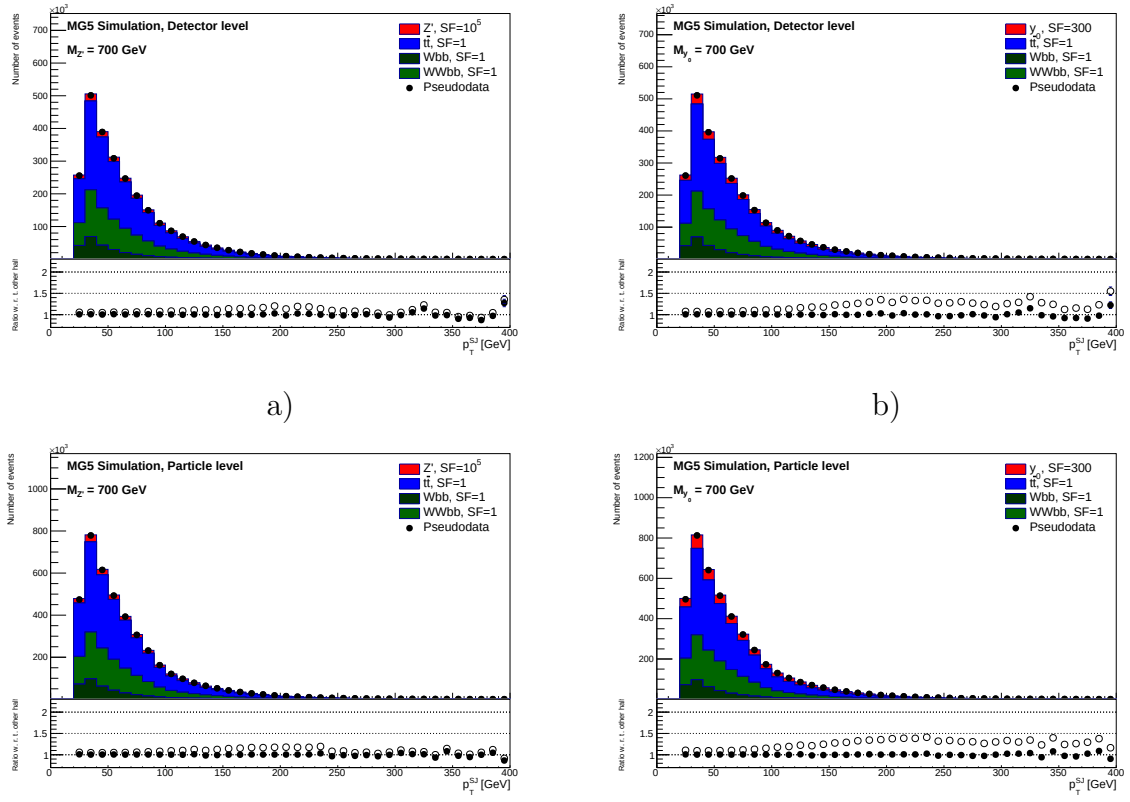


Figure 5.4: Combined spectrum over all samples of the transverse momentum of all small jets p_{T}^{SJ} at the detector (top row) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{\mathrm{Z}'} = 700$ GeV (a), c)) and with $M_{\mathrm{y}_0} = 700$ GeV (b), d)). All samples are weighted to the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

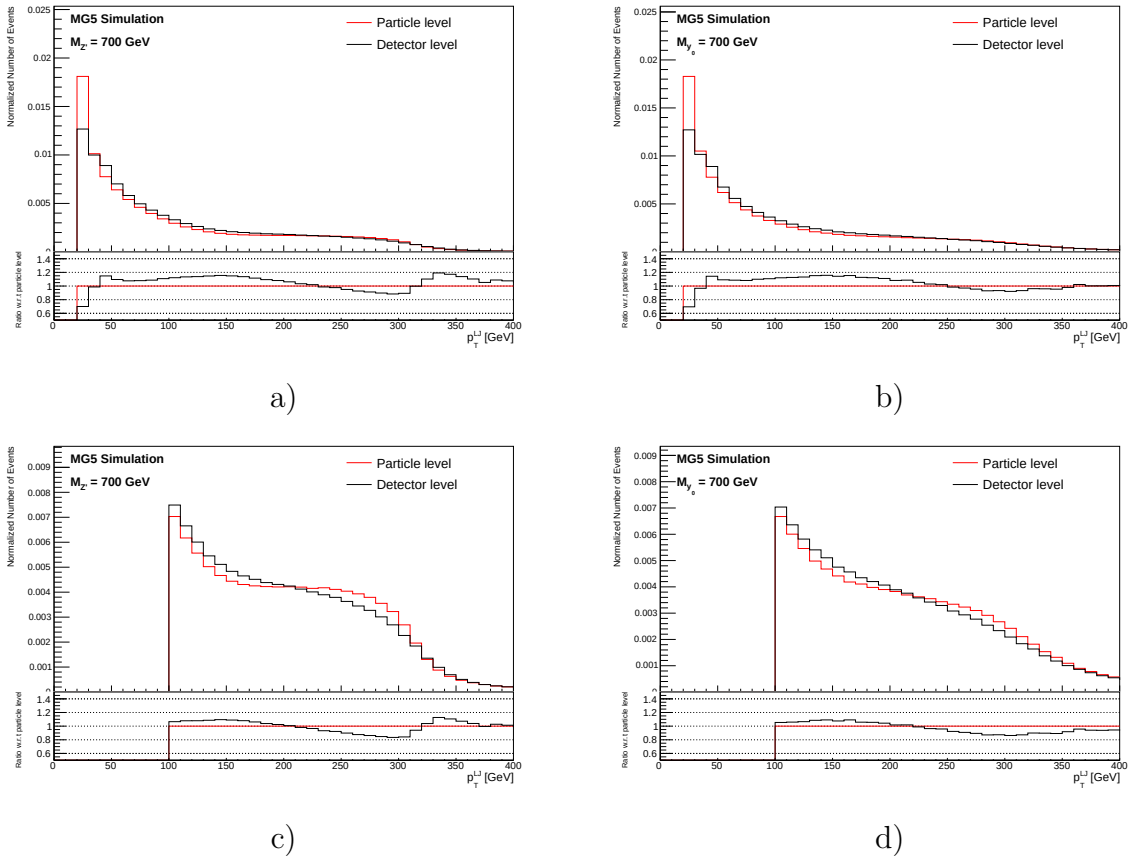


Figure 5.5: Normalized spectrum of transverse momentum of large jets in all events (p_{T}^{LJ}) before (a), b)) and after (c), d)) the application of the selection conditions for both detector and particle levels for the sample with $M_{Z'} = 700$ GeV (a), c)) and for the sample with $M_{\mathrm{y}_0} = 700$ GeV (b), d)). Lower pads show the ratio between detector and particle level.

and weighted spectrum from the statistically independent sample. The bottom part shows the ratio between the combined spectrum and pseudodata and the strength of the signal.

Electron or muon used in the event reconstruction of the leptonic side of the event is chosen based on the transverse momentum condition as specified in the previous chapter. The transverse momentum of all electrons in the event is shown in Fig. 5.7 and in Fig. 5.9 for muons. The shape of the electron and muon transverse momentum spectra in both signal samples is similar. The main difference between particle and detector level is in the number of leptons at the beginning of the spectra before the application of selection, where there is peak at the particle level, which is unseen on the detector level. This corresponds to the detection threshold of the detector simulation for lower transverse momentum, which is well modeled in the ATLAS simulation by the DELPHES package. The combined spectrum of lepton transverse momenta from all samples is shown in Fig 5.8 and in Fig 5.10, weighted to the luminosity of the $t\bar{t}$ sample. The ratio between the combined spectrum and pseudodata and the strength of the signal is shown in the bottom part of the plot.

Missing transverse energy is used in kinematic reconstruction of the escaping neutrino from the leptonic decay of the W boson. Events with low missing transverse

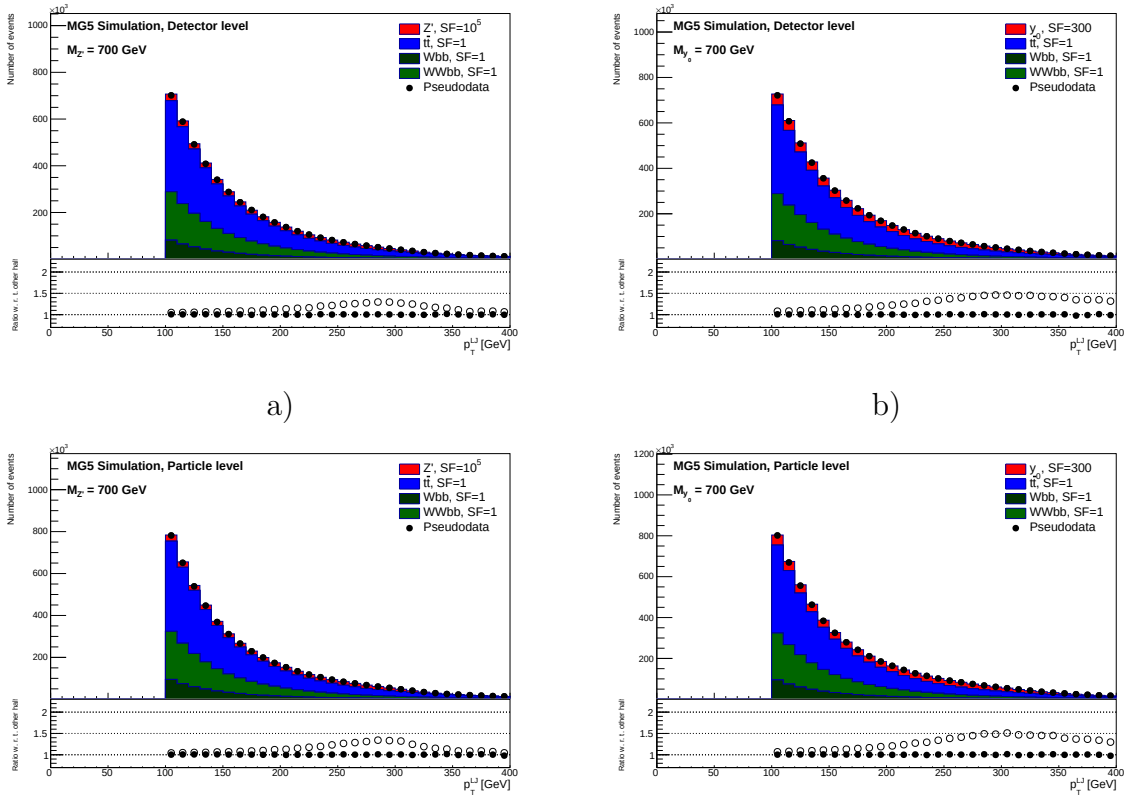


Figure 5.6: Combined spectrum over all samples of the transverse momentum of all large jets p_{T}^{LJ} at the detector (top row) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{\mathrm{Z}'} = 700$ GeV (a), c)) and with $M_{\mathrm{y}_0} = 700$ GeV (b), d)). All samples are weighted to the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

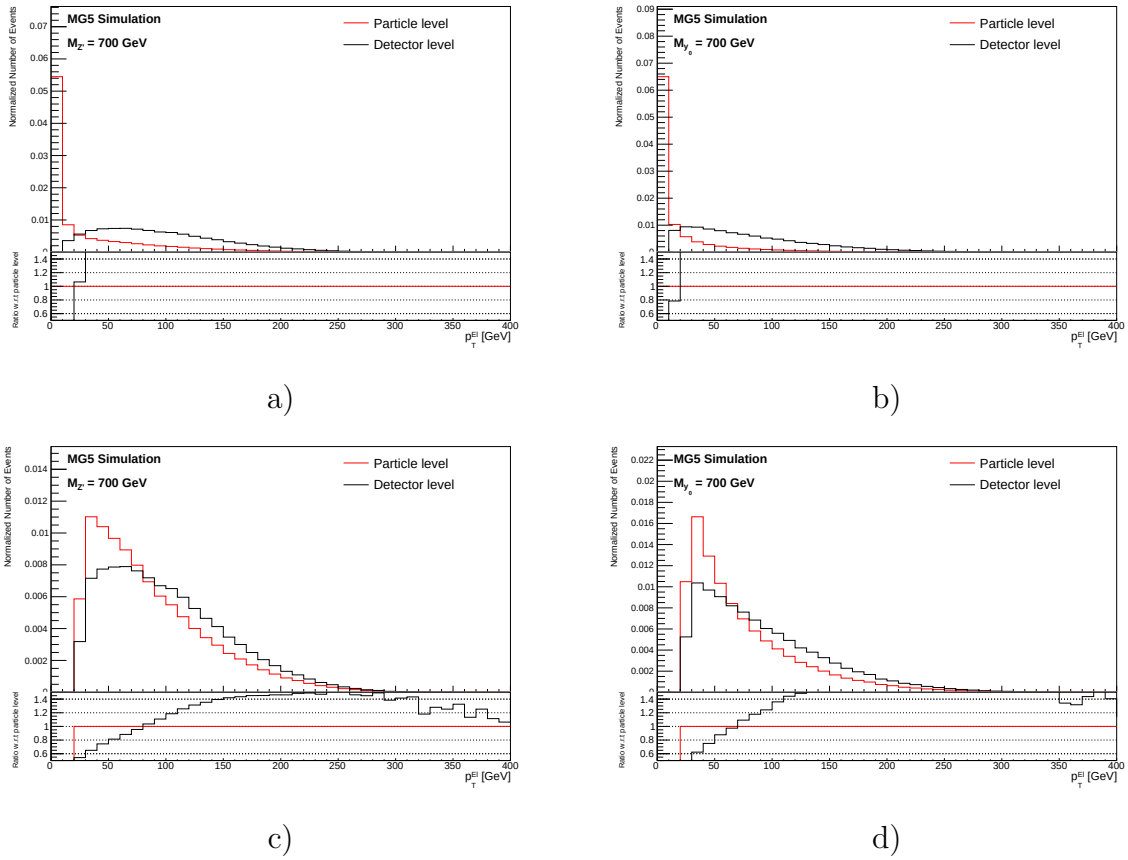


Figure 5.7: Normalized spectrum of transverse momentum of electrons in all events (p_{T}^{El}) before (a), b)) and after (c), d)) the application of the selection conditions for both detector and particle levels for the sample with $M_{Z'} = 700$ GeV (a), c)) and for the sample with $M_{\mathrm{y}_0} = 700$ GeV (b), d)). Lower pads show the ratio between detector and particle level.

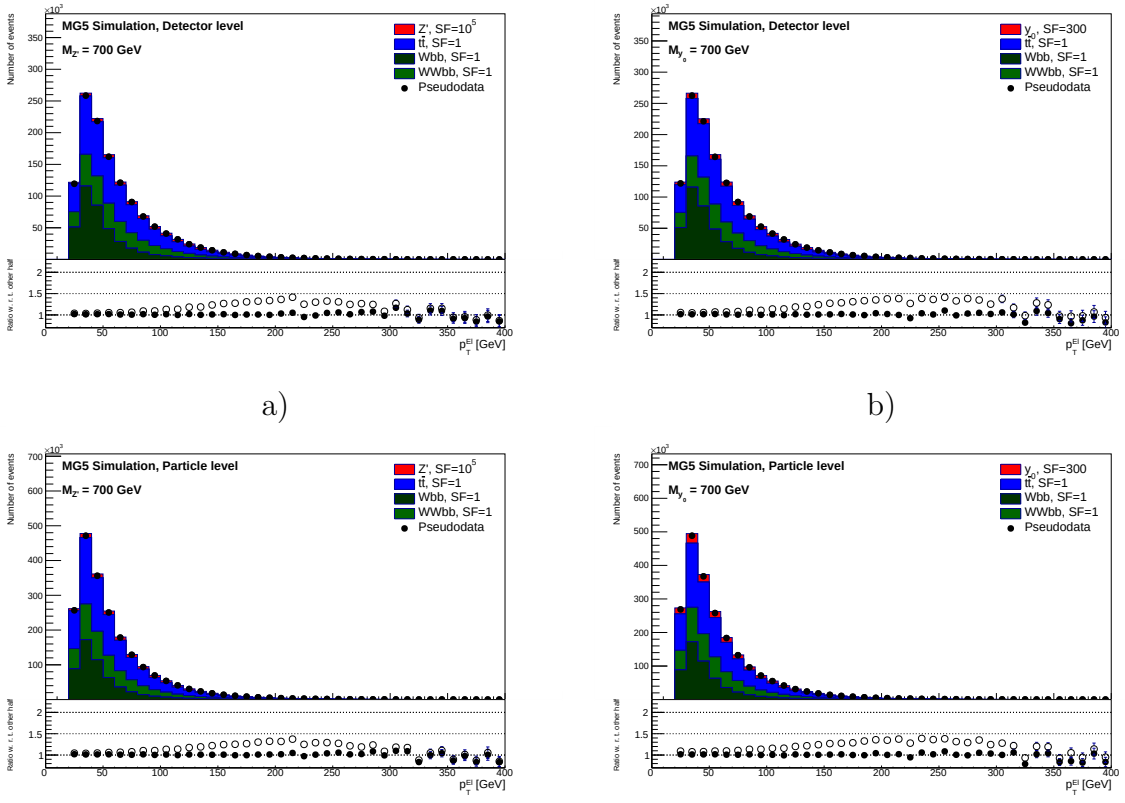
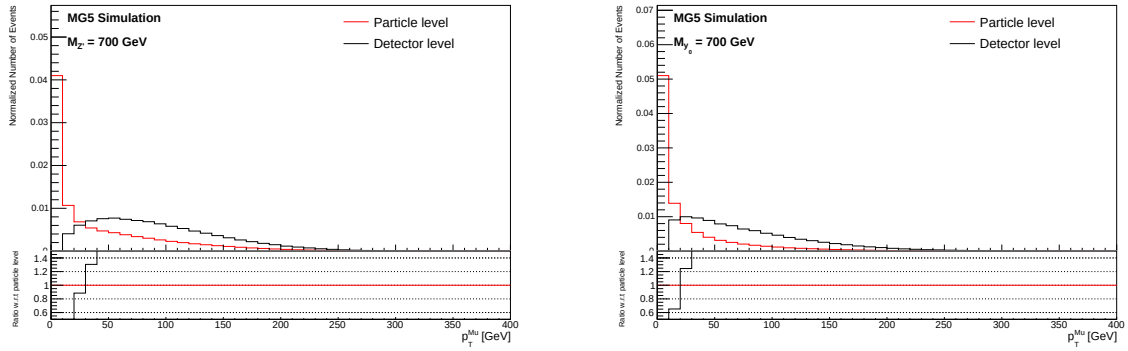
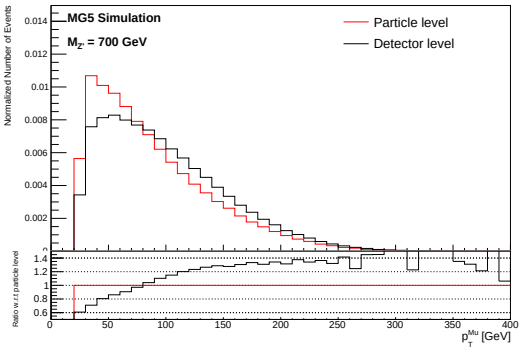


Figure 5.8: Combined spectrum over all samples of the electron transverse momentum p_{T}^{El} at the detector (top row) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{Z'} = 700$ GeV (a), c)) and with $M_{\mathrm{y}_0} = 700$ GeV (b), d)). All samples are weighted to the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

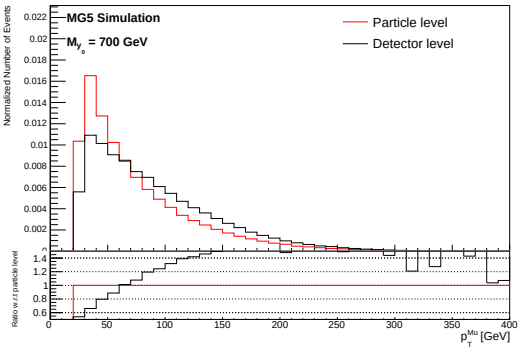


a)

b)



c)



d)

Figure 5.9: Normalized spectrum of muon transverse momentum in all events (p_{T}^{μ}) before (a), b)) and after (c), d)) the application of the selection conditions for both detector and particle levels for the sample with $M_{\mathrm{Z}'} = 700$ GeV (a), c)) and for the sample with $M_{\mathrm{y}_0} = 700$ GeV (b), d)). Lower pads show the ratio between detector and particle level.

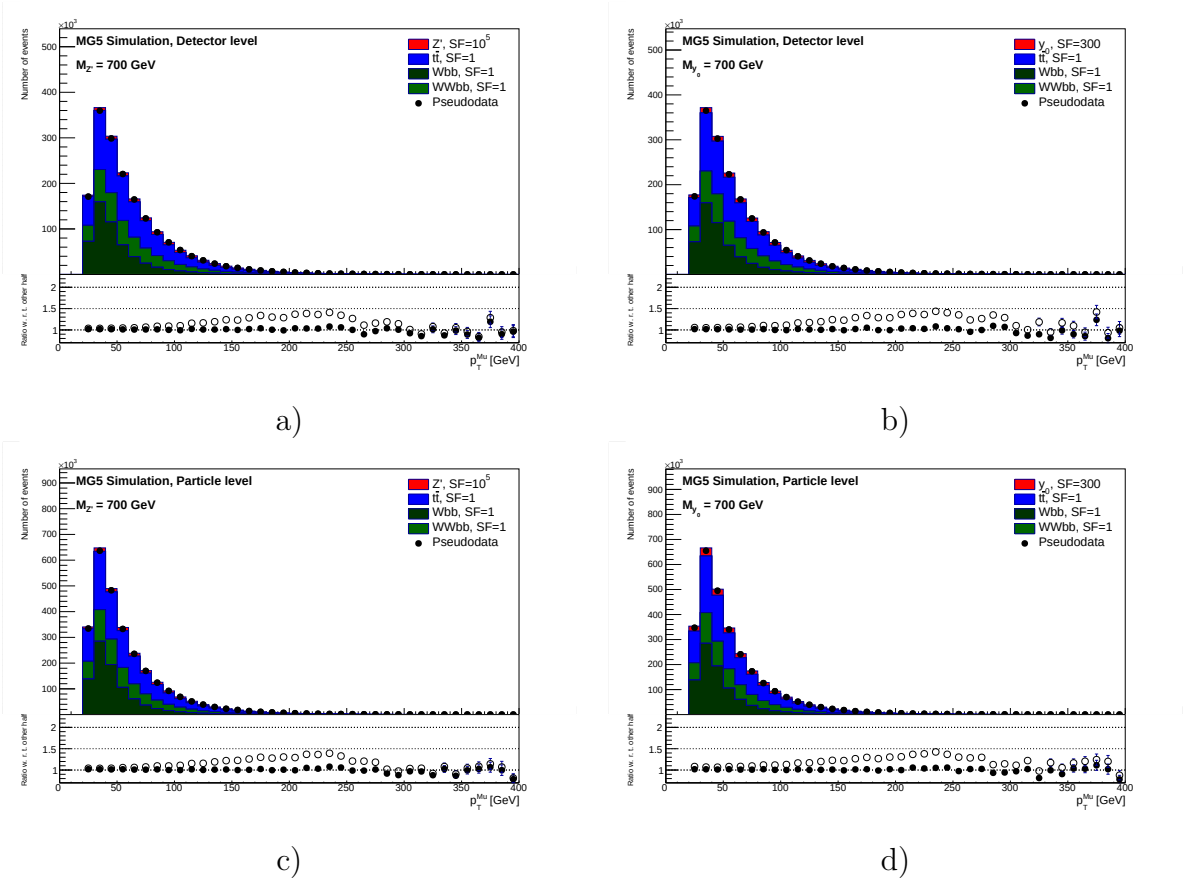


Figure 5.10: Combined spectrum over all samples of the muon transverse momentum p_{T}^{μ} at the detector (top row) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{\mathrm{Z}'} = 700$ GeV (a), c)) and with $M_{\mathrm{y_0}} = 700$ GeV (b), d)). All samples are weighted to the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

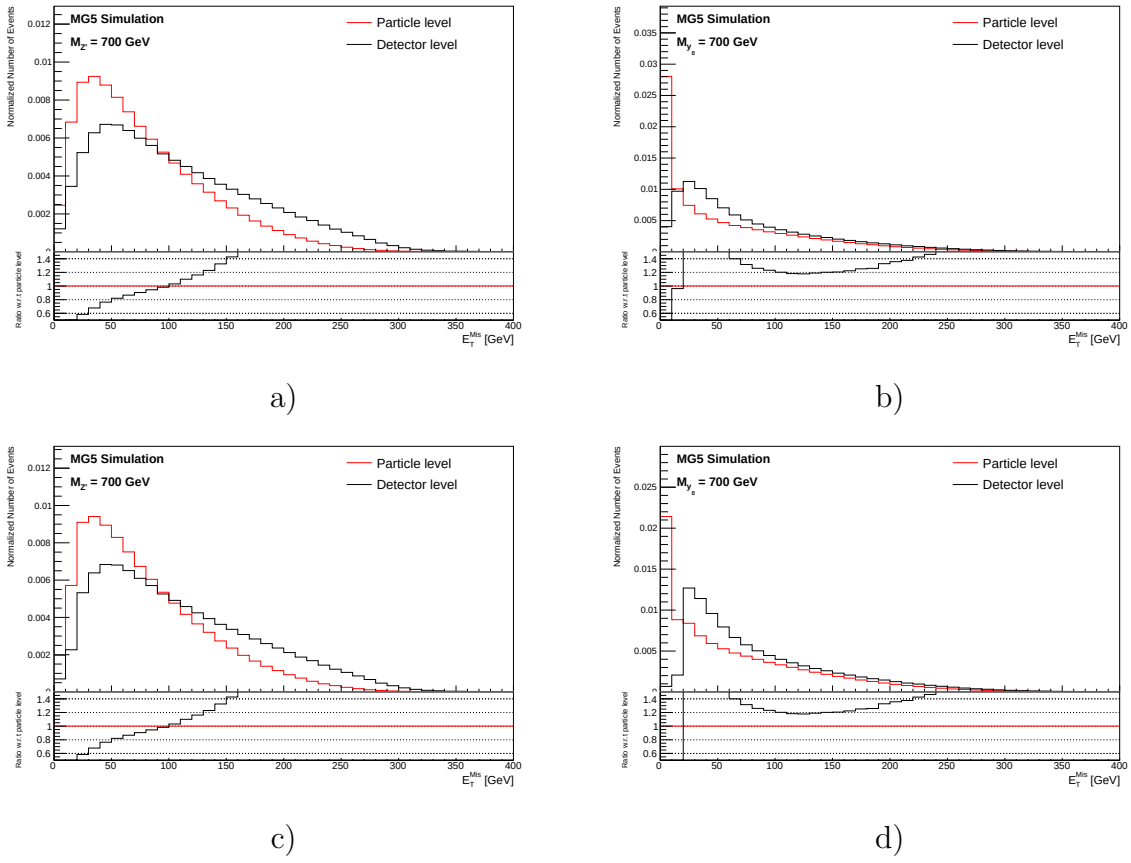


Figure 5.11: Normalized spectrum of missing transverse energy in all events ($E_{\mathrm{T}}^{\mathrm{Mis}}$) before (a), b)) and after (c), d)) the application of the selection conditions for both detector and particle levels for the sample with $M_{\mathrm{Z}'} = 700$ GeV (a), c)) and for the sample with $M_{\mathrm{y}_0} = 700$ GeV (b), d)). Lower pads show the ratio between detector and particle level.

energy are usually not interesting for the studied process, thus there is a condition on the missing transverse energy. The energy spectrum of this variable is shown in Fig. 5.11. The shape of this distribution is slightly different for the two signal samples are compared. The shape differences originating from the simulation of the process where Z' is simulated in the semi-leptonic channel exclusively while the y_0 simulation also contains events where both top and anti-top quarks decay hadronically or both decay leptonically. This phenomenon is unavoidable as the simulation of decay products of top quark pair generated through the triangle loop as it is not yet incorporated in the simulation package. Thus, the y_0 sample contains events which in general have lower missing transverse energy due to the non-existent neutrinos from leptonic top quark decay. The combined spectrum of missing transverse energy or momentum from all samples is shown in Fig 5.12 and in Fig 5.12, weighted to the luminosity of the $t\bar{t}$ sample. The ratio between the combined spectrum and pseudodata and the strength of the signal is shown in the bottom part of the plot. The threshold for the $E_{\mathrm{T}}^{\mathrm{Mis}} > 20$ GeV needs to be fulfilled at least for one reconstruction level, which enter the reconstruction, therefore there are events reconstructed only in one reconstructed level to enhance statistics in separate topologies.

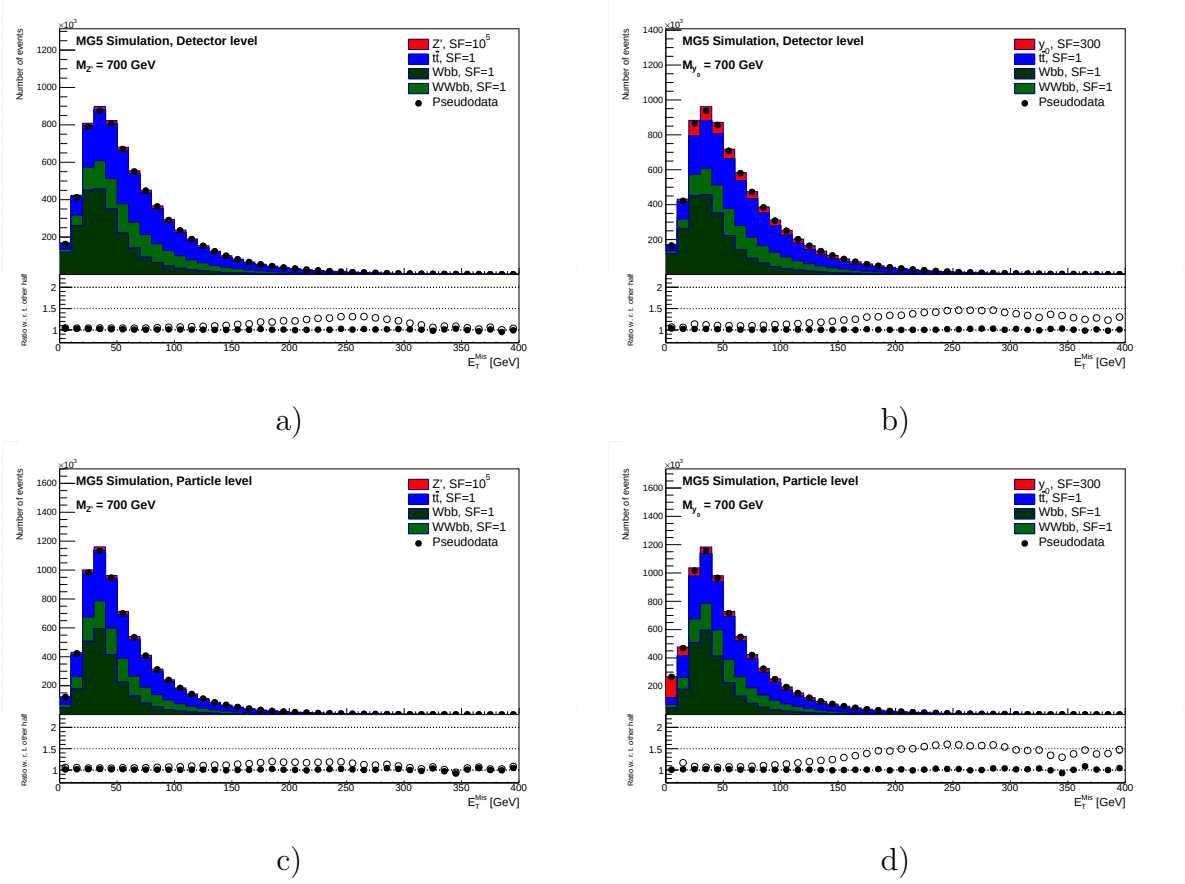


Figure 5.12: Combined spectrum over all samples of the missing transverse momentum E_{T}^{Mis} at the detector (top row) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{\mathrm{Z}'} = 700$ GeV (a), c)) and with $M_{\mathrm{y_0}} = 700$ GeV (b), d)). All samples are weighted to the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

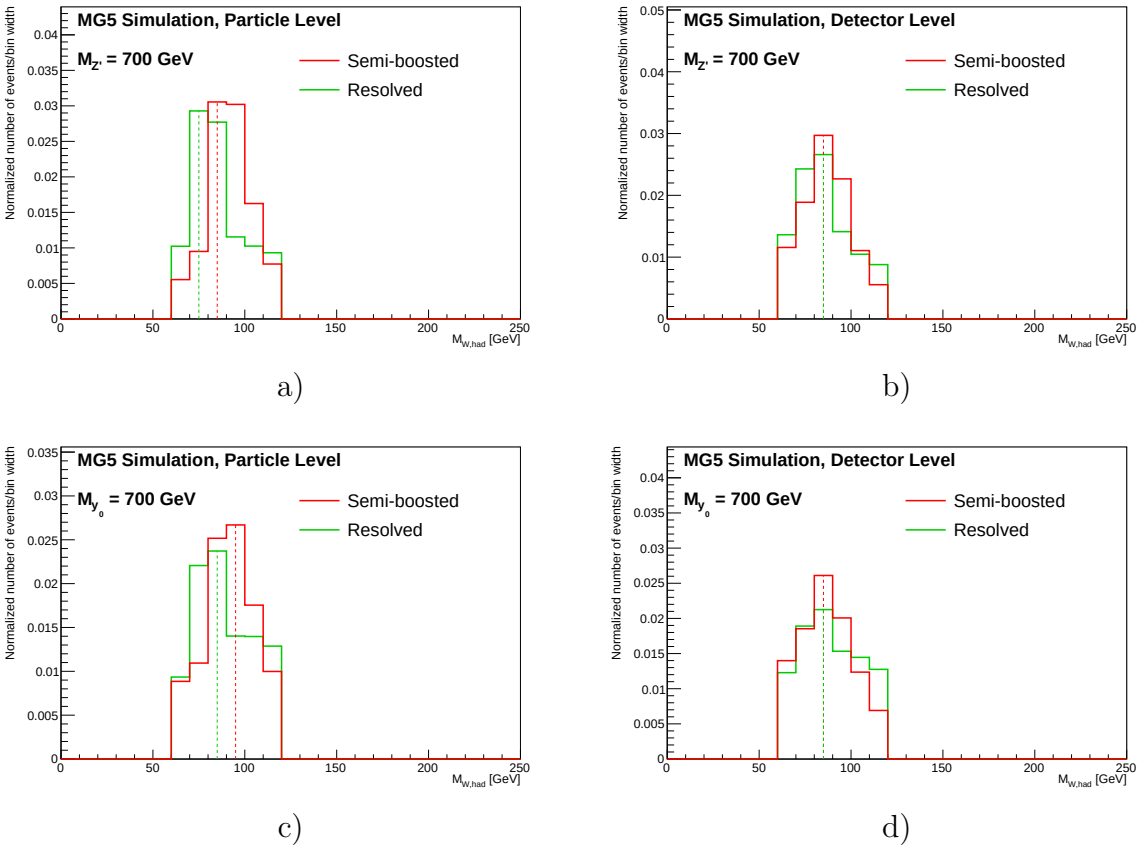


Figure 5.13: Comparison of the reconstructed hadronically decaying W boson invariant mass spectra ($M_{\mathrm{W,had}}$) between the particle (a), c)) and the detector (b), d)) level for both the semiboosted and the resolved topology using the sample with $M_{Z'} = 700$ GeV (a), b)) and $M_{\mathrm{y_0}} = 700$ (c), d)).

5.3 Oservables Shape Comparison

The shape comparison of chosen variables of reconstructed particles from signal samples with with $M_{Z'} = 700$ GeV and $M_{\mathrm{y_0}} = 700$ GeV is presented in this section for both detector and particle levels.

The reconstructed invariant mass of the hadronically decayed W boson is presented in Fig. 5.13, but only in the semi-boosted and the resolved topology is the hadronically decaying W boson reconstructed, in other cases there is no direct connection to the W boson. The position of peaks of the spectra highlighted by the dashed line roughly correspond to the theoretical invariant mass of the W boson.

The invariant mass of the reconstructed leptonically decaying top quark compared between the detector and the particle level is shown in Fig. 5.14. The invariant mass corresponds well with the theoretical invariant top quark mass value used for generation of the samples over all the topologies. The shape of the spectra is corresponding between topologies with exception for the resolved topology on the detector level, which is slightly different.

The combined spectrum of reconstructed leptonically decaying top quark mass from all samples is shown in Fig 5.15 in the semi-boosted topology only all spectra are weighted to the luminosity of the $t\bar{t}$ sample for the comparison. The black dotted spectrum denoted as pseudodata corresponds to the same combined and weighted

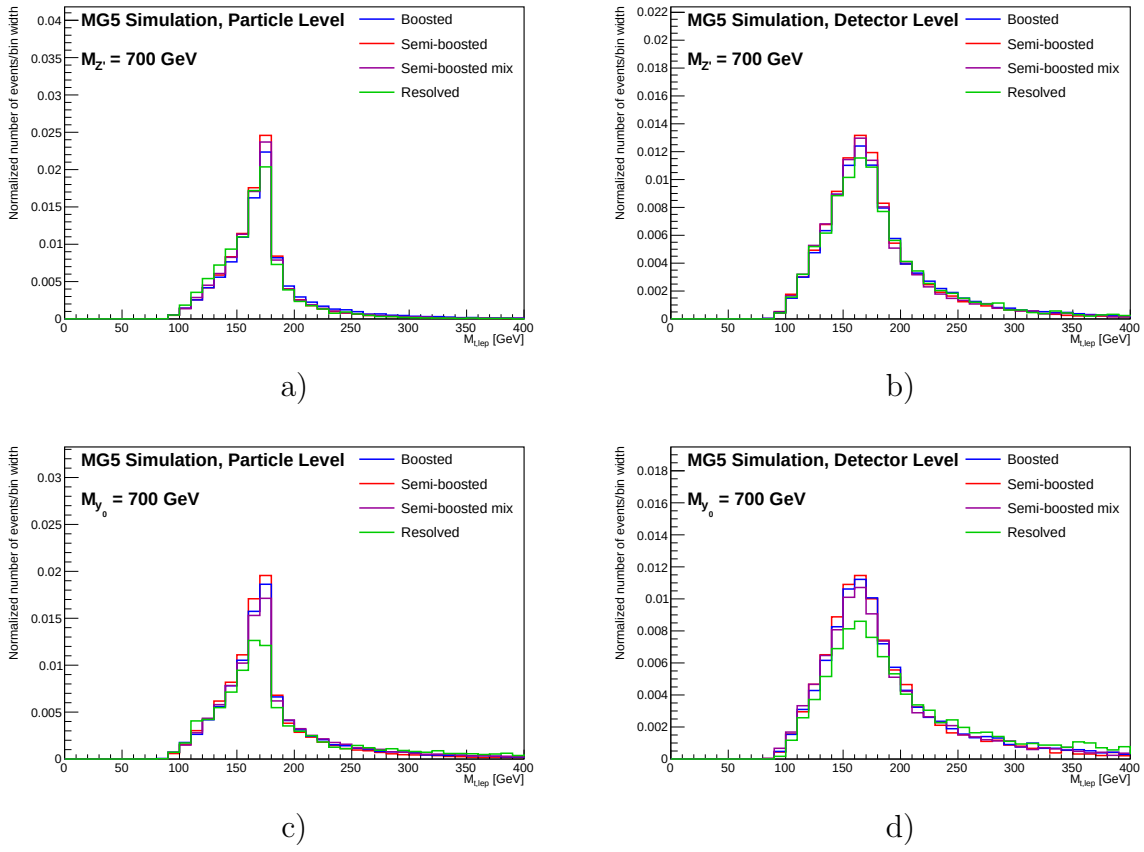


Figure 5.14: Comparison of the reconstructed leptonically decaying top quark mass spectra ($M_{\mathrm{t,lep}}$) between the particle (a), c)) and the detector (b), d)) level for all the topologies using the sample with $M_{Z'} = 700$ GeV (a), b)) and $M_{\mathrm{y_0}} = 700$ (c), d)).

spectrum from the statistically independent sample. The bottom part shows the ratio between the combined spectrum and pseudodata and ratio of combined sample with contribution from the hypothetical particle over the combined spectrum without the signal contribution, thus showing the strength of the signal.

The invariant mass of the reconstructed hadronically decaying top quark compared between the detector and the particle level is shown in Fig. 5.16. The reconstructed invariant mass of the hadronically decaying top quark is peaking around the theoretical value as in the case of the leptonically decaying top quark over all topologies. However, the shape of the spectra differs in mass peak width over the topologies, which is caused by the different approach of the reconstruction of the hadronic side for each topology. The invariant mass peak is in good agreement in comparison between the Z' and y_0 signal samples.

The combined spectrum of the reconstructed hadronically decaying top quark mass from all samples is shown in Fig 5.17 in the semi-boosted topology. The bottom part shows the ratio between the combined spectrum and pseudodata and the strength of the signal.

The rapidity of the hadronically decaying top quark compared between the detector and the particle level is shown in Fig. 5.18 for all topologies. This angular variable is comparable over the topologies. The combined spectrum of reconstructed hadronically decaying top quark rapidity is shown in Fig. 5.19 in the semi-boosted topology. The

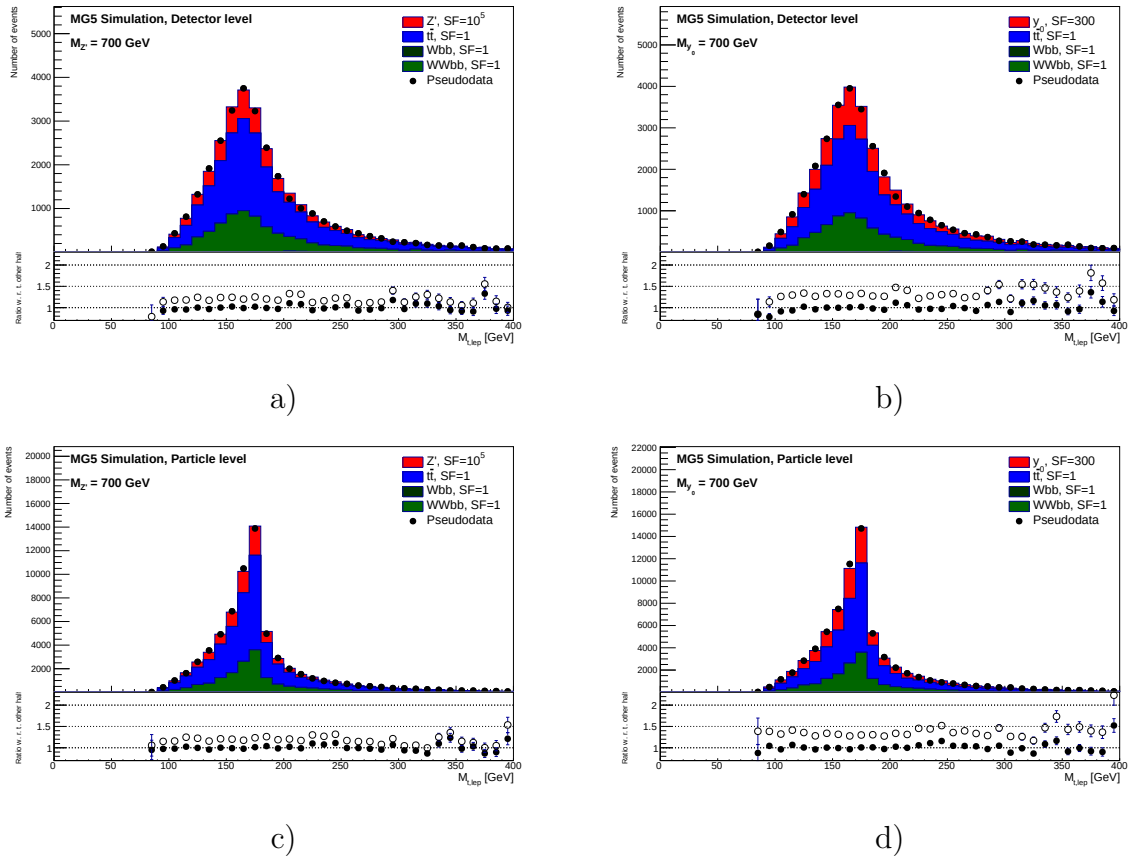


Figure 5.15: Combined spectrum over all samples of the reconstructed leptonically decayed top quark mass ($M_{\mathrm{t,lep}}$) at the detector (top row) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{\mathrm{Z'}} = 700$ GeV (a), c)) and with $M_{\mathrm{y_0}} = 700$ GeV (b), d)). All samples are weighted to the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

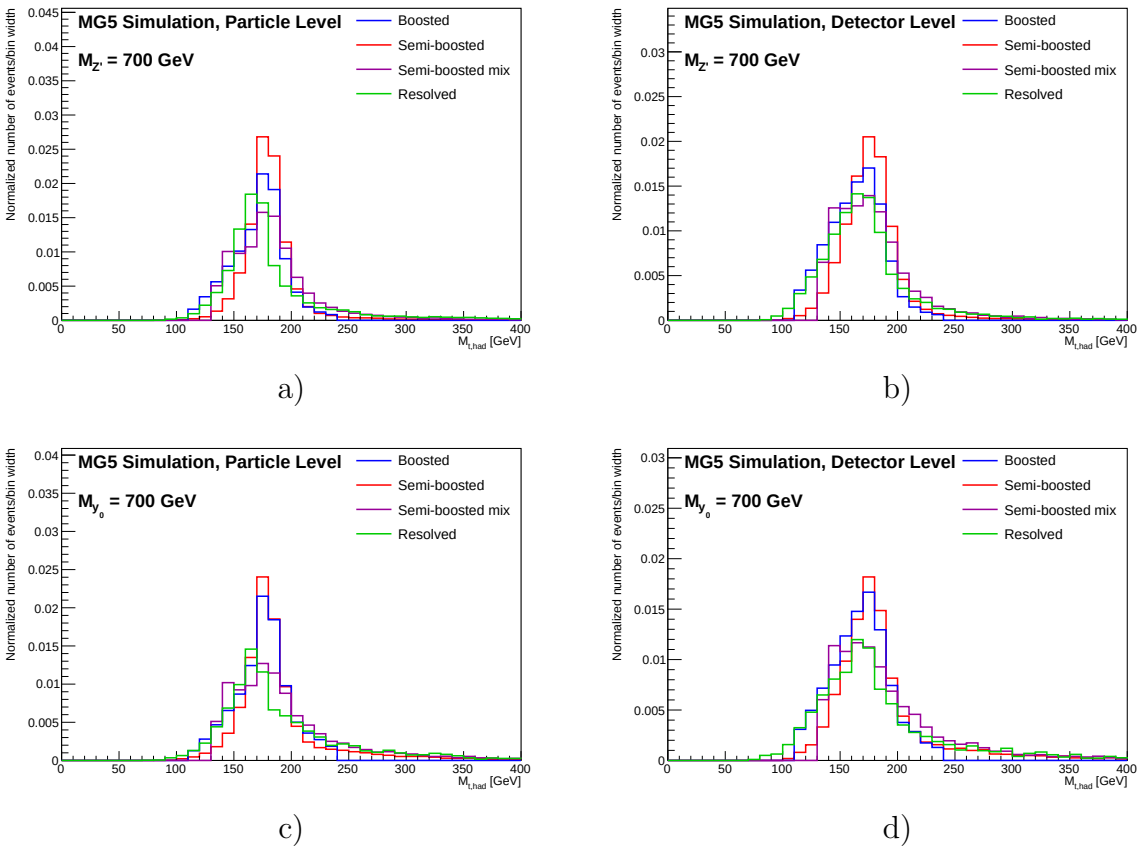


Figure 5.16: Comparison of the reconstructed hadronically decaying top quark mass spectra ($M_{\mathrm{t,had}}$) between the particle (a), c)) and the detector (b), d)) level for all the topologies using the sample with $M_{Z'} = 700$ GeV(a), b)) and $M_{\mathrm{y_0}} = 700$ (c), d)).

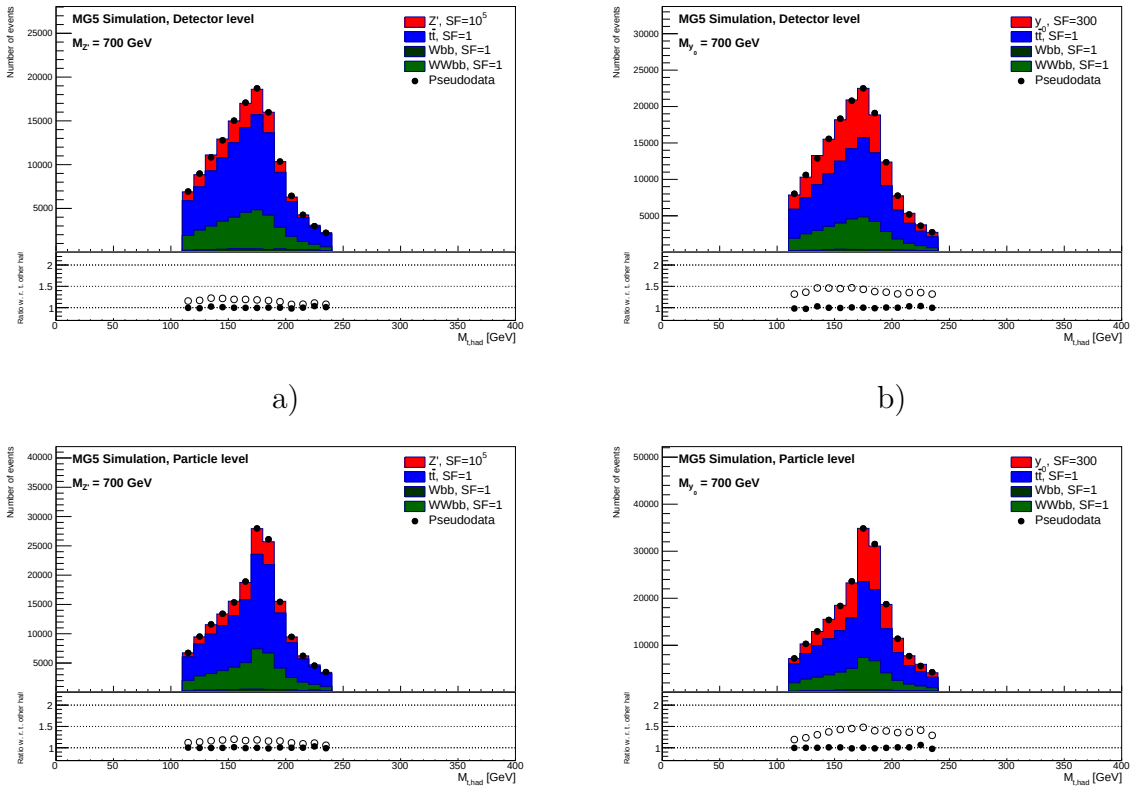


Figure 5.17: Combined spectrum over all samples of the reconstructed hadronically decayed top quark mass ($M_{\mathrm{t,had}}$) at the detector (top row) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{\mathrm{Z'}} = 700$ GeV (a), c)) and with $M_{\mathrm{y_0}} = 700$ GeV (b), d)). All samples are weighted to the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

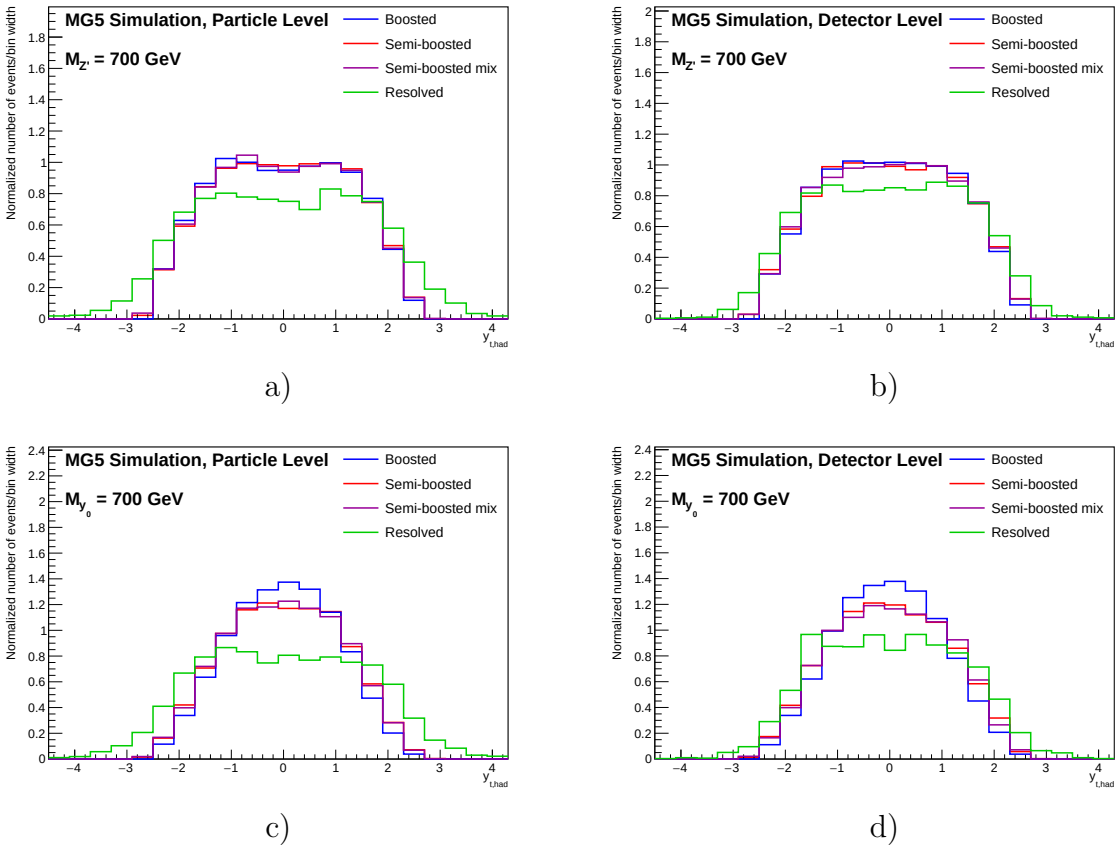


Figure 5.18: Comparison of the rapidity of the reconstructed hadronically decaying top quark ($y_{\mathrm{t,had}}$) between the particle (a), c)) and the detector (b), d)) level for all the topologies using the sample with $M_{Z'} = 700$ GeV(a), b)) and $M_{\mathrm{y_0}} = 700$ (c), d)).

bottom part shows the ratio between the combined spectrum and pseudodata and the strength of the signal.

The cosine of the production angle θ^* of the hadronically decayed top quark, see section 3.2.1 for details, compared between the detector and the particle level, is shown in Fig. 5.20 for all topologies. The difference between the topologies is well visible, where the boosted topology peaks around the zero value whereas the resolved topology prefers higher values of the cosine of the production angle, although shapes are similar when comparing between the Z' and the y_0 signal samples. The energy of the reconstructed event in boosted topology is higher, which may lead to the higher transverse momentum and the higher probability of detection in the central part of the detector. The resolved topology event performance is the opposite due to the same reasoning.

The combined spectrum of the reconstructed hadronically decaying top quark cosine of the production angle from all samples is shown in Fig 5.21 in the semi-boosted topology. The bottom part shows the ratio between the combined spectrum and pseudodata and the strength of the signal.

The reconstructed invariant mass of the top anti-top quark pair compared between the detector and the particle level is shown in Fig. 5.22. The invariant mass of the reconstructed system is well corresponding to the original generated hypothetical particle, in this case the invariant mass value is 700 GeV for both signal samples. The shapes are wider in all spectra at the detector level, which depicts the finite resolution

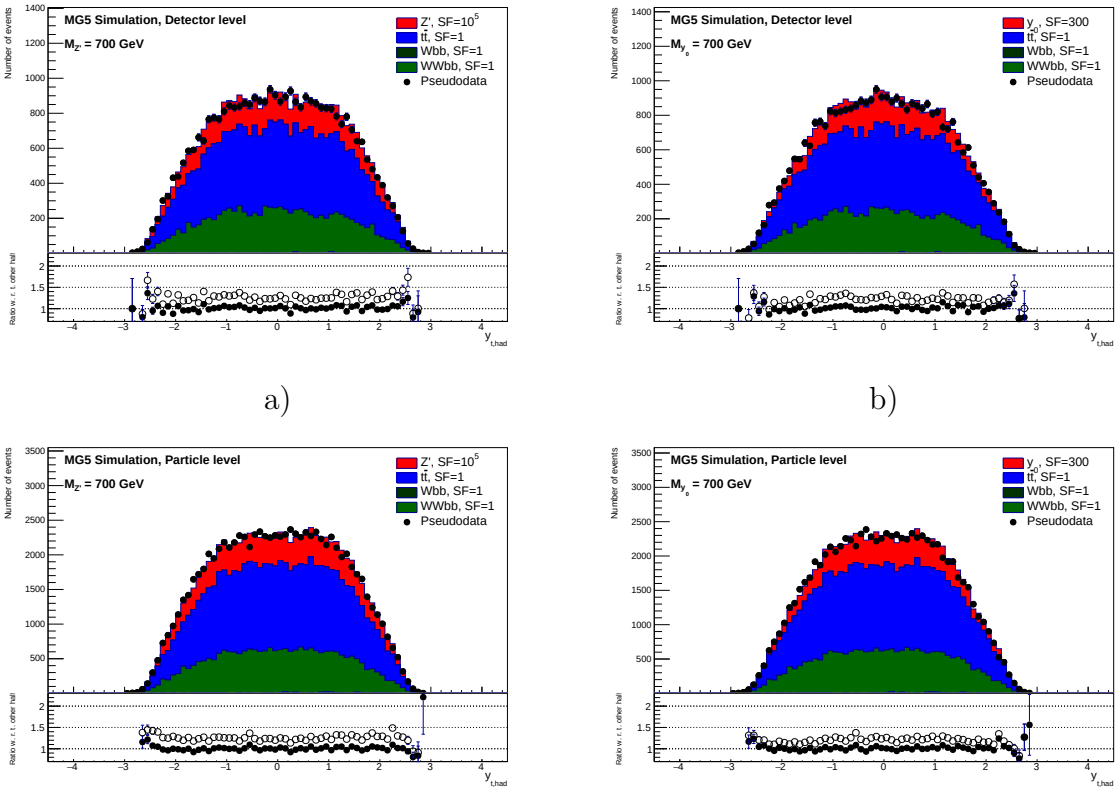


Figure 5.19: Combined spectrum of the rapidity of the reconstructed hadronically decaying top quark ($y_{\mathrm{t,had}}$) at the detector (top row) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{\mathrm{Z'}} = 700$ GeV (a), c)) and with $M_{\mathrm{y_0}} = 700$ GeV (b), d)). All samples are weighted to the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

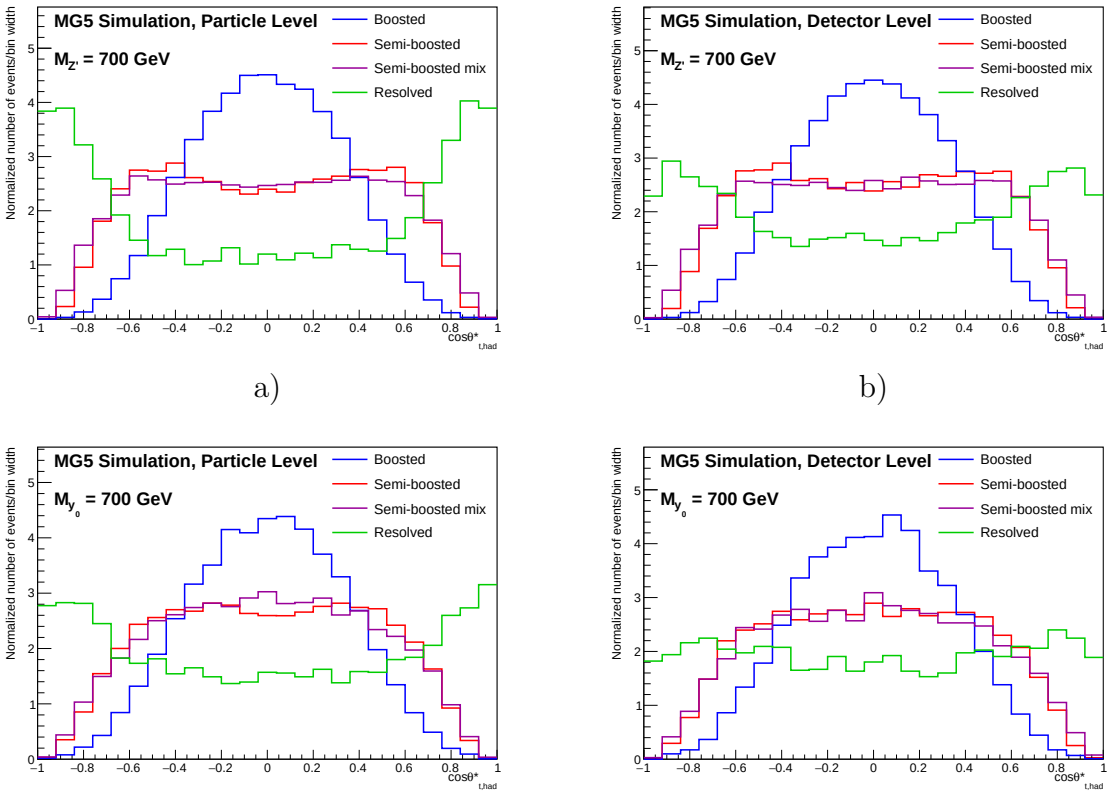


Figure 5.20: Comparison of the cosine of the production angle of the reconstructed hadronically decaying top quark ($\cos\theta^*_{\mathrm{t,had}}$) between the particle (a), c)) and the detector (b), d)) level for all the topologies using the sample with $M_{Z'} = 700$ GeV(a), b)) and $M_{\mathrm{y_0}} = 700$ (c), d)).

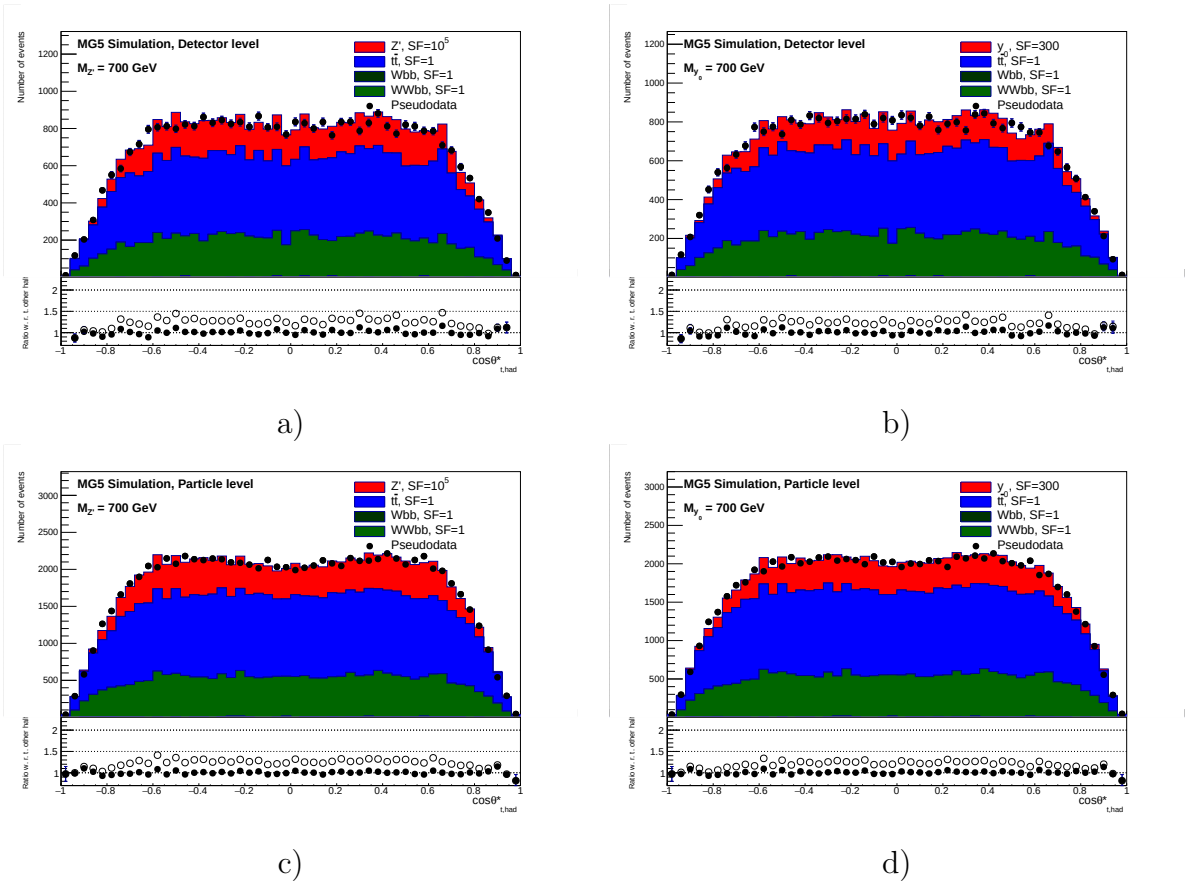


Figure 5.21: Combined spectrum of the cosine of the production angle of the reconstructed hadronically decaying top quark ($\cos\theta^*_{\mathrm{t,had}}$) at the detector (top row) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{\mathrm{Z}'} = 700$ GeV (a), c)) and with $M_{\mathrm{y_0}} = 700$ GeV (b), d)). All samples are weighted to the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

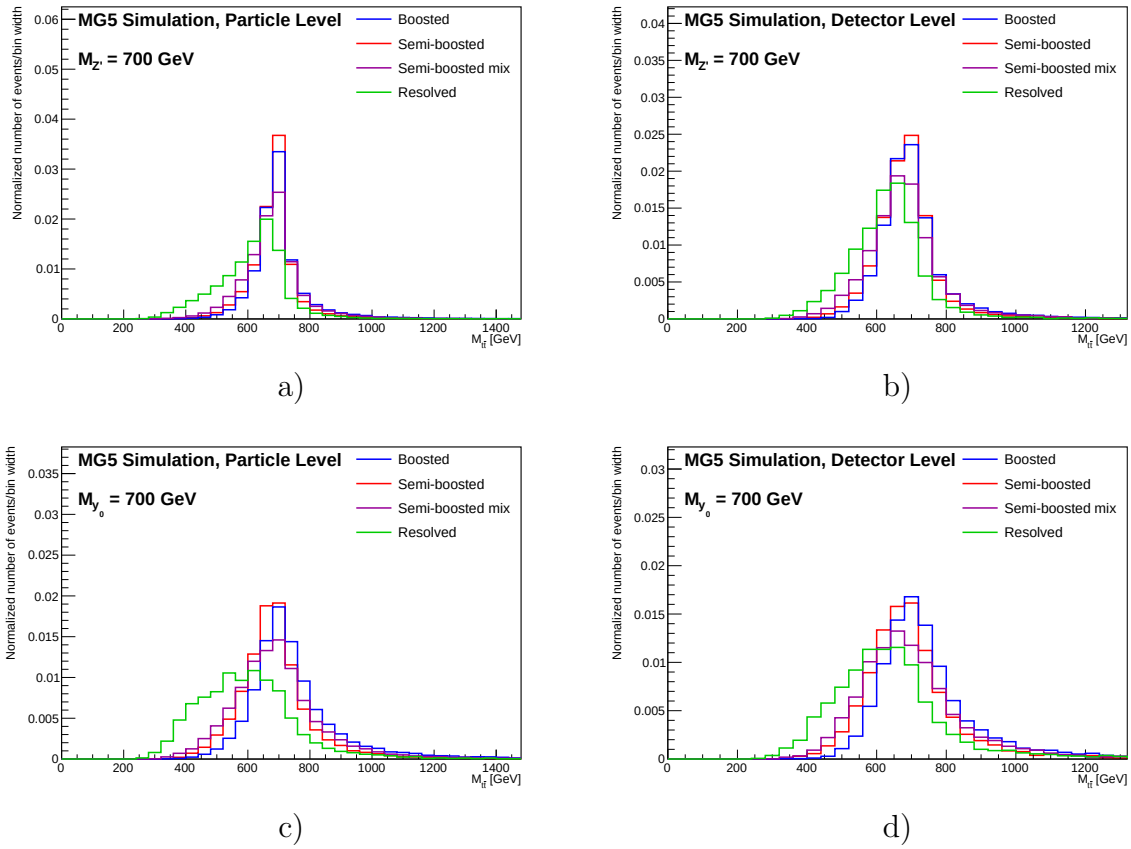


Figure 5.22: Comparison of the reconstructed invariant $t\bar{t}$ pair mass spectra ($m_{\mathrm{t\bar{t}}}$) between the particle (a), c)) and the detector (b), d)) level for all the topologies using the sample with $M_{\mathrm{Z'}} = 700$ GeV (a), b)) and $M_{\mathrm{y_0}} = 700$ GeV (c), d)).

of the detector in the DELPHES simulation. Only the resolved topology spectrum is deviating from other topologies in shape, which is probably caused by the large combinatorial background for this topology.

The combined spectrum of the reconstructed hadronically decaying top quark cosine of the production angle from all samples is shown in Fig 5.21 in the semi-boosted topology. The bottom part shows the ratio between the combined spectrum and pseudodata and the strength of the signal. Combined spectra of the reconstructed invariant mass of the top anti-top quark pair are shown in detail in Appendix A for various signal samples and topologies.

The event yields for each sample is presented in Table 5.1 at the detector and particle levels for each topology after weighting to the same luminosity, thus corresponding to the Fig. 5.23 and also with signal event fractions over the topologies in Fig. 5.27.

Examples of reconstructed events in the $\eta - \phi$ plane or different topologies is shown in Fig. 5.24 for illustration of location of the selected jets and lepton in the detector. The event arrangement in the detector for the semi-boosted mixed topology differs from the semi-boosted topology one only in a switching of the b-tag between the selected small jet and the large jet on hadronic side, thus, the localization of objects is very similar.

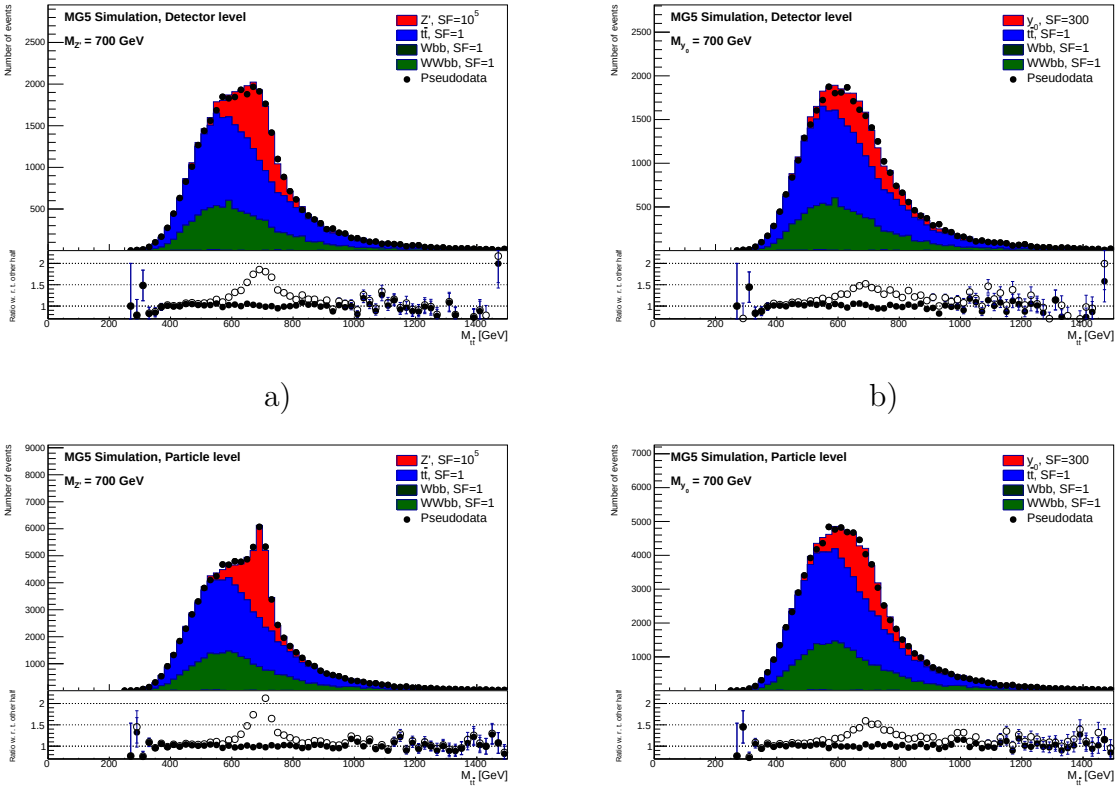


Figure 5.23: Combined spectrum of the reconstructed invariant $t\bar{t}$ pair mass spectra ($m_{\mathrm{t\bar{t}}}$) at the detector (top row) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{\mathrm{Z'}} = 700$ GeV (a), c)) and with $M_{\mathrm{y_0}} = 700$ GeV (b), d)). All samples are weighted to the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

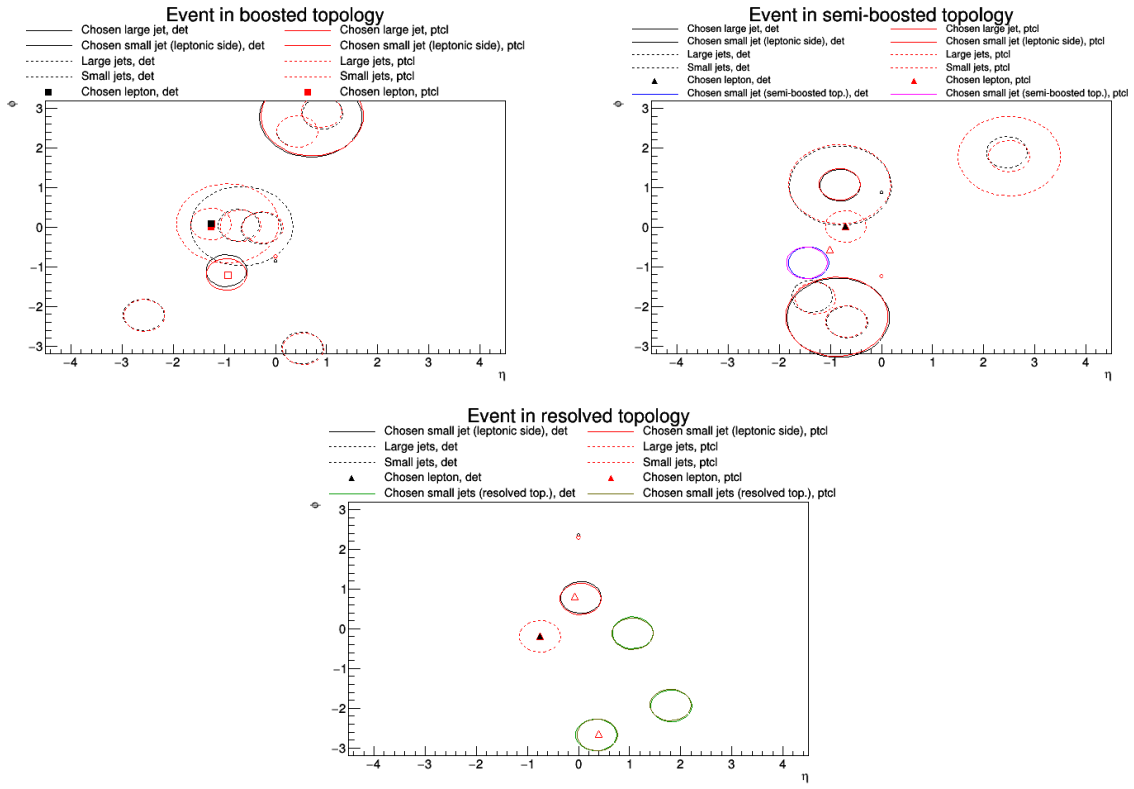


Figure 5.24: The example of reconstructed events in the boosted (left), the semi-boosted (right) and the resolved topologies with locations of selected jets and a lepton in the $\eta - \phi$ plane of the ATLAS detector. Large circles denotes large jets, small circles small jets, where the selected jets are drawn with a solid line and smallest circles is the direction of the missing transverse energy. The circles in black are drawn for the detector level selection whereas red is describing the particle level constituents. The green circles denote selected jets in the resolved topology, while the blue and magenta circles denotes selected b-tagged jet for the semi-boosted and resolved topology on the detector and particle levels respectively. The selected lepton is denoted by a triangle or a square full marker depicting electrons and muons respectively. Jets, which are not selected by the algorithm are drawn by a dashed line and not selected leptons with empty markers.

Table 5.1: The rounded event yields for samples at the detector and particle levels after weighting to the $t\bar{t}$ sample luminosity, where R, SBM, SB and B denote the resolved, semi-boosted mixed, semi-boosted and boosted topologies.

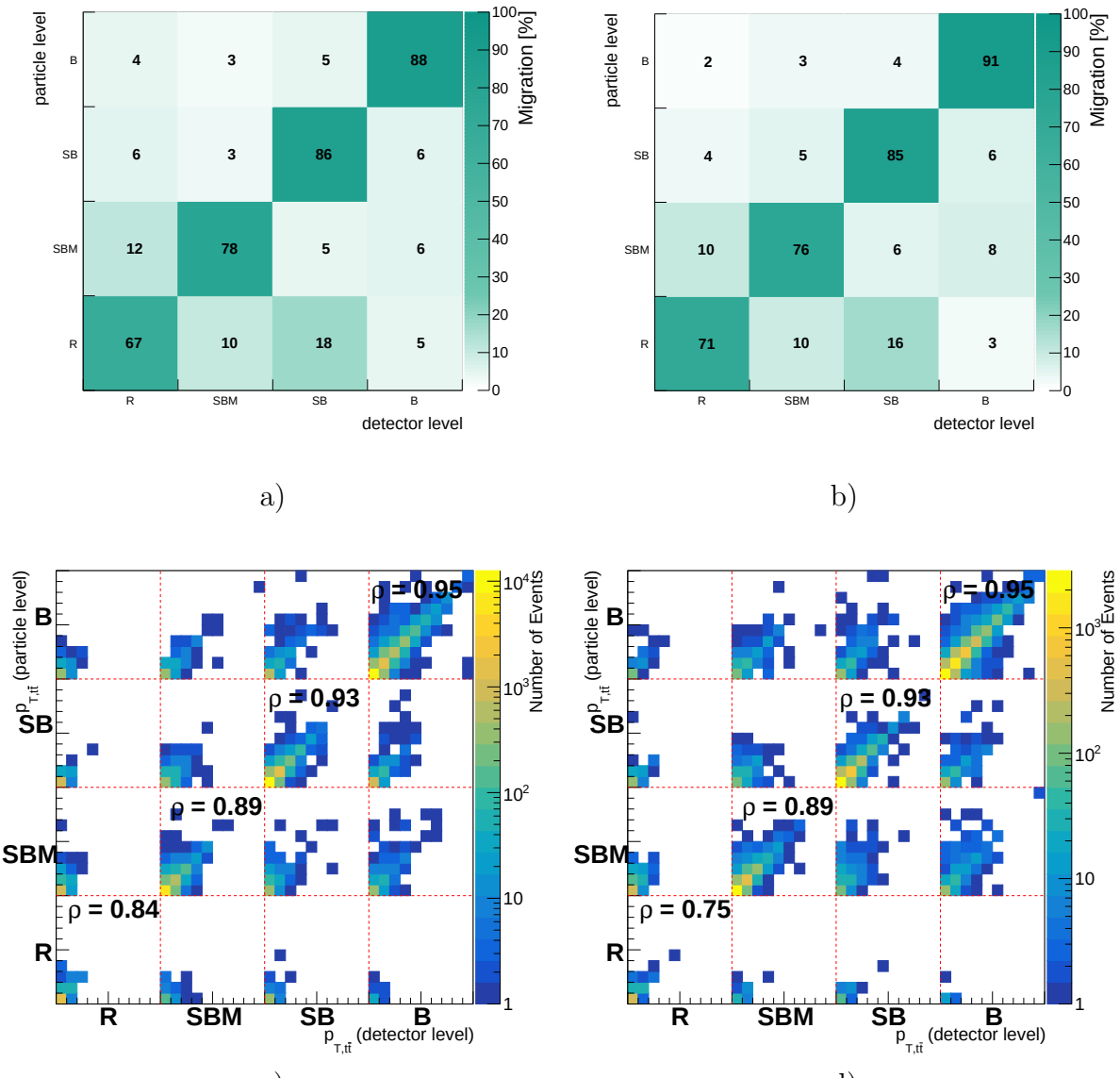
Level	Particle level				Detector level			
Sample, Topology	R	SBM	SB	B	R	SBM	SB	B
Wbb + jets	417	176	84	110	104	88	34	52
$WWbb$ + jets	24699	18191	15380	11020	11774	4963	5917	5184
Z'(M = 1000 GeV)	2406	19117	19465	89097	2370	5176	7605	40940
Z'(M = 900 GeV)	3191	20088	24044	71654	2720	5363	9410	33248
Z'(M = 800 GeV)	3617	20616	28906	47560	3144	5510	11162	22444
Z'(M = 700 GeV)	4415	20614	32358	20508	3478	5631	12379	9796
Z'(M = 600 GeV)	6602	16920	25217	5054	4589	4959	9762	2314
Z'(M = 500 GeV)	11627	9849	8637	1258	6500	2913	3348	535
y_0(M = 1000 GeV)	1064	11481	6463	31812	999	2574	2613	13717
y_0(M = 900 GeV)	1223	11487	7435	26375	1165	2720	2962	11721
y_0(M = 800 GeV)	1459	11206	8593	19219	1336	2653	3381	8646
y_0(M = 700 GeV)	1826	10483	9079	11053	1413	2570	3611	5017
y_0(M = 600 GeV)	2440	8773	8098	4681	1777	2157	3107	2122
y_0(M = 500 GeV)	3897	5749	4374	1776	2266	1581	1745	743
$t\bar{t}$ sample	66236	54530	46485	37190	32379	14806	17852	17015

5.4 Topology Event Fractions and Resolution

The reconstruction is performed independently at the detector and particle levels, thus it can lead to different reconstructed topology on each level. Migration of events between topologies and between the detector level (horizontal axis) and particle levels (vertical axis) is shown in Fig. 5.25 for the two main signal samples. The same is shown for the transverse momentum of reconstructed top anti-top quark pair ($p_{\mathrm{T,t\bar{t}}}$) and its migration Fig. 5.25.

Contributions from particular topologies to the number of reconstructed events for the mass spectrum of top anti-top quark pair is shown in Fig. 5.26 at the particle (a), b)) and detector levels (c), d)) and for different mass values of hypothetical particles in Fig. 5.27. The resolution for particular topologies can be extracted from widths of the mass spectra of the top anti-top quark pair and can be compared between topologies. The resolution of the spectra is wider at the detector level in general, which is expected due to the energy fluctuations of reconstructed jets in the simulated detector. However, the spectra are similar between the signal samples for both the particle and detector levels, although the peaks are wider for the sample with $M\mathrm{y}_0 = 700$ GeV.

The resolution was extracted as the width of the fitted of normal distribution, this information is shown in Fig. 5.28 in absolute values (a), b)) and relative values (c), d)) with respect to mass of hypothetical particle Z' (a), c)) and y_0 (b), d)). The Landau distribution was also tested for fitting, but the resulting error of fitting was larger than in case of fitting with normal distribution function. The resolution heavily depends on the combination of topology and mass of the studied hypothetical particle. For example the resolved topology has the best resolution for mass 500 GeV and then deteriorates with increasing mass. On the other hand, the boosted topology resolution performance



a) b)

c) d)

Figure 5.25: Migration of reconstructed events between reconstructed levels and topologies (a), b)) and migration of events between topologies and bins for transverse momentum of reconstructed top anti-top quark pair $p_{\mathrm{T,t\bar{t}}}$ (c), d)) for samples with $M_{Z'} = 700$ GeV (a), c)) and $M_{\mathrm{y_0}} = 700$ (b), d)).

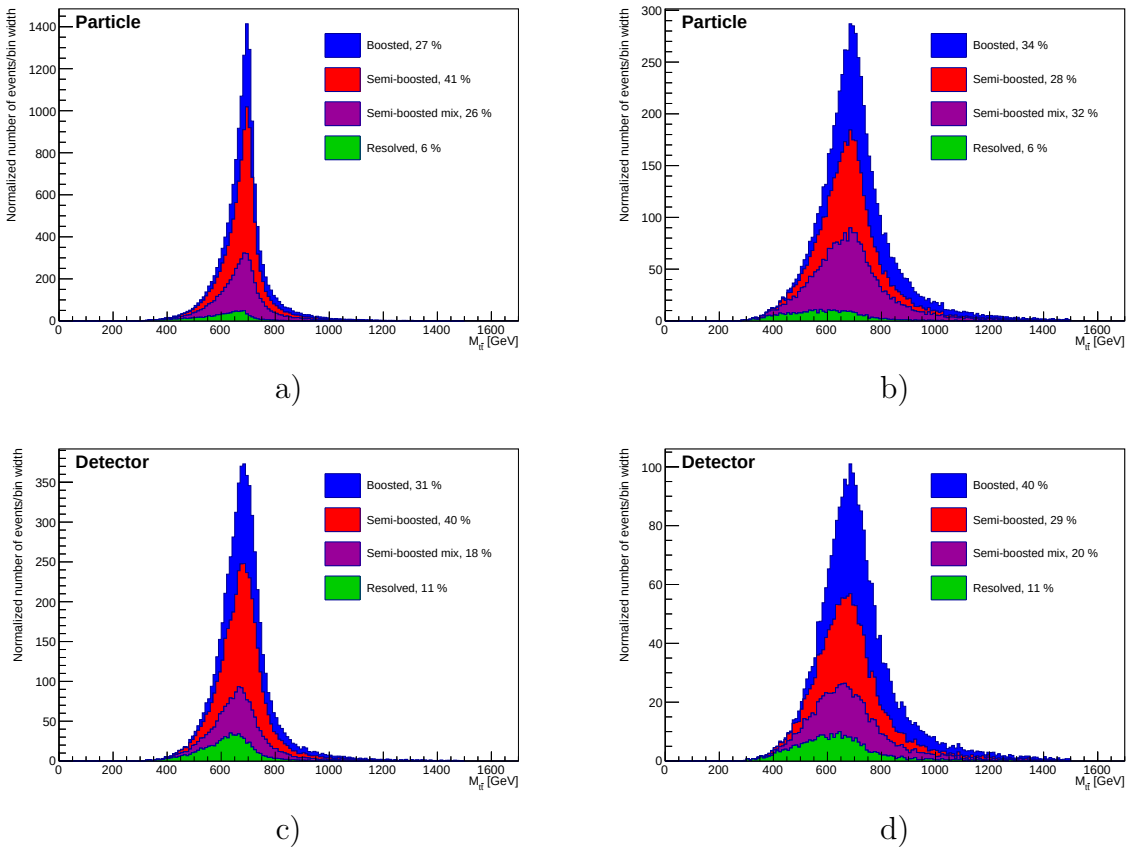


Figure 5.26: Contributions from particular topologies to the mass spectrum of reconstructed top anti-top quark pair $m_{\mathrm{t\bar{t}}}$ on the particle level (a), b)) and the detector levels (c), d)) for samples with $M_{Z'} = 700$ GeV (a), c)) and $M_{\mathrm{y_0}} = 700$ (b), d).

has an opposite trend, where the samples with higher mass have better resolution for this topology, which corresponds to the increasing number of reconstructed events. In general, the resolution is slightly better for the samples with Z' as probing particle.

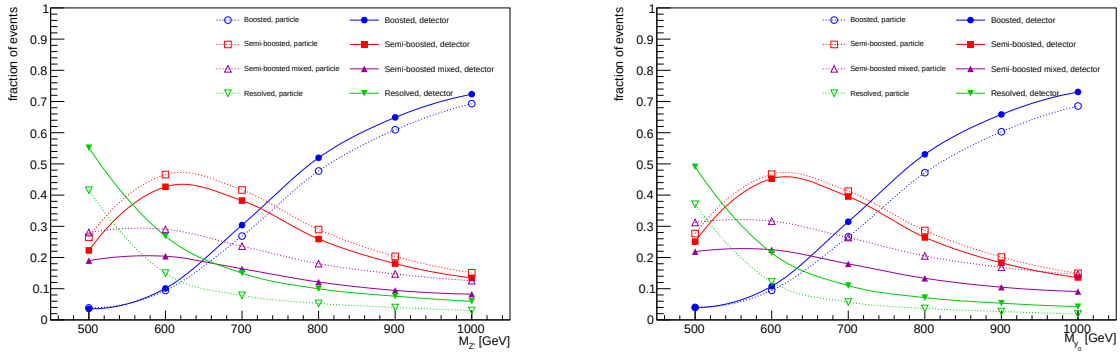


Figure 5.27: The fraction of events contributing to the $t\bar{t}$ reconstruction from each topology over samples with various masses of the hypothetical Z' (left) and y_0 (right) particle ($M_{Z'}$and M_{y_0}) at the detector (solid lines, full markers) and particle (dotted lines, open markers) levels.

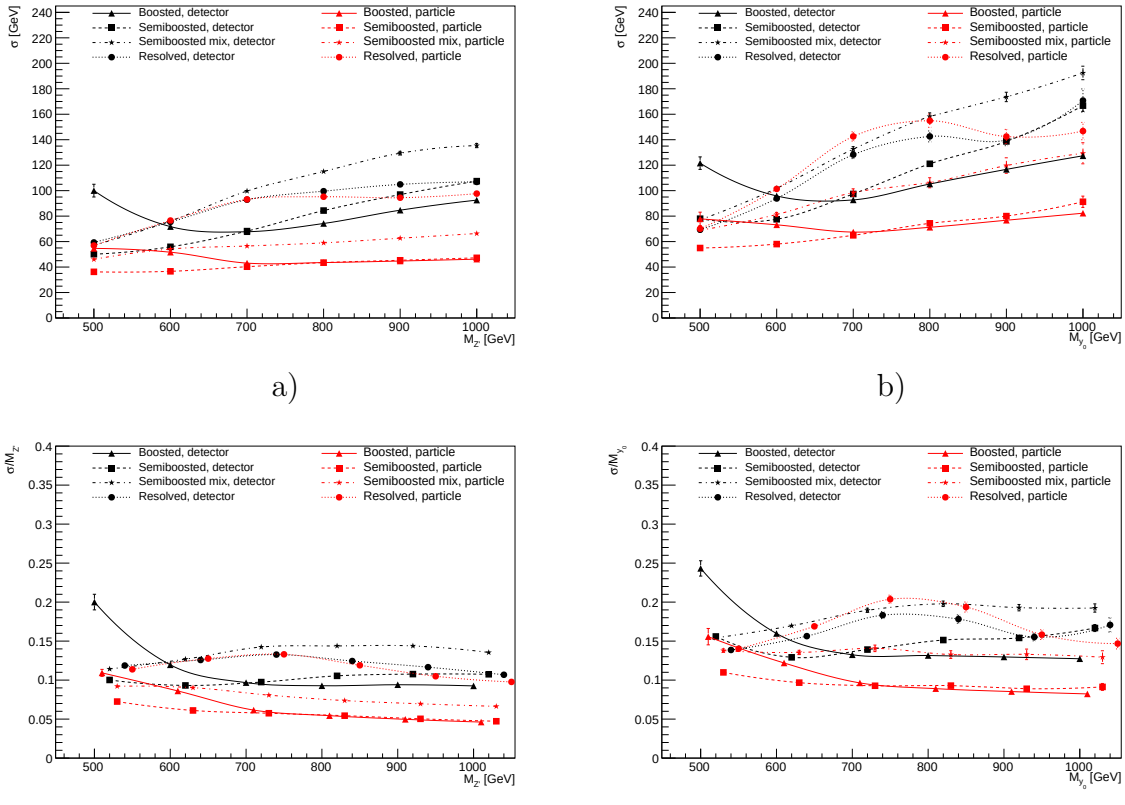


Figure 5.28: Comparison of the $t\bar{t}$ resolution for samples with different mass of hypothetical particle Z' (a),c)) and y_0 (b),d)) in particular topologies in absolute values (a), b)) and relative values (c),d).

Chapter 6

The Unfolding Procedure and Signal Significance

The unfolding procedure and its performance on generated samples is presented in this chapter. The change in the strength of the hypothetical Z' or y_0 particle signal is studied before and after the unfolding procedure.

The first section describes the theory behind the unfolding procedure, followed by definition of the significance of the signal. Results of closure tests and the check of the unfolding procedure performance are presented before the results with chosen variables and samples are shown. Final results of the unfolding procedure are performed on an admixture of the signal Z' or y_0 sample, the $t\bar{t}$ sample and two background samples. Finally the comparison of significances between different topologies is described in the last section.

6.1 Unfolding Procedure

Spectra measured in every detector suffer in precision by the influences of the detector, in our case the simulation of the ATLAS detector. The unfolding procedure corrects for these influences, for example resolution of the detector and granularity of the separate subdetectors. There are several unfolding methods, which could possibly yield different results. The Fully Bayesian Unfolding (FBU) is chosen to mediate the unfolding procedure and was performed by the PyMC3 package [102, 103]. The FBU method estimates the probability density functions (posteriors) from the detector level spectrum values in each bin (priors) under conditions given by correction factors and the migration matrix. The most probable value in each bin can be taken as the unfolded value of the spectra of interest in the bin coresponding to the particle level spectrum. Among the main advantages of this method are that the migration matrix is not modified as in the singular value decomposition [104], absence of iterations in calculation in comparison with iterative unfolding [105], which also uses the Bayesian theorem, and the control over the result as all extremes of the probability density function are revealed and not only the one to which the procedure converges as in the iterative Bayesian unfolding. A general unfolding process is described by the following formula

$$\hat{T_{\mathrm{i}}} = \frac{1}{f_{\mathrm{i,eff}}} M^{-1}_{\mathrm{ij}} f_{\mathrm{j,acc}} (D_{\mathrm{j}} - B_{\mathrm{j}}), \tag{6.1}$$

where $\hat{T_{\mathrm{i}}}$ is the estimated value in the bin i of the particle level spectrum, $f_{\mathrm{i,eff}}$ and $f_{\mathrm{j,acc}}$

are efficiency and acceptance correction factors, defined in the next formula, M_{ij}^{-1} is embodiment of the unfolding procedure [1], D_{j} is the measured detector level spectra in bin j and finally B_{j} is the subtracted background contribution to the measured detector level spectrum. The efficiency and acceptance correction factors are defined as follows

$$f_{\mathrm{i,eff}} = \frac{P^{\mathrm{match}}_{\mathrm{t\bar{t},i}}}{P_{\mathrm{t\bar{t},i}}} \quad \text{and} \quad f_{\mathrm{j,acc}} = \frac{D^{\mathrm{match}}_{\mathrm{t\bar{t},j}}}{D_{\mathrm{t\bar{t},j}}}, \tag{6.2}$$

where $P^{\mathrm{match}}_{\mathrm{t\bar{t},i}}$ is the particle level spectrum value in bin i, which fulfills the particle level selection criteria and its detector level matched counterpart fulfills the detector level selection criteria; $P_{\mathrm{t\bar{t},i}}$ is the particle level spectrum value in bin i, which fulfill the particle level selection criteria regardless on its detector level counterpart, both evaluated on the $t\bar{t}$ sample. Similarly, $D^{\mathrm{match}}_{\mathrm{t\bar{t},j}}$ is the detector level spectrum value in bin j fulfilling particle level selection criteria and is matched to the detector level counterpart and $D_{\mathrm{t\bar{t},j}}$ is reconstructed detector level spectrum value in bin j without any matching to the particle level counterpart, both evaluated on the $t\bar{t}$ sample. The events contributing to the matched spectra are fulfilling the corresponding selection criteria and are reconstructed at both detector and particle levels in the same topology, thus the matching descriptor.

6.2 Significance of the Signal before and after the Unfolding

The spectra entering the unfolding procedure have the addition of the Z' or y_0 signal sample with an amplified cross section by an ad hoc number ($5 * 10^5$ and 300 respectively) to study the impact of the unfolding on the strength of the signal. The strength of the signal is quantified by a significance, which considers statistical errors of used samples in the given bin. The signal significance S in the given bin i before unfolding is defined for the hypothetical particle Z' as

$$S_{i,\mathrm{det}} = (P_{\mathrm{i}}^{t\bar{t}+Z'+Bg} - T_{\mathrm{i}}^{t\bar{t}} - Bg_{\mathrm{i},1} - ... - Bg_{\mathrm{i},n})/(\sqrt{\sigma_{P_{\mathrm{i}}}^2 + \sigma_{T_{\mathrm{i}}}^2}), \tag{6.3}$$

where P_{i} are the pseudo data consisting from the signal and the background added to the expected $t\bar{t}$ sample in the given bin, T_{i} is the detector level spectrum from the statistically independent $t\bar{t}$ sample and $Bg_{\mathrm{i},n}$ is the background contribution to the studied spectra from the n-th background sample, $\sigma_{P_{\mathrm{i}}}$ is the statistical error of the pseudo data sample in the given bin and $\sigma_{T_{\mathrm{i}}}$ is the statistical error of the bin value of the statistically independent $t\bar{t}$ sample.

A similar significance is defined after the unfolding procedure

$$S_{i,\mathrm{unf}} = (U_{\mathrm{i}}^{t\bar{t}+Z'} - T_{\mathrm{i}}^{t\bar{t}})/(\sqrt{\sigma_{U_{\mathrm{i}}^2} + \sigma_{T_{\mathrm{i}}}^2}), \tag{6.4}$$

where U_{i} are the unfolded data in the given bin, T_{i} is the particle level spectrum from the statistically independent $t\bar{t}$ sample, $\sigma_{U_{\mathrm{i}}}$ is the statistical error of the unfolded spectra in the given bin and $\sigma_{T_{\mathrm{i}}}$ is the statistical error of bin value of the statistically independent $t\bar{t}$ sample.

[1] The FBU unfolding does not use inverted migration matrix. The notation is traditional way to incorporated the unfolding procedure in the formula.

The detector and the unfolded significance plots are shown under the ratio plots in the unfolded spectra.

The integral significance, strength of the signal over the whole spectrum, is defined similarly for both the detector and the unfolded level. The detector integral significance is defined as

$$S_{\mathrm{I,det}} = \sum_{i=0}^{m} (P_{\mathrm{i}}^{t\bar{t}+Z'+Bg} - T_{\mathrm{i}}^{t\bar{t}} - Bg_{\mathrm{i},1} - ... - Bg_{\mathrm{i},n}) / \sqrt{\sum_{i=0}^{m} (\sigma_{P_{\mathrm{i}}}^2 + \sigma_{T_{\mathrm{i}}}^2)}, \tag{6.5}$$

where m is the number of bins in the given spectrum, P_{i} are the pseudo data consisting from the signal and the background added to the expected $t\bar{t}$ sample in the given bin, T_{i} is the detector level spectrum from the statistically independent $t\bar{t}$ sample and $Bg_{\mathrm{i},n}$ is the background contribution to the studied spectra from the n-th background sample, $\sigma_{P_{\mathrm{i}}}$ is the statistical error of the pseudo data sample in the given bin and $\sigma_{T_{\mathrm{i}}}$ is the statistical error of bin value of the statistically independent $t\bar{t}$ sample.

The detector integral significance is the same for all variables for given particle, *i.e.* the $M^{\mathrm{t\bar{t}}}$ spectrum and $p_{\mathrm{T}}^{\mathrm{t\bar{t}}}$ spectrum have the same value of the detector integral significance.

The integral significance for unfolded level or unfolded integral significance is defined similarly as

$$S_{\mathrm{I,unf}} = \sum_{i=0}^{m} (U_{\mathrm{i}}^{t\bar{t}+Z'} - T_{\mathrm{i}}^{t\bar{t}}) / \sqrt{\sum_{i=0}^{m} (\sigma_{U_{\mathrm{i}}^2} + \sigma_{T_{\mathrm{i}}}^2)}, \tag{6.6}$$

where m is number of bins in the given spectrum, U_{i} are the unfolded data in the given bin, T_{i} is the particle level spectrum from the statistically independent $t\bar{t}$ sample, $\sigma_{U_{\mathrm{i}}}$ is the statistical uncertainty of the unfolded spectra in the given bin and $\sigma_{T_{\mathrm{i}}}$ is the statistical error of bin value of the statistically independent $t\bar{t}$ sample.

The unfolded integral significance is varying slightly as the unfolding procedure is not able to reproduce exactly the same results but the resulting quantities are slightly shifting in the range of statistical error. Significances for the hypothetical particle y_0 are defined in the same manner as for the Z' particle.

The uncertainty of the detector level significance is estimated from 100 pseudoexperiments, where the detector level spectrum is smeared in each bin with the random number taken from a gaussian curve with the σ parameter set to the statistical uncertainty of the studied bin and mean parameter set to zero value. Pseudoexperiment values are then studied in each bin fitted by a gaussian curve and its σ parameter is the detector level significance uncertainty for the given bin. The unfolded significance uncertainty estimation is performed in a similar way, the spectrum is first smeared at the detector level and then unfolded and each bin is studied as in the detector level significance uncertainty estimation case.

6.3 Closure Test on Signal Samples

The closure test is used to validate the unfolding process, where the unfolded detector level spectrum is compared to the particle level one. The expected outcome of the test is that the ratio between the unfolded spectrum and particle level spectrum is close to one in range of the statistical uncertainty. The variable $p_{\mathrm{T}}^{\mathrm{t,had}}$ in semi-boosted

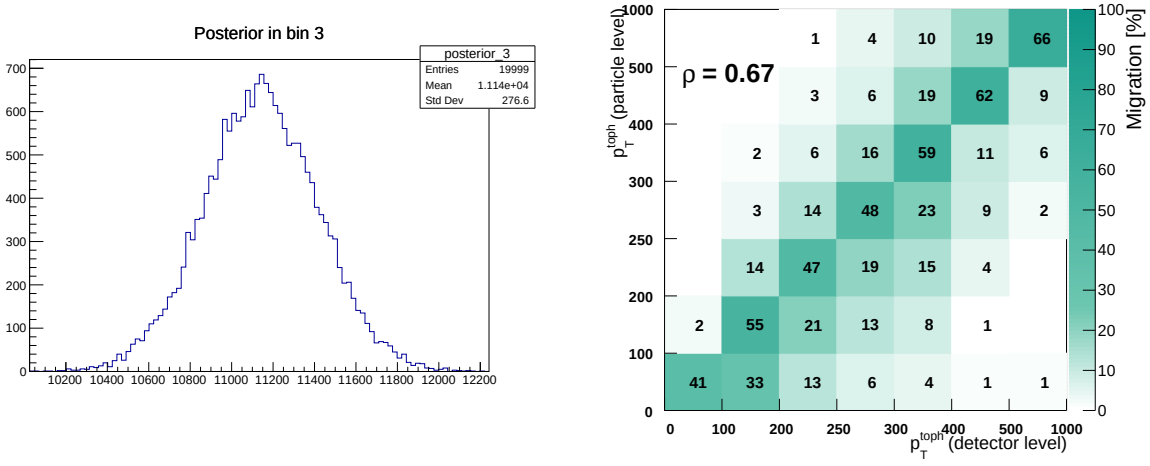


Figure 6.1: Example of a posterior of unfolding process in the third bin of $p_{\mathrm{T}}^{\mathrm{t,had}}$ spectrum (left) in semi-boosted topology and corresponding migration matrix for same variable for sample with $M_{\mathrm{Z}'} = 700$ GeV.

topology is chosen as example for the unfolding procedure from the Z' signal sample and the corrections from the $t\bar{t}$ sample are used. The software package Fully Bayesian Unfolding (FBU) was used for the unfolding [103]. The example of posterior in chosen bin is in Fig. 6.1 on the left, corresponding migration matrix is in the same figure on the right. Resulting unfolded spectrum and relative unfolded spectrum with respect to the particle level is shown in Fig. 6.2. The brown belt in the relative spectrum is set to the width of posterior distribution in each bin of the spectrum.

6.4 Unfolding with Background

The results for the unfolding of the signal samples with hypothetical particles Z' and y_0 are presented in this section for chosen variables. The focus is set to the samples with $M_{\mathrm{Z}'} = 700$ and $M_{\mathrm{y_0}} = 700$ where the studied semi-boosted topologies fraction of events is high. The transverse momentum, rapidity and cosine of the production angle of the hadronically decaying top quark ($p_{\mathrm{T}}^{\mathrm{t,had}}$, $y_{\mathrm{t,had}}$ and $\cos\theta^{*}_{\mathrm{t,had}}$) and the top anti-top quark pair invariant mass ($M_{\mathrm{t\bar{t}}}$) are chosen for examples of the unfolding procedure in the semi-boosted topology. The first step in the unfolding process is to mix spectra with background and $t\bar{t}$ sample and weight them to same luminosity, the luminosity of the $t\bar{t}$ sample was used in our case. The resulting mixed spectra entering the unfolding are shown in Fig. 6.3 and in Fig. 6.4 for the aforementioned variables. The black dotted spectrum corresponds to the same mixed spectra from the statistically independent sample part for verification of the mixed spectrum.

The ad hoc weighting parameter of signal samples was introduced to scale up the normalization of the sinal. Signal samples would have small fraction of events due to their small cross section. An add hoc scaling parameter is set to 300 for the samples with hypothetical particle y_0 and further by five for the resolved topology for the sample with $M_{\mathrm{y_0}} = 700$ GeV. The value of this parameter for the sample with $M_{\mathrm{Z}'} = 700$ GeV is set to 10 000 for all topologies.

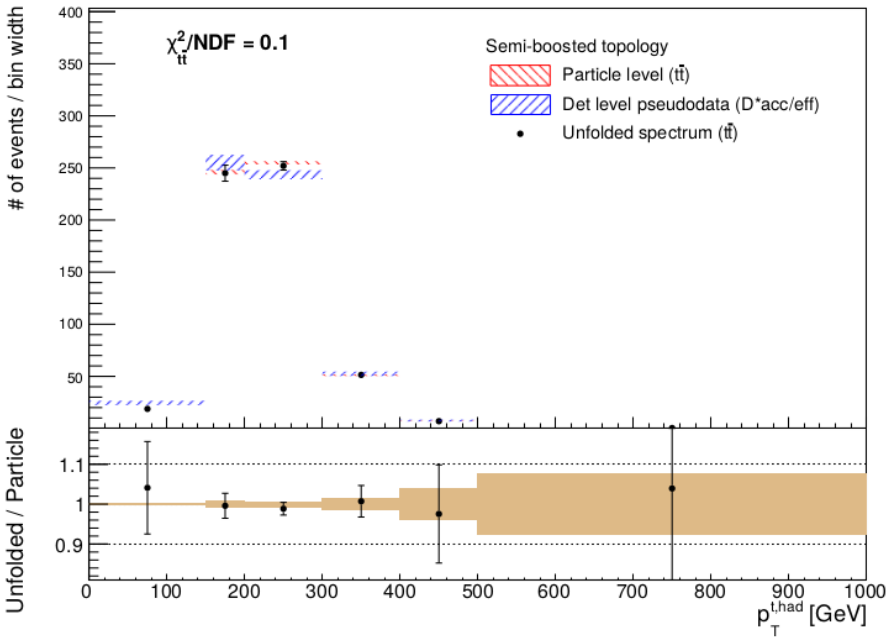


Figure 6.2: Example of spectra of variable $p_{\mathrm{T}}^{\mathrm{t,had}}$ for semi-boosted topology compared between detector level, particle level and unfolded spectra for $t\bar{t}$ sample. The hatched area corresponds to the statistical uncertainty of the spectrum. The $\chi^2_{\mathrm{t\bar{t}}}$ test is performed between $t\bar{t}$ particle level spectrum and unfolded $t\bar{t}$ spectrum from the statistically independent sample. The lower part of the plot describes ratio between spectrum from unfolding procedure and particle level spectrum form the statistically independent $t\bar{t}$ sample also known as the closure test and the yellow bar area shows the statistical uncertainty of the particle level spectra of the $t\bar{t}$ sample.

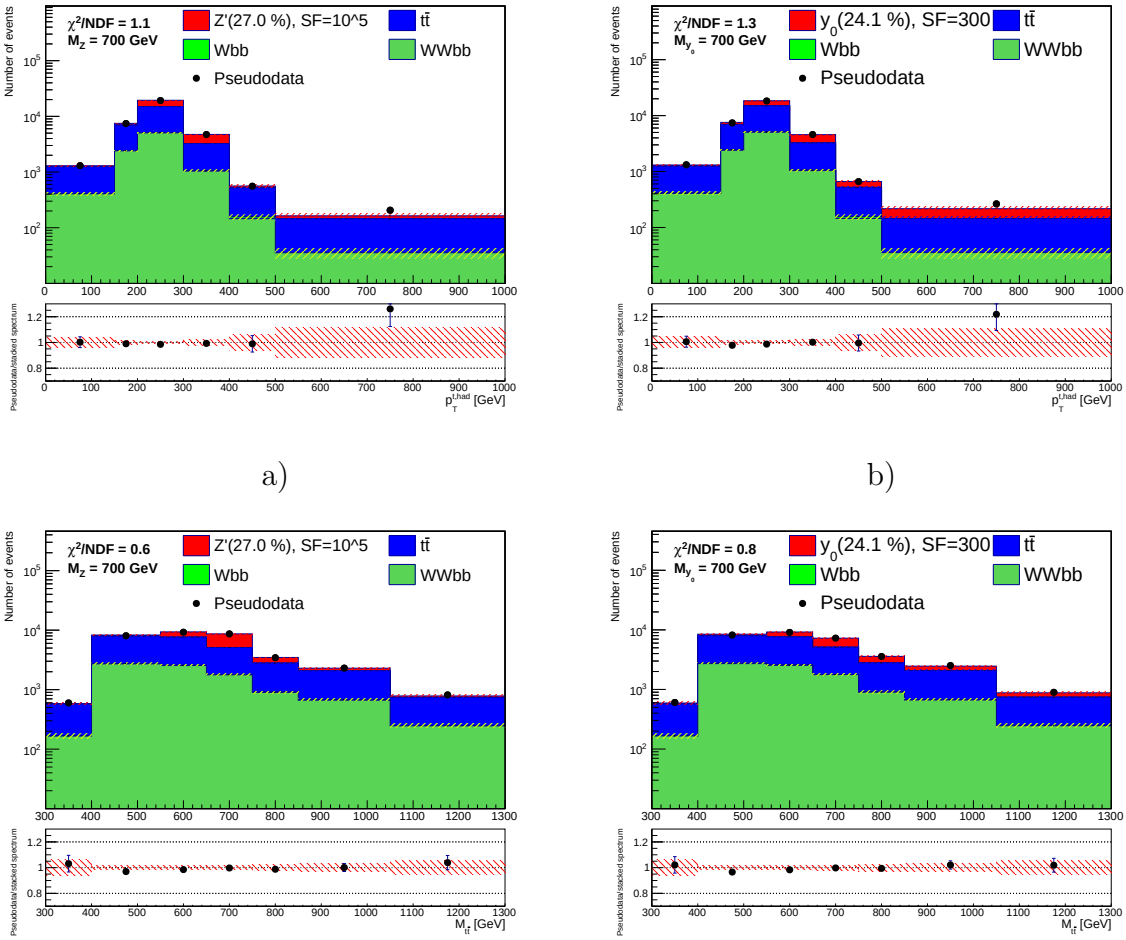


Figure 6.3: Comparison of the detector level spectra from two statistically independent parts (full markers and filled stack) for the $t\bar{t}$ sample with the addition of Wbb and $WWbb$ backgrounds and an admixture of events from the sample with $M_{\mathrm{y_0}} = 700$ GeV (a), c)) and the sample with $M_{\mathrm{Z'}} = 700$ GeV (b), d)) for the transverse momentum of the hadronically decaying top quark ($p_{\mathrm{T}}^{\mathrm{t,had}}$, a), b)) and for the top anti-top quark pair invariant mass ($M_{\mathrm{t\bar{t}}}$, c), d)), all spectra are reconstructed in the semi-boosted topology. The hatched bands in the top panel represent the statistical uncertainty in each sample.

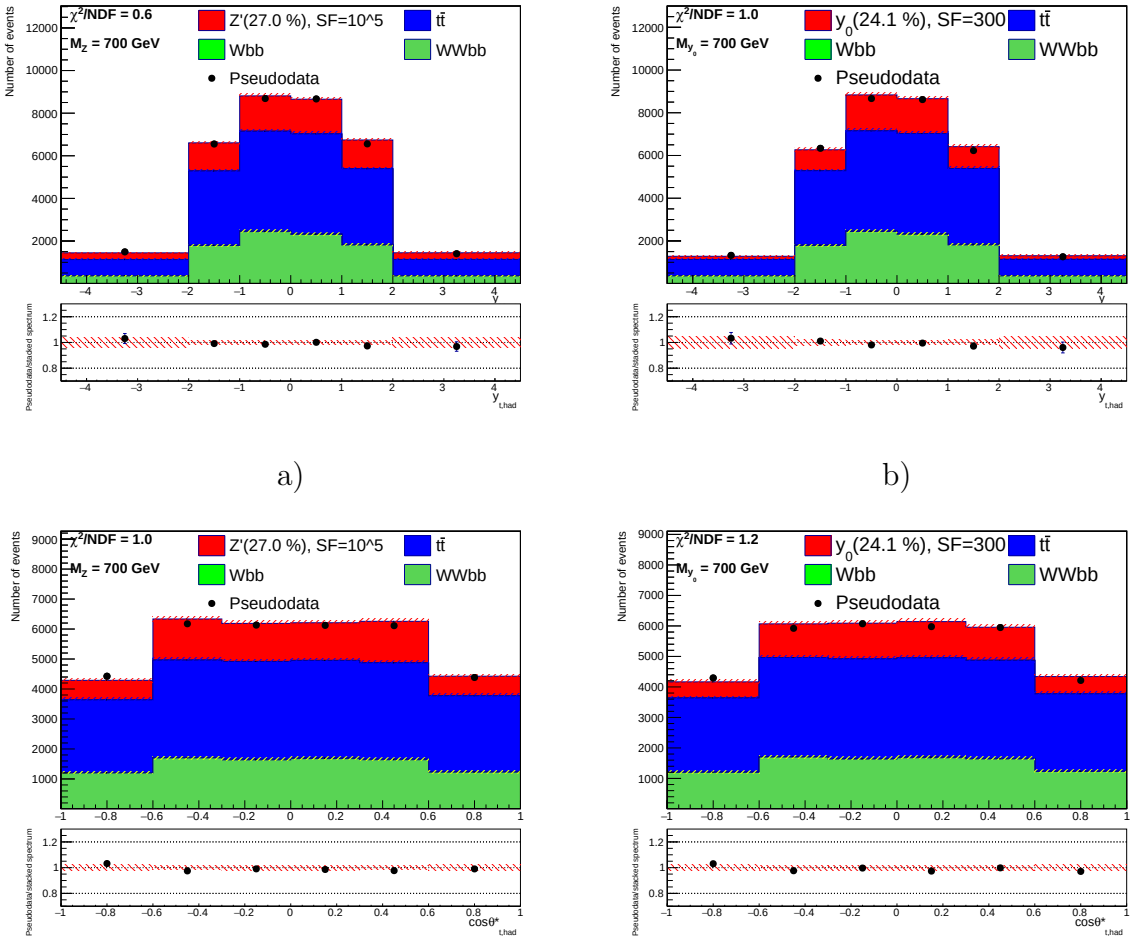


Figure 6.4: Comparison of the detector level spectra from two statistically independent parts (full markers and filled stack) for the $t\bar{t}$ sample with the addition of Wbb and $WWbb$ backgrounds and an admixture of events from the sample with $M_{\mathrm{y_0}} = 700$ GeV (a), c)) and the sample with $M_{\mathrm{Z'}} = 700$ GeV (b), d)) for the rapidity of the hadronically decaying top quark ($y_{\mathrm{t,had}}$, a), b)) and for the cosine of the production angle of the hadronically decaying top quark ($\cos\theta^*_{\mathrm{t,had}}$, c), d)), all spectra are reconstructed in the semi-boosted topology. The hatched bands in the top panel represent the statistical uncertainty in each sample.

Bin-by-bin corrections, also known as acceptance and efficiency corrections, are essential for unfolding as they correct for two phenomena. First, at the the detector level by removing non-signal events and at the particle level for undetected events. Those corrections are evaluated on the $t\bar{t}$ sample based on the equations shown in section 6.1. The $t\bar{t}$ sample was used for the determination of the corrections to increase correspondence to the real measurements for searches of unknown particles in the top quark spectra. Bin-by-bin corrections for chosen spectra are shown in Fig. 6.5.

Bin-by-bin migration of events between the detector and the particle levels is described by the migration matrix separately for each topology. The migration matrices for the transverse momentum, rapidity and the cosine of the production angle of hadronically decaying top quark ($p_{\mathrm{T}}^{\mathrm{t,had}}$, $y_{\mathrm{t,had}}$ and $\cos\theta^{*}_{\mathrm{t,had}}$) and the top anti-top quark pair invariant mass ($M_{\mathrm{t\bar{t}}}$) are shown in Fig. 6.6. Migration matrices are evaluated on the $t\bar{t}$ sample to achieve consistency with acceptance and efficiency corrections.

The results from the unfolding procedure for the chosen variables, namely the transverse momentum ($p_{\mathrm{T}}^{\mathrm{t,had}}$), rapidity of the hadronically decaying top quark ($y_{\mathrm{t,had}}$) and cosine of the production angle of hadronically decaying top quark ($\cos\theta^{*}_{\mathrm{t,had}}$) and the top anti-top quark pair invariant mass ($M_{\mathrm{t\bar{t}}}$), are shown in Fig. 6.7 and Fig. 6.8. The unfolding results are compared with detector level spectra corrected for the acceptance and efficiency to show the influence of the unfolding in the spectra in the top pad. The ratio between the unfolded and particle level spectra from the statistically independent sample is shown in middle tab. The bottom part of each plot shows the detector and unfolded significances, in general, the signal is stronger in the spectrum before unfolding than in the unfolded spectrum. This phenomenon manifests over all samples, topologies and variables, with exception for the cosine of the production angle of hadronically decaying top quark spectrum in the boosted topology.

The unfolding results show good agreement in the "closure" chi-squared test ($\chi^2_{\mathrm{t\bar{t}+Z'}}$ and $\chi^2_{\mathrm{t\bar{t}+y_0}}$), while the chi-squared test between unfolded and $t\bar{t}$ shows non-negligible influence of the signal. In general, the detector significance is higher in all studied spectra of all variables in all samples and topologies. This is manly caused by the difference between detector level spectra statistical uncertainty and width of the posteriors in the unfolding procedure, which is larger roughly by factor of two and is discussed in the next section.

6.5 Significance of the Signal and Influence of Unfolding Procedure

The significance of the signal with respect to the $t\bar{t}$ sample is evaluated on the detector and unfolded levels by using formula from previous section 6.2. The uncertainty of the significance is evaluated using 100 pseudoexperiments, which are unfolded afterwards. Each pseudoexperiment spectrum is prepared from the admixture of spectra entering the unfolding by smearing content in each bin by amount randomly picked from the statistical uncertainty interval. The gaussian distribution function is used for the random selection of the smearing value from the aforementioned interval. Example of spectra of pseudoexperiments for the semi-boosted topology for the transverse momentum of the hadronically decaying top quark $p_{\mathrm{T}}^{\mathrm{t,had}}$ and for the invariant mass of the reconstructed top anti-top quark pair $M_{\mathrm{t\bar{t}}}$ are shown in Fig. 6.9, for the rapidity of the hadronically decaying top quark $y_{\mathrm{t,had}}$ and for the cosine of the production angle of the hadronically decaying top quark $\cos\theta^{*}_{\mathrm{t,had}}$ in Fig. 6.10.

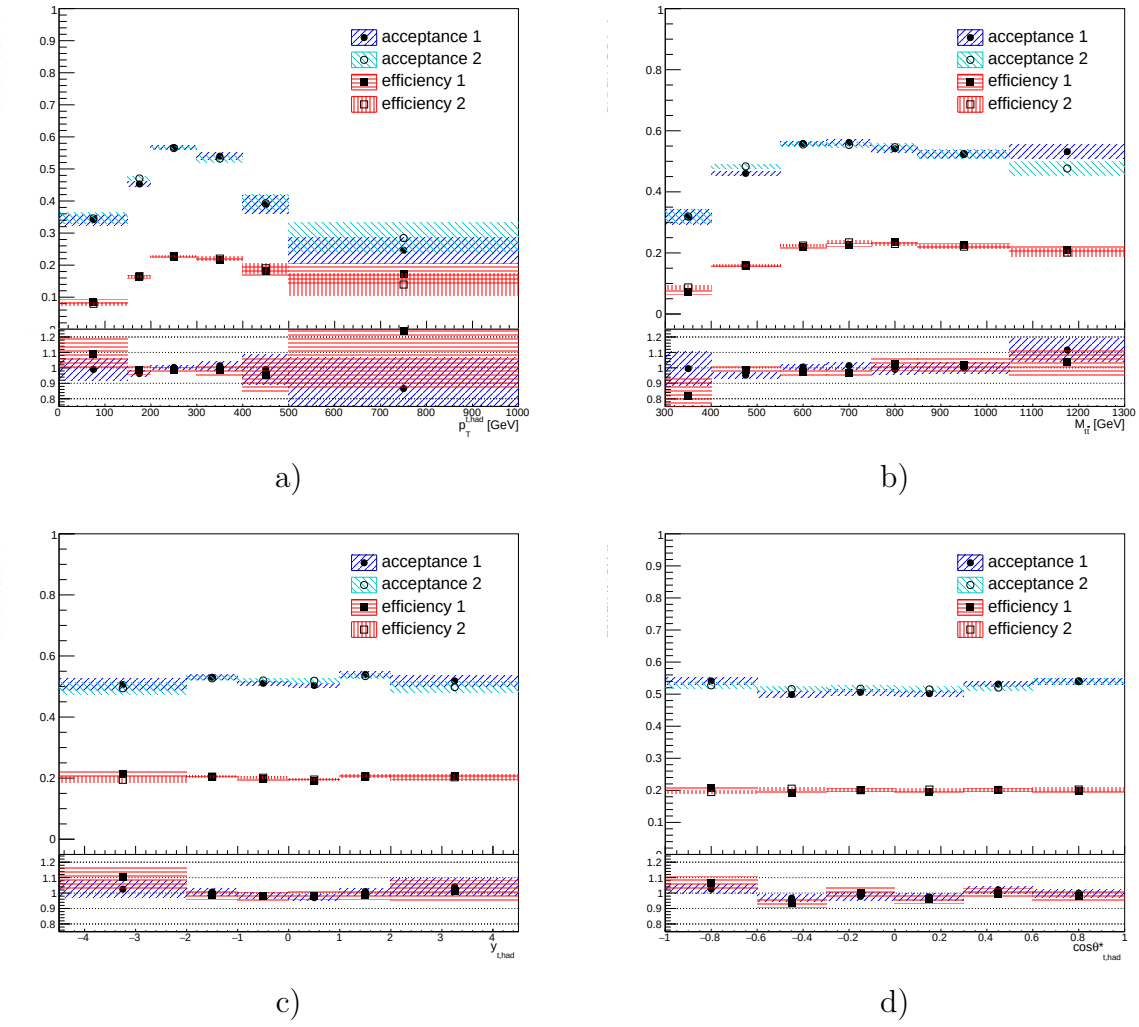


Figure 6.5: The acceptance and the efficiency for the two statistically independent $t\bar{t}$ samples for the reconstructed hadronically decaying top quark transverse momentum ($p_{\mathrm{T}}^{\mathrm{t,had}}$, a)), the invariant mass of the reconstructed $t\bar{t}$ system ($M_{\mathrm{t\bar{t}}}$, b)), for the rapidity of the hadronically decaying top quark ($y_{\mathrm{t,had}}$, c)) and for the cosine of the production angle of the hadronically decaying top quark ($\cos\theta^{*}_{\mathrm{t,had}}$, d)), all in the semi-boosted topology. Indices 1 and 2 denote the two statistically independent samples.

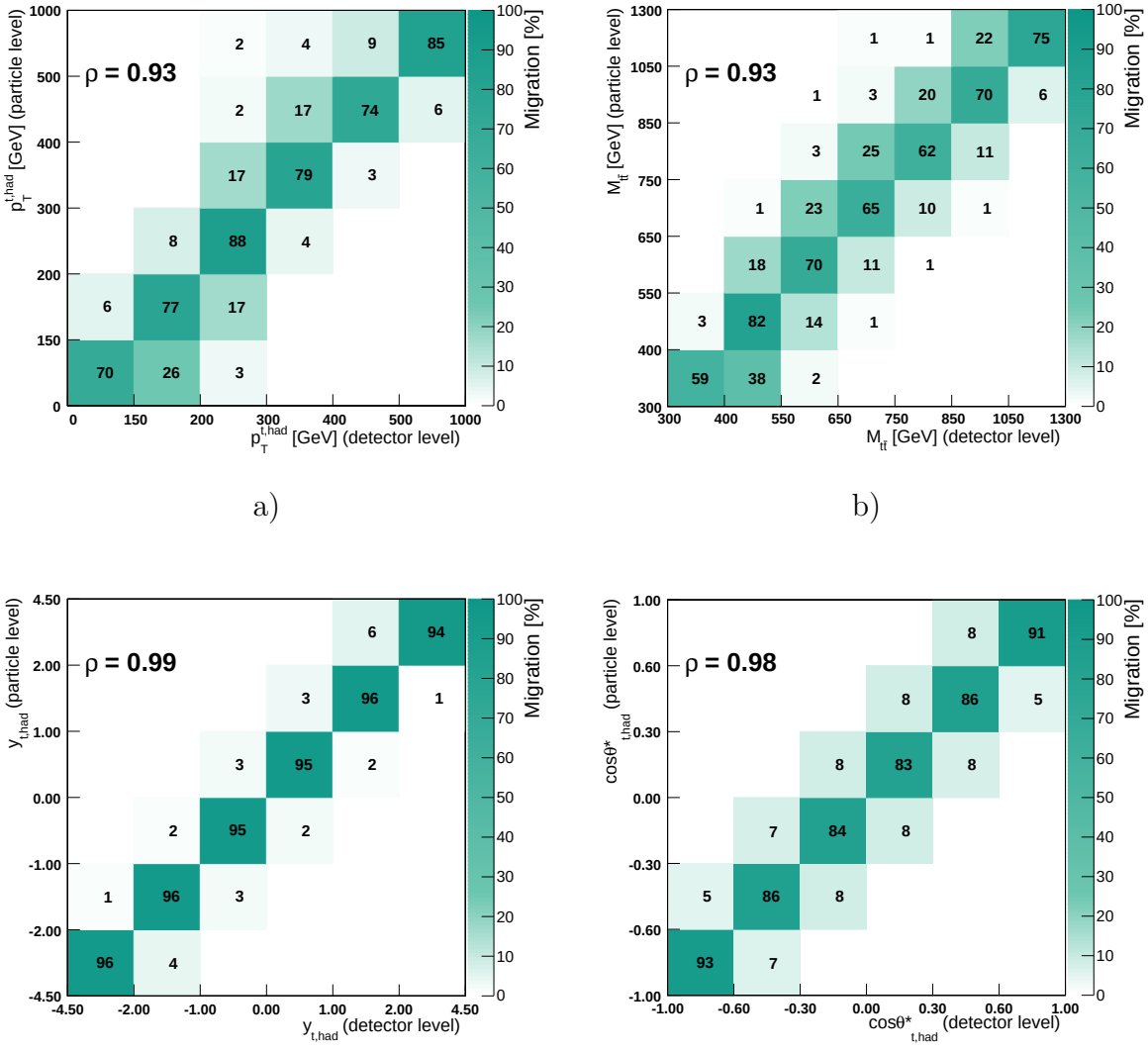


Figure 6.6: The migration matrices for the reconstructed transverse momentum of the hadronically decaying top quark ($p_{\mathrm{T}}^{\mathrm{t,had}}$, a)), invariant mass of $t\bar{t}$ system ($M_{\mathrm{t\bar{t}}}$, b)), for the rapidity of the hadronically decaying top quark $y_{\mathrm{t,had}}$, c)) and for the cosine of the production angle of the hadronically decaying top quark ($\cos\theta^*_{\mathrm{t,had}}$, d)), all in the semi-boosted topology. All matrices were derived from the $t\bar{t}$ sample and used in the unfolding procedure. The correlation factor ρ is calculated between the detector and particle levels.

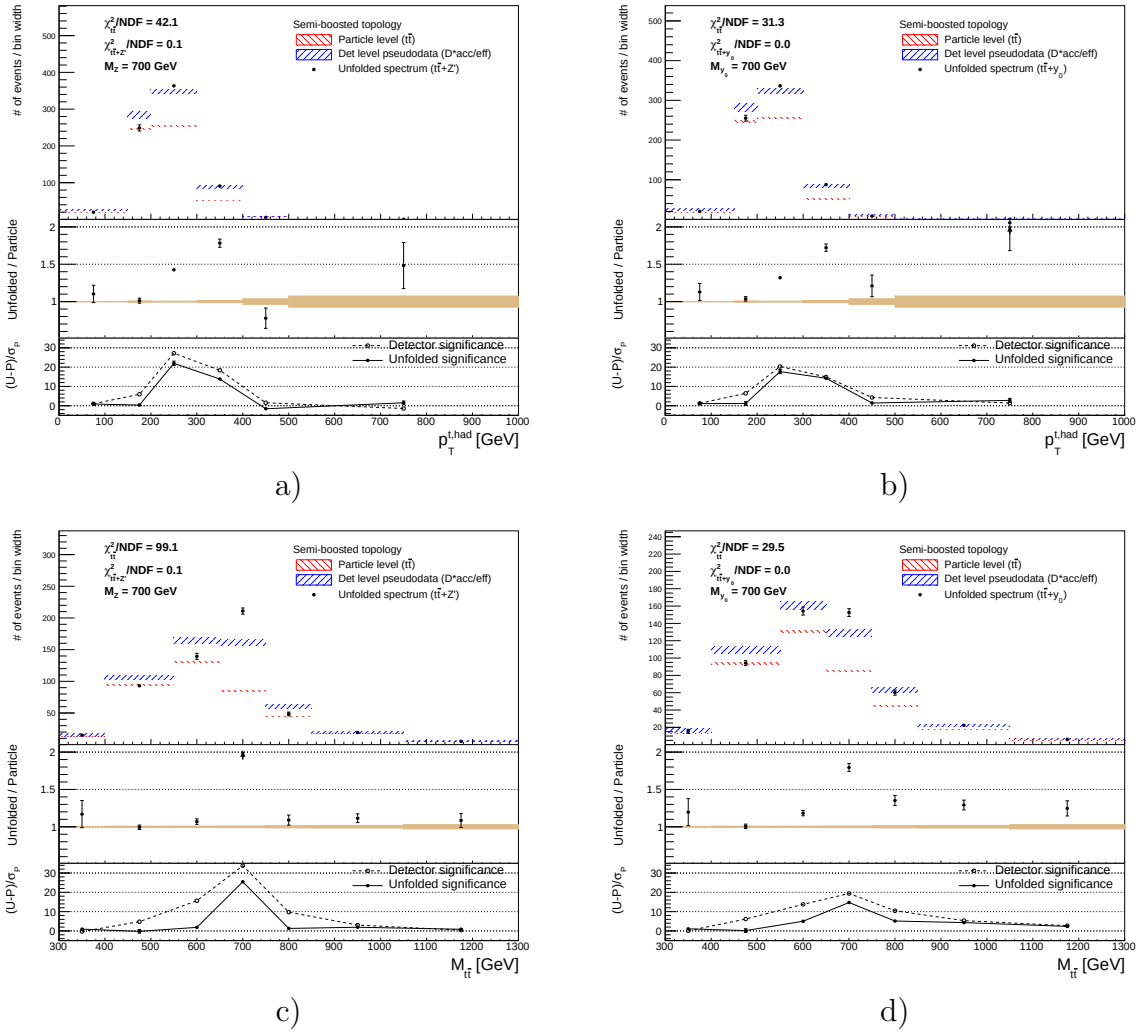


Figure 6.7: Unfolding results for the transverse momentum of the hadronically decaying top quark $p_{\mathrm{T}}^{\mathrm{t,had}}$ (a), b)) and for the invariant mass of the reconstructed top anti-top quark pair $M_{\mathrm{t\bar{t}}}$ (c), d)) for samples with mass of the hypothetical particle $M_{\mathrm{Z'}} = 700$ GeV (a), c)) and $M_{\mathrm{y0}} = 700$ GeV (b), d)) all in the semi-boosted topology. The top part of each plot shows results unfolding on the admixture of spectra (black full markers) compared with the particle level spectrum from statistically independent sample (red) and with detector level spectrum from $t\bar{t}$ sample (blue) corrected by the acceptance and efficiency factors to achieve similar phase space. The hatched area corresponds to the statistical uncertainty of the spectrum. The $\chi^2_{\mathrm{t\bar{t}}}$ test is performed between $t\bar{t}$ particle level spectrum and unfolded spectrum while $\chi^2_{\mathrm{t\bar{t}+Z'}}$ and $\chi^2_{\mathrm{t\bar{t}+y_0}}$ are performed between unfolded and $t\bar{t}$ particle level spectrum with respective signal particle level addition (closure test). The middle part of each plot shows the ratio between the unfolded spectrum and the particle level one from the statistically independent $t\bar{t}$ sample only. The yellow bar area shows the statistical uncertainty of the particle level spectra of the $t\bar{t}$ sample. The bottom part of the plot shows comparison between the unfolded (full markers, solid line) and the detector level (open markers, dashed line) significances.

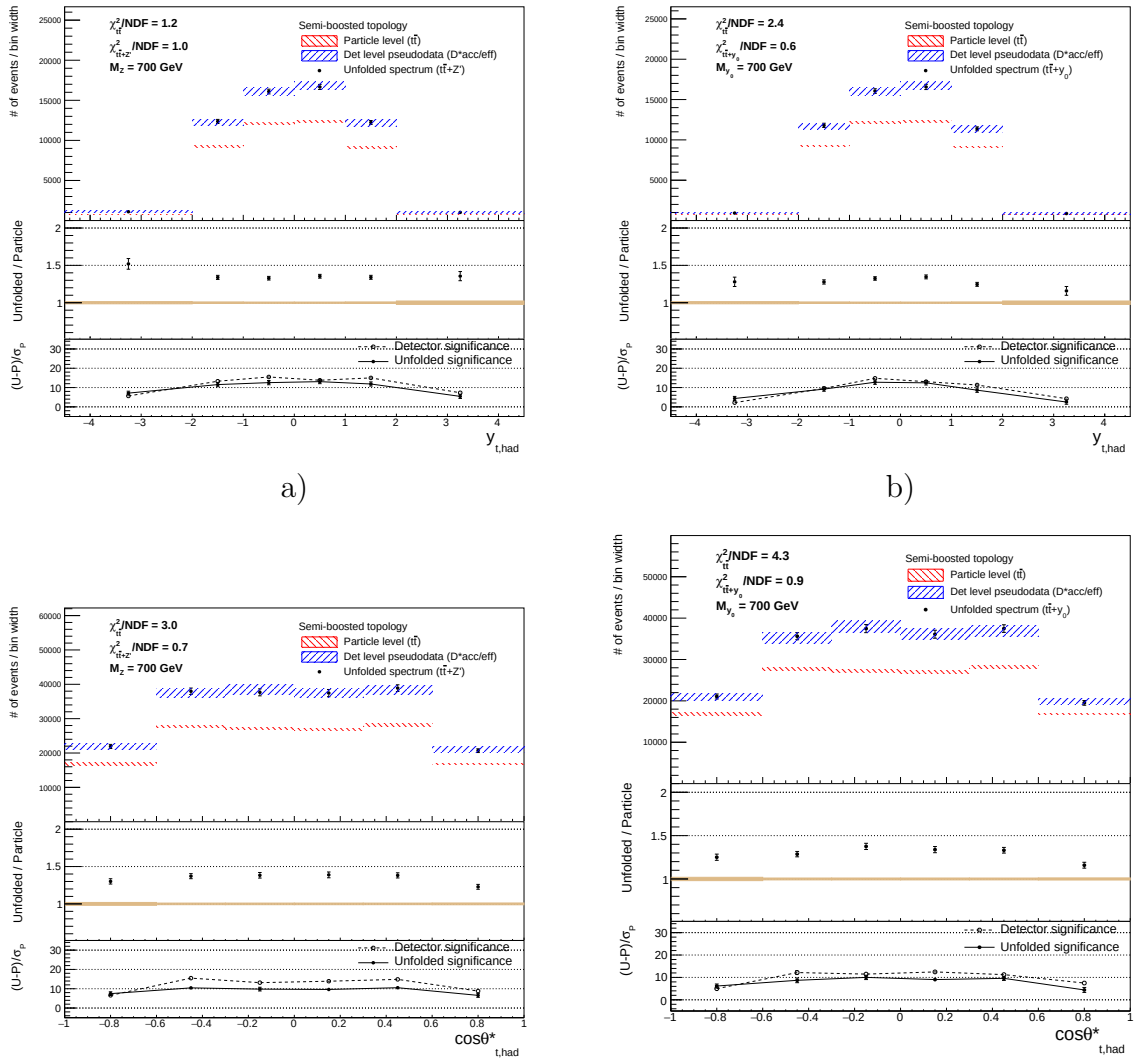


Figure 6.8: Unfolding results for the rapidity of the hadronically decaying top quark $y_{\mathrm{t,had}}$ (a), b)) and for the cosine of the production angle of the hadronically decaying top quark $\cos\theta^{*}_{\mathrm{t,had}}$ (c), d)) for samples with mass of the hypothetical particle $M_{\mathrm{Z'}} = 700$ GeV (a), c)) and $M_{\mathrm{y_0}} = 700$ GeV (b), d)) all in the semi-boosted topology. The top part of each plot shows unfolding results on the admixture of spectra (black full markers) compared with the particle level spectrum from statistically independent sample (red) and with detector level spectrum from $t\bar{t}$ sample (blue) corrected by the acceptance and efficiency factors to achieve similar phase space. The hatched area corresponds to the statistical uncertainty of the spectrum. The $\chi^2_{\mathrm{t\bar{t}}}$ test is performed between $t\bar{t}$ particle level spectrum and unfolded spectrum while $\chi^2_{\mathrm{t\bar{t}+Z'}}$ and $\chi^2_{\mathrm{t\bar{t}+y_0}}$ are performed between unfolded and $t\bar{t}$ particle level spectrum with respective signal particle level addition (closure test). The middle part of each plot shows the ratio between the unfolded spectrum and the particle level spectrum one from the statistically independent $t\bar{t}$ sample only. The yellow bar area shows the statistical uncertainty of the particle level spectra of the $t\bar{t}$ sample. The bottom part of the plot shows comparison between the unfolded (full markers, solid line) and the detector level (open markers, dashed line) significances.

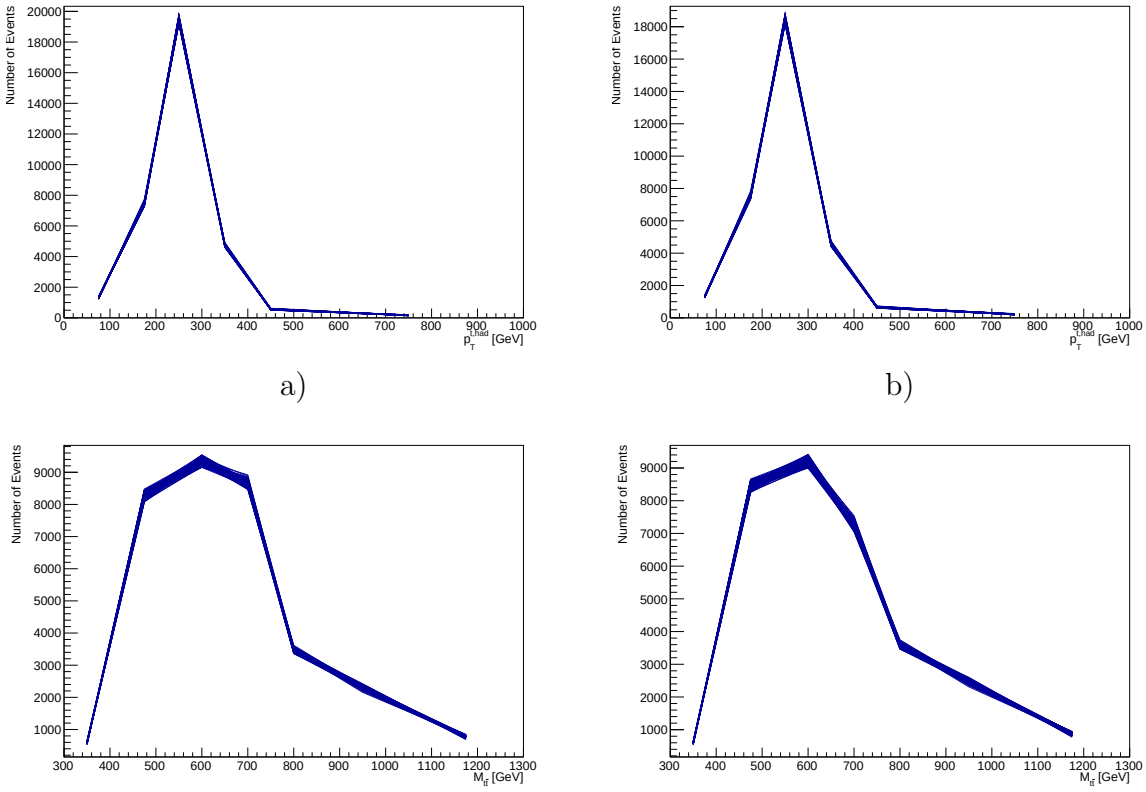


Figure 6.9: Spectra in pseudoexperiments for the evaluation of the unfolded significance statistical uncertainty for the transverse momentum of the hadronically decaying top quark $p_{\mathrm{T}}^{\mathrm{t,had}}$ (a), b)) and for the invariant mass of the reconstructed top anti-top quark pair $M_{\mathrm{t\bar{t}}}$ (c), d)) for samples with mass of the hypothetical particle $M_{\mathrm{Z'}} = 700$ GeV (a), c)) and $M_{\mathrm{y_0}} = 700$ GeV (b), d)), all in the semi-boosted topology.

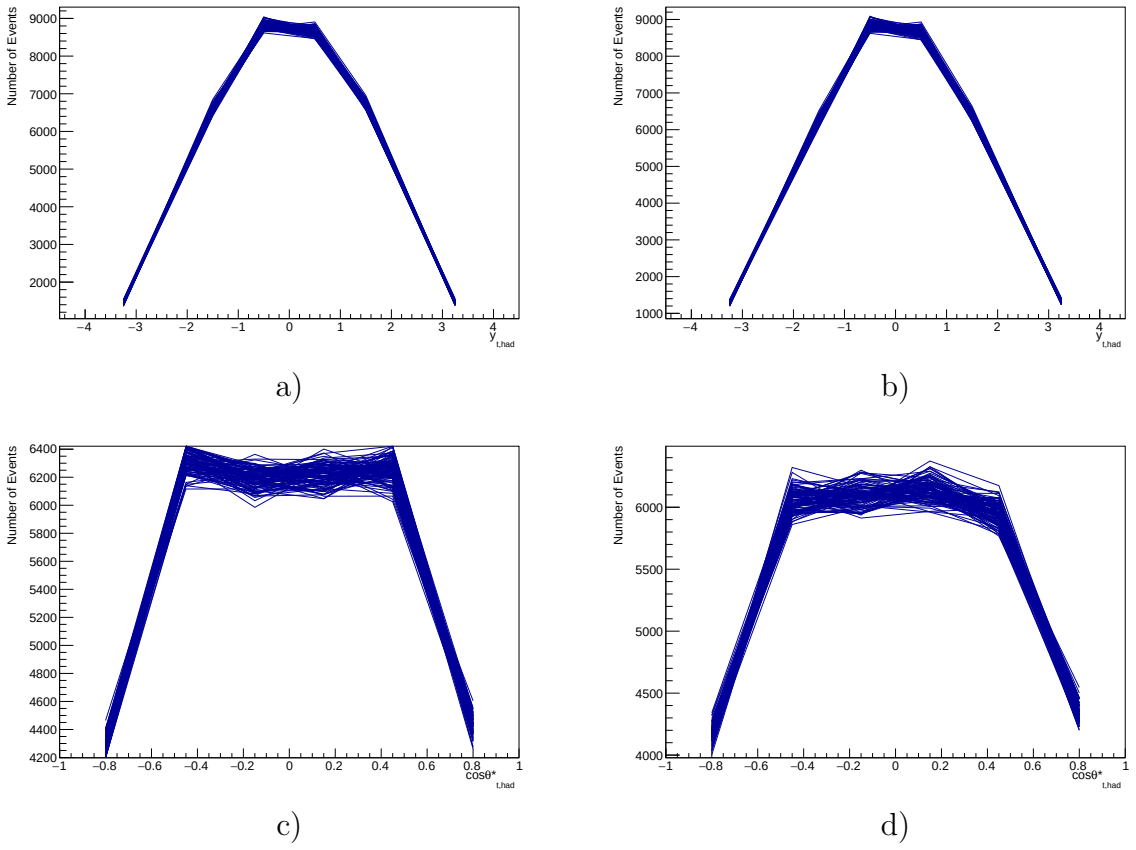


Figure 6.10: Spectra in pseudoexperiments for the evaluation of the unfolded significance statistical uncertainty for the rapidity of the hadronically decaying top quark $y_{\mathrm{t,had}}$ (a), b)), for the cosine of the production angle of the hadronically decaying top quark $\cos\theta^*_{\mathrm{t,had}}$ (c), d)) and for the invariant mass of the reconstructed top anti-top quark pair $M_{\mathrm{t\bar{t}}}$ (e), f)) for samples with mass of the hypothetical particle $M_{\mathrm{Z'}} = 700$ GeV (a), c), e)) and $M_{\mathrm{y_0}} = 700$ GeV (b), d), f)), all in the semi-boosted topology.

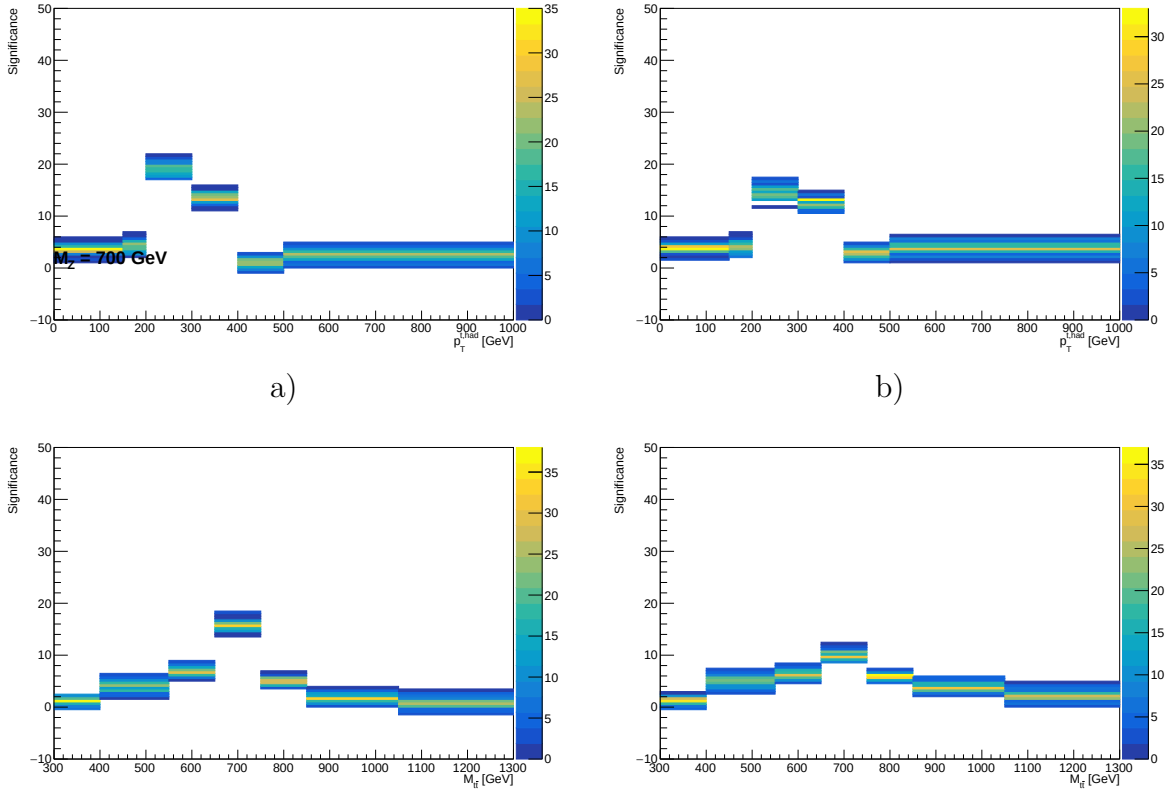


Figure 6.11: Depiction of significances of unfolded pseudoexperiments for the transverse momentum of the hadronically decaying top quark $p_{\mathrm{T}}^{\mathrm{t,had}}$ (a), b)) and for the invariant mass of the reconstructed top anti-top quark pair $M_{\mathrm{t\bar{t}}}$ (c), d)) for samples with mass of the hypothetical particle $M_{Z'} = 700$ GeV (a), c)) and $M_{\mathrm{y_0}} = 700$ GeV (b), d)), all in the semi-boosted topology.

Each pseudoexperiment results in a slightly different significance due to the smearing of the spectrum, thus, calculated significances for each pseudoexperiment differ. The distribution of significances in each bin is then fitted by a gaussian function and width of this fit is considered as statistical uncertainty of the significance in each bin. Distribution of the significances extracted from unfolded pseudoexperiments is shown in Fig. 6.11 for the transverse momentum of the hadronically decaying top quark $p_{\mathrm{T}}^{\mathrm{t,had}}$ and for the invariant mass of the reconstructed top anti-top quark pair $M_{\mathrm{t\bar{t}}}$ and in Fig 6.12 for the the rapidity of the hadronically decaying top quark $y_{\mathrm{t,had}}$ and cosine of the production angle of the hadronically decaying top quark $\cos\theta^*_{\mathrm{t,had}}$, all in the semi-boosted topology. An example of significances for samples with various hypothetical particles masses are presented in Appendix B.

6.5.1 Results on Significance

The signal significances in the $p_{\mathrm{T}}^{\mathrm{t,had}}$ spectrum peaks at different values of $p_{\mathrm{T}}^{\mathrm{t,had}}$ depending on the topology as the event selection in each topology biases the spectrum and effectively selects different ranges in $M_{\mathrm{t\bar{t}}}$, too, both shown in Fig 6.13. On the other hand, significance in the $M_{\mathrm{t\bar{t}}}$ spectrum peaks around the value of the generated y_0 mass or Z' mass of 700 GeV as expected, with a slight tail to lower values in the

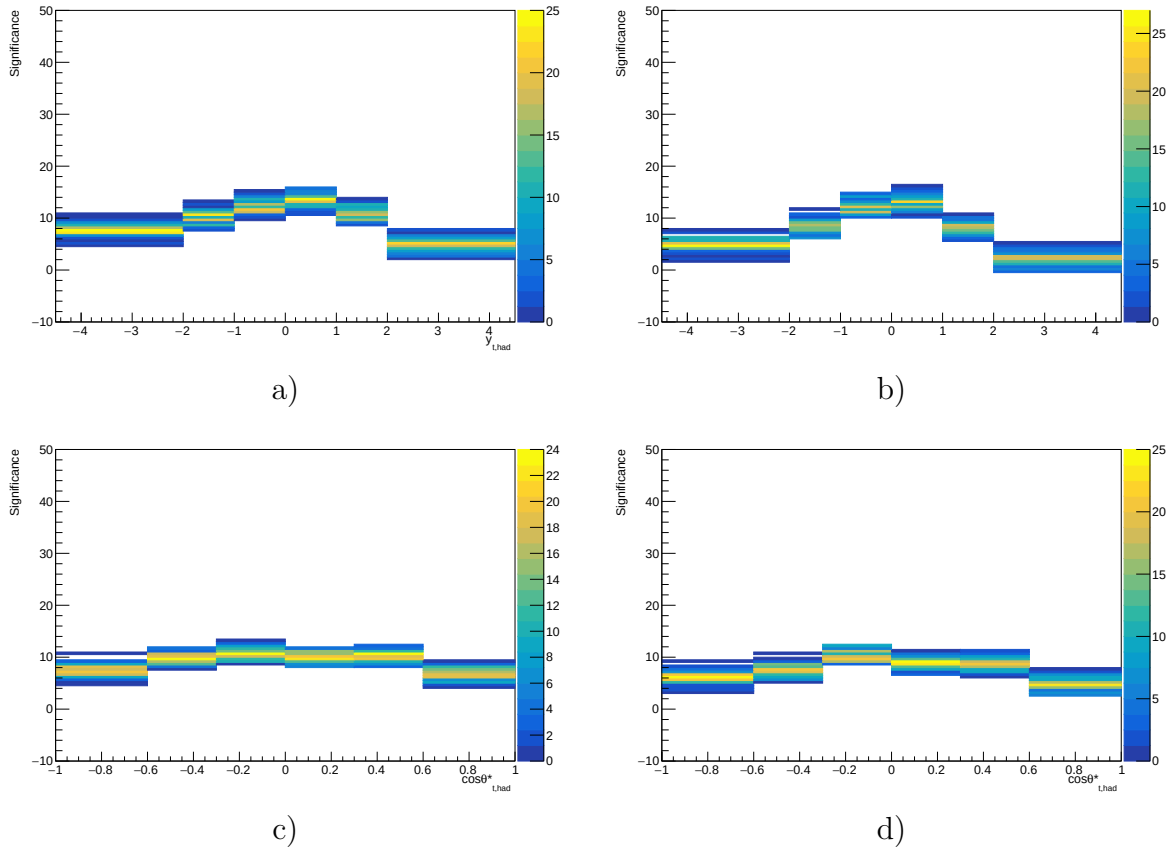


Figure 6.12: Depiction of significances of unfolded pseudoexperiments for the rapidity of the hadronically decaying top quark $y_{\mathrm{t,had}}$ (a), b)), for the cosine of the production angle of the hadronically decaying top quark $\cos\theta^*_{\mathrm{t,had}}$ (c), d)) and for the invariant mass of the reconstructed top anti-top quark pair $M_{\mathrm{t\bar{t}}}$ (e), f)) for samples with mass of the hypothetical particle $M_{Z'} = 700$ GeV (a), c), e)) and $M_{\mathrm{y_0}} = 700$ GeV (b), d), f)), all in the semi-boosted topology.

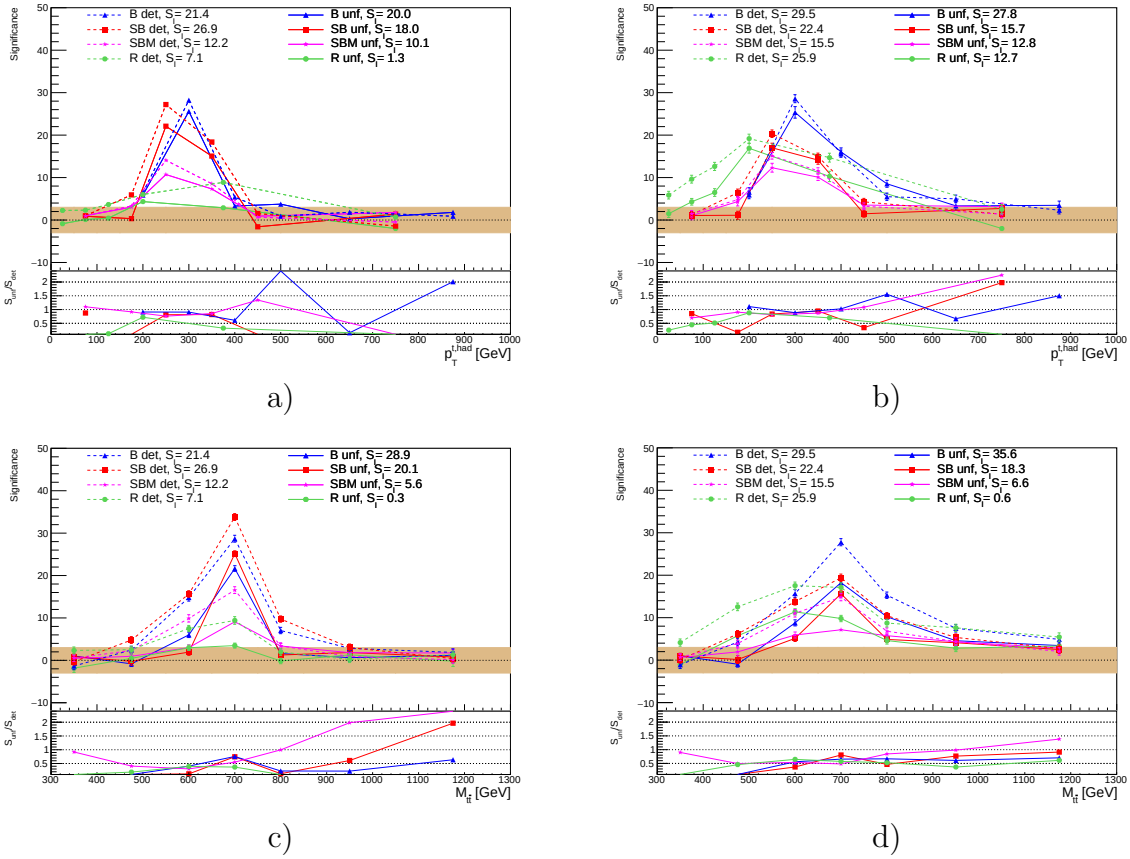


Figure 6.13: The detector (open markers, dashed line) and unfolded (full markers, solid line) significances for the transverse momentum of the hadronically decaying top quark ($p_{\mathrm{T}}^{\mathrm{t,had}}$, a), b)) and invariant mass of the reconstructed $t\bar{t}$ pair ($M_{\mathrm{t\bar{t}}}$, c), d)) plotted for all the topologies in each bin for samples with mass of the hypothetical particle $M_{\mathrm{Z'}} = 700$ GeV (a), c)) and $M_{\mathrm{y_0}} = 700$ GeV (b), d)). The band defines the area where the absolute value of the significance is below three, corresponding to the 3-σ interval. The lower pads present the ratios of the unfolded over the detector level significances, without uncertainties which are highly correlated.

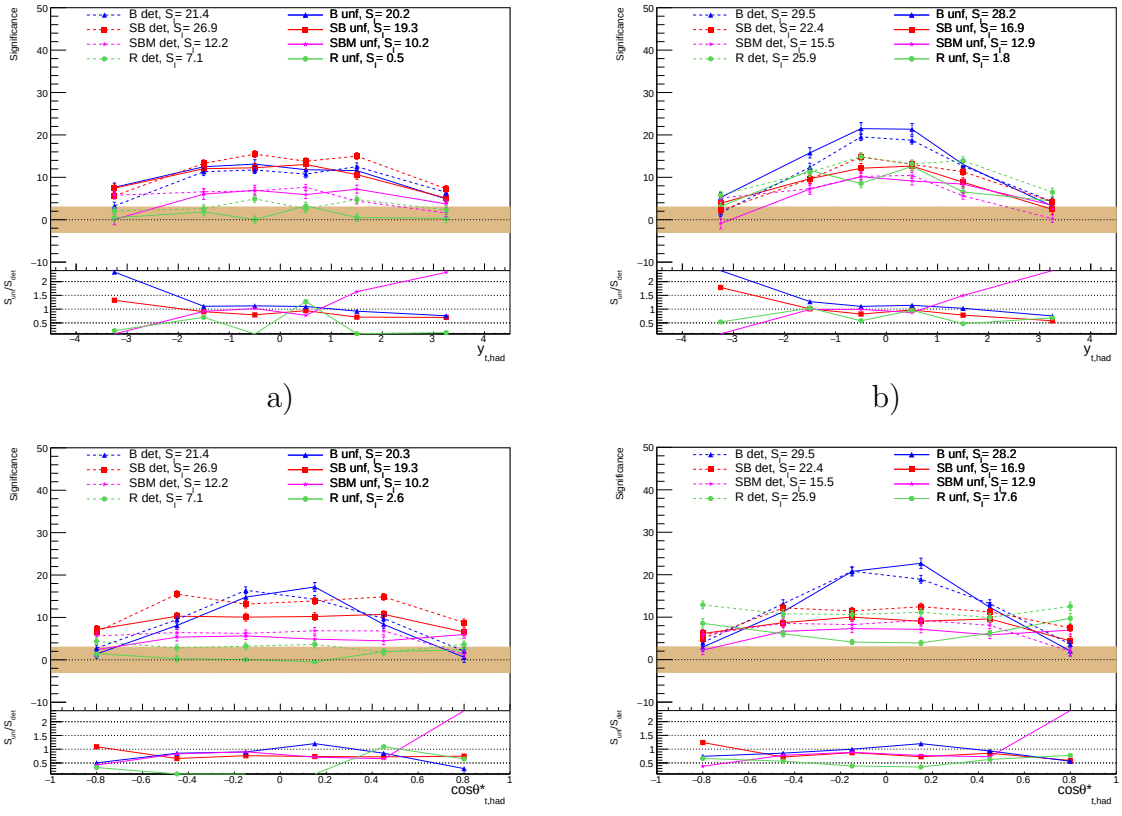


Figure 6.14: The detector (open markers, dashed line) and unfolded (full markers, solid line) significances for the rapidity of the hadronically decaying top quark ($y_{\mathrm{t,had}}$, a), b)) and the cosine of the production angle of the hadronically decaying top quark ($\cos\theta^{*}_{\mathrm{t,had}}$, c), d)) plotted for all the topologies in each bin for samples with mass of the hypothetical particle $M_{\mathrm{Z'}} = 700$ GeV (a), c)) and $M_{\mathrm{y_0}} = 700$ GeV (b), d)). The band defines the area where the absolute value of the significance is below three, corresponding to the 3-σ interval. The lower pads present the ratios of the unfolded over the detector level significances, without uncertainties which are highly correlated.

resolved topology which is the least suitable one to reconstruct a resonance of such a large mass. In contrast, the $y_{\mathrm{t,had}}$ and the $\cos\theta^*_{\mathrm{t,had}}$ signal significance is very flat also for the signal sample and there is no clear isolated excess of signal events in this spectrum, with the exception of the boosted topology which selects, by construction, high-p_{T} large jets and thus also top quarks, consequently more localized in the central rapidity region, producing a peak around zero, both shown in Fig 6.14.

The four selected spectra are thus good candidate observables to illustrate different behavior and spread of significances over bins, also presenting a selection of observables of a dimension of energy as well as dimensionless (angular). The binned significances are in general lower after the unfolding, for which an explicit proof is delivered by this study. The cause of this is as follows.

While a sharper spectrum may be recovered by unfolding, the procedure in general correlates information among bins by maximizing a likelihood function in case of FBU, or minimizing (possibly modified and regularized) χ^2 or iterating and sequentially improving the result for the case of other methods. An increase in the correlation across bins of the unfolded spectrum is a known and important fact and a correlation matrix should preferably be published along with unfolded spectra from real experiments, as done *e.g.* in [106, 107] where a correlation matrix between the observables was also evaluated. We observe that in case of the FBU method the posteriors variance usually increases, leading to larger absolute as well as relative uncertainties of the unfolded spectrum w.r.t. the particle level one. This effectively decreases the significance of the observed signal excess. An increase of the statistical uncertainties with the number of iterations in case of the Iterative Bayesian Unfolding [105] was also reported by other analyses [108]. We note that in our case the BSM signal significances decrease by 20–40%.

Conclusions

The results of the $t\bar{t}$ spectra shapes, their resolution and hypothetical BSM signal significance studies in the ℓ + jets channel on simulated signal and background samples were presented in this work. The selection criteria are chosen as close to the real analyses as possible corresponding to the ATLAS top working group analysis, thus, the mutual comparison is possible. The addition of the semi-boosted and semi-boosted mixed topology enhances yields between 20% and 50% in the $t\bar{t}$ mass range from 500 GeV to 1000 GeV. The resolution of the reconstructed $t\bar{t}$ mass peak at the detector level in the semi-boosted topology is comparable with the resolution for boosted and resolved topology. The semi-boosted mixed topology has slightly worse resolution, roughly by 50%. The results of the unfolding procedure on the $t\bar{t}$ sample corresponds well to the simulated particle level in closure tests as well as with the addition of the weighted background samples and hypothetical Z' and y_0 particles signal samples. The strength of the signal quantified by the significance is well visible after the unfolding in the admixture of the $t\bar{t}$, background and the hypothetical particle sample. However, the significance is diminished by the unfolding procedure for spectra of transverse momentum of hadronically decaying top quark, reconstructed top anti-top quark mass and production angle with the exception for angular variables in the boosted topology. Namely, the cosine of the production angle and the rapidity show only a small gap between the detector and unfolded significances for both hypothetical particles. This phenomenon is not visible for other topologies in samples, where these topologies have dominant fraction of events, for example in the sample with $m_{Z'} = 700$ GeV for the semi-boosted topology. The explanation lies in the ingredients of the unfolding procedure, where the migration matrix of the angular variables is diagonal and the events does not migrate in large numbers between the bins. The width of the posterior in unfolding in the studied bin is then smaller, which leads to the lesser diminishing of the signal. Unfolding procedure also strengthens the peak in spectra and combination of those two factors may lead to the higher significance at the unfolded level for the cosine of the production angle. This quantification of the significance of the signal before and after unfolding is, to our knowledge, fully demonstrated for the first time in this study. The unfolded spectra results still show sensitivity to the presence of the possible BSM signal is the $t\bar{t}$ spectrum.

The author exploited experience gained during the qualification task focused on the JES corrections preparation, which leaded to the internal note on which the article [92] was based, and from contribution to the analysis in measurement of the normalized differential cross-section of the $t\bar{t}$ pair in ℓ + jets channel, which was later published [47]. The private analysis results were also published, but only for the scalar hypothetical particle y_0 [109].

Acknowledgments

The author gratefully acknowledges the support from following projects: Research Infrastructure for experiments at CERN, LM2018104, Getting new knowledge of the microworld using the CERN infrastructure, MSMT CERN TRANSFER_UP LTT17018, Czech Science Foundation project GAČR 19-21484S and Palacky University scholarship program Student Mobility - Traineeship.

Bibliography

[1] Armin Hermann, Lanfranco Belloni, Gerhard John Krige, Ulrike Mersits, and Dominique Pestre. *History of CERN*. North-Holland, Amsterdam, 1987.

[2] Franco Bonaudi. The cern sc machine. *Physics Reports*, 225(1):31–43, 1993.

[3] Giuseppe Fidecaro. The Discoveries of Rare Pion Decays at the CERN Synchrocyclotron. *Adv. Ser. Dir. High Energy Phys.*, 23:397–414, 2015.

[4] Luigi Di Lella and Carlo Rubbia. The Discovery of the W and Z Particles. *Adv. Ser. Dir. High Energy Phys.*, 23:137–163, 2015.

[5] Ewa Lopienska. The CERN accelerator complex, layout in 2022. Complexe des accélérateurs du CERN en janvier 2022. 2022. General Photo.

[6] Oliver Sim Brüning et al. *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004.

[7] I. Zurbano Fernandez et al. High-Luminosity Large Hadron Collider (HL-LHC): Technical design report. 10/2020, 12 2020.

[8] Joao Pequenao and Paul Schaffner. How ATLAS detects particles: diagram of particle paths in the detector. Jan 2013.

[9] Joao Pequenao. Computer generated image of the ATLAS inner detector. Mar 2008.

[10] Heinz Pernegger. The Pixel Detector of the ATLAS Experiment for LHC Run-2. Technical Report 06, CERN, Geneva, 2015.

[11] Joao Pequenao. Computer generated images of the Pixel, part of the ATLAS inner detector. 2008.

[12] Georges Aad et al. Operation and performance of the ATLAS semiconductor tracker in LHC Run 2. *JINST*, 17(01):P01013, 2022.

[13] A. Boldyrev et al. The atlas transition radiation tracker. *Instruments and Experimental Techniques*, 55, 05 2012.

[14] Devin James Mahon and ATLAS LAr Collaboration. ATLAS LAr Calorimeter Performance in LHC Run-2. Technical Report 06, CERN, Geneva, 2020.

[15] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 1996.

[16] G. Aad et al. Readiness of the ATLAS Tile Calorimeter for LHC collisions. *Eur. Phys. J. C*, 70:1193–1236, 2010.

[17] Experiment ATLAS. An instrumentation drawer of the barrel tile calorimeter of the ATLAS experiment. General Photo, Apr 2014.

[18] L Pontecorvo. The ATLAS Muon Spectrometer. 2004. revised version number 1 submitted on 2003-07-27 16:31:16.

[19] Joao Pequenao. Computer generated image of the ATLAS Muons subsystem. Mar 2008.

[20] Gregorio Bernardi, T. R. Junk, and M. Carena. Higgs bosons, theory and searches. 2008.

[21] Georges Aad et al. Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV. *JHEP*, 08:045, 2016.

[22] A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery. *Nature*, 607(7917):52–59, 2022.

[23] C. N. Yang and R. L. Mills. Conservation of isotopic spin and isotopic gauge invariance. *Phys. Rev.*, 96:191–195, Oct 1954.

[24] Sheldon L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579–588, 1961.

[25] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.

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

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

[28] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585–587, Nov 1964.

[29] G. Zweig. *An SU(3) model for strong interaction symmetry and its breaking. Version 2*, pages 22–101. 2 1964.

[30] M Gell-Mann. The eightfold way: A theory of strong interaction symmetry. 3 1961.

[31] Steven Weinberg. The Making of the Standard Model. *Eur. Phys. J. C*, 34:5–13, 2004.

[32] J. E. Augustin, A. M. Boyarski, M. Breidenbach, F. Bulos, J. T. Dakin, G. J. Feldman, G. E. Fischer, D. Fryberger, G. Hanson, B. Jean-Marie, R. R. Larsen, V. Lüth, H. L. Lynch, D. Lyon, C. C. Morehouse, J. M. Paterson, M. L. Perl, B. Richter, P. Rapidis, R. F. Schwitters, W. M. Tanenbaum, F. Vannucci, G. S. Abrams, D. Briggs, W. Chinowsky, C. E. Friedberg, G. Goldhaber, R. J. Hollebeek, J. A. Kadyk, B. Lulu, F. Pierre, G. H. Trilling, J. S. Whitaker, J. Wiss, and J. E. Zipse. Discovery of a narrow resonance in e^+e^- annihilation. *Phys. Rev. Lett.*, 33:1406–1408, Dec 1974.

[33] J. J. Aubert, U. Becker, P. J. Biggs, J. Burger, M. Chen, G. Everhart, P. Goldhagen, J. Leong, T. McCorriston, T. G. Rhoades, M. Rohde, Samuel C. C. Ting, Sau Lan Wu, and Y. Y. Lee. Experimental observation of a heavy particle j. *Phys. Rev. Lett.*, 33:1404–1406, Dec 1974.

[34] M. L. Perl, G. S. Abrams, A. M. Boyarski, M. Breidenbach, D. D. Briggs, F. Bulos, W. Chinowsky, J. T. Dakin, G. J. Feldman, C. E. Friedberg, D. Fryberger, G. Goldhaber, G. Hanson, F. B. Heile, B. Jean-Marie, J. A. Kadyk, R. R. Larsen, A. M. Litke, D. Lüke, B. A. Lulu, V. Lüth, D. Lyon, C. C. Morehouse, J. M. Paterson, F. M. Pierre, T. P. Pun, P. A. Rapidis, B. Richter, B. Sadoulet, R. F. Schwitters, W. Tanenbaum, G. H. Trilling, F. Vannucci, J. S. Whitaker, F. C. Winkelmann, and J. E. Wiss. Evidence for anomalous lepton production in $e^+ - e^-$ annihilation. *Phys. Rev. Lett.*, 35:1489–1492, Dec 1975.

[35] CDF experiment collaboration. Observation of top quark production in $\bar{\text{p}}$p collisions with the collider detector at fermilab. *Phys. Rev. Lett.*, 74:2626–2631, Apr 1995.

[36] D0 Experment Collaboration. Observation of the top quark. *Phys. Rev. Lett.*, 74:2632–2637, Apr 1995.

[37] G. Aad 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.

[38] S. Chatrchyan 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.

[39] Stephen P. Martin. A Supersymmetry primer. *Adv. Ser. Direct. High Energy Phys.*, 18:1–98, 1998.

[40] Katrin Becker, Melanie Becker, and John H. Schwarz. *String Theory and M-Theory: A Modern Introduction*. Cambridge University Press, 2006.

[41] Florian Goertz. Composite Higgs theory. *PoS*, ALPS2018:012, 2018.

[42] Wikimedia Commons. File:standard model of elementary particles.svg — wikimedia commons, the free media repository, 2022. [Online; accessed 18-December-2022].

[43] Combination of ATLAS and CMS results on the mass of the top quark using up to 4.9 inverse femtobarns of data. 2013.

[44] R. L. Workman and Others. Review of Particle Physics. *PTEP*, 2022:083C01, 2022.

[45] Giuseppe Degrassi, Stefano Di Vita, Joan Elias-Miró, José R. Espinosa, Gian F. Giudice, Gino Isidori, and Alessandro Strumia. Higgs mass and vacuum stability in the standard model at nnlo. *Journal of High Energy Physics*, 2012(8), 2012.

[46] S. Mandelstam. Determination of the pion-nucleon scattering amplitude from dispersion relations and unitarity. general theory. *Phys. Rev.*, 112:1344–1360, Nov 1958.

[47] Georges Aad et al. Measurements of top-quark pair differential cross-sections in the lepton+jets channel in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector. *Eur. Phys. J. C*, 76(10):538, 2016.

[48] M. Aaboud et al. Measurement of the top quark mass in the $tt \rightarrow$lepton + jets channel from $\sqrt{s} = 8$ TeV ATLAS data and combination with previous results. *The European Physical Journal C*, 79(4), mar 2019.

[49] V. Khachatryan et al. Measurement of the top quark mass using proton-proton data at $\sqrt{(}s) = 7$ and 8 tev. *Phys. Rev. D*, 93:072004, Apr 2016.

[50] A. M. Sirunyan et al. Measurement of the top quark mass with lepton+jets final states using pp collisions at $\sqrt{s} = 13$ tev. *The European Physical Journal C*, 78(11), nov 2018.

[51] A. M. Sirunyan et al. Measurement of the top quark mass in the all-jets final state at $\sqrt{s} = 13$ tev and combination with the lepton+jets channel. *The European Physical Journal C*, 79(4), apr 2019.

[52] M. Aaboud et al. Top-quark mass measurement in the all-hadronic $t\bar{t}$ decay channel at $\sqrt{s} = 8$ tev with the atlas detector. *Journal of High Energy Physics*, 2017(9), sep 2017.

[53] Tevatron Electroweak Working Group. Combination of cdf and d0 results on the mass of the top quark using up 9.7 fb^{-1} at the tevatron, 2016.

[54] Victor Mukhamedovich Abazov et al. Measurement of top quark polarization in $t\bar{t}$ lepton+jets final states. *Phys. Rev. D*, 95(1):011101, 2017.

[55] Victor Mukhamedovich Abazov et al. Measurement of Spin Correlation between Top and Antitop Quarks Produced in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Lett. B*, 757:199–206, 2016.

[56] David J. Mietlicki. Measurements of Top Quark Properties at the Tevatron. 4 2012.

[57] Morad Aaboud et al. Measurements of top-quark pair spin correlations in the $e\mu$ channel at $\sqrt{s} = 13$ TeV using pp collisions in the ATLAS detector. *Eur. Phys. J. C*, 80(8):754, 2020.

[58] Albert M Sirunyan et al. Measurement of the top quark polarization and $t\bar{t}$ spin correlations using dilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Phys. Rev. D*, 100(7):072002, 2019.

[59] Werner Bernreuther, Dennis Heisler, and Zong-Guo Si. A set of top quark spin correlation and polarization observables for the LHC: Standard Model predictions and new physics contributions. *JHEP*, 12:026, 2015.

[60] Morad Aaboud et al. Measurements of top quark spin observables in $t\bar{t}$ events using dilepton final states in $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector. *JHEP*, 03:113, 2017.

[61] T. Aaltonen et al. Evidence for d0-d0bar mixing using the cdf ii detector. *Physical Review Letters*, 100(12), mar 2008.

[62] T. Aaltonen et al. Observation of s-channel production of single top quarks at the tevatron. *Phys. Rev. Lett.*, 112:231803, Jun 2014.

[63] G. Aad et al. Comprehensive measurements of t-channel single top-quark production cross sections at $\sqrt{s} = 7$ tev with the atlas detector. *Physical Review D*, 90(11), dec 2014.

[64] S. Chatrchyan et al. Measurement of the single-top-quark t-channel cross section in pp collisions at $\sqrt{s} = 7$ tev. *Journal of High Energy Physics*, 2012(12), dec 2012.

[65] M. Aaboud et al. Fiducial, total and differential cross-section measurements of t-channel single top-quark production in pp collisions at $\sqrt{s} = 8$ TeV using data collected by the ATLAS detector. *The European Physical Journal C*, 77(8), aug 2017.

[66] V. Khachatryan et al. Measurement of the t-channel single-top-quark production cross section and of the $|vtb|$ ckm matrix element in pp collisions at $\sqrt{s} = 8$ tev. *Journal of High Energy Physics*, 2014(6), jun 2014.

[67] M. Aaboud et al. Measurement of the inclusive cross-sections of single top-quark and top-antiquark t-channel production in pp collisions at $\sqrt{s} = 13$ tev with the atlas detector. *Journal of High Energy Physics*, 2017(4), apr 2017.

[68] A.M. Sirunyan et al. Cross section measurement of t-channel single top quark production in pp collisions at $\sqrt{s} = 13$ tev. *Physics Letters B*, 772:752–776, sep 2017.

[69] Georges Aad et al. Evidence for single top-quark production in the s-channel in proton-proton collisions at $\sqrt{s}$ =8 TeV with the ATLAS detector using the Matrix Element Method. *Phys. Lett. B*, 756:228–246, 2016.

[70] Vardan Khachatryan et al. Search for s channel single top quark production in pp collisions at $\sqrt{s} = 7$ and 8 TeV. *JHEP*, 09:027, 2016.

[71] Georges Aad et al. Measurement of the production cross-section of a single top quark in association with a W boson at 8 TeV with the ATLAS experiment. *JHEP*, 01:064, 2016.

[72] Morad Aaboud et al. Measurement of the cross-section for producing a W boson in association with a single top quark in pp collisions at $\sqrt{s} = 13$ TeV with ATLAS. *JHEP*, 01:063, 2018.

[73] Albert M Sirunyan et al. Measurement of the production cross section for single top quarks in association with W bosons in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 10:117, 2018.

[74] Victor Mukhamedovich Abazov et al. Measurement of the Inclusive $t\bar{t}$ Production Cross Section in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV and Determination of the Top Quark Pole Mass. *Phys. Rev. D*, 94:092004, 2016.

[75] T. Aaltonen et al. Measurement of the top-quark pair production cross-section in events with two leptons and bottom-quark jets using the full CDF data set. *Phys. Rev. D*, 88:091103, 2013.

[76] Timo Antero Aaltonen et al. Combination of Measurements of the Top-Quark Pair Production Cross Section from the Tevatron Collider. *Phys. Rev. D*, 89(7):072001, 2014.

[77] Georges Aad et al. Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b-tagged jets in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS detector. *Eur. Phys. J. C*, 74(10):3109, 2014. [Addendum: Eur.Phys.J.C 76, 642 (2016)].

[78] Serguei Chatrchyan et al. Measurement of the $t\bar{t}$ Production Cross Section in the Dilepton Channel in pp Collisions at $\sqrt{s} = 7$ TeV. *JHEP*, 11:067, 2012.

[79] Combination of ATLAS and CMS top-quark pair cross section measurements using proton-proton collisions at sqrt(s) = 7 TeV. Technical report, CERN, Geneva, 2013.

[80] Serguei Chatrchyan et al. Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 8$ TeV. *JHEP*, 02:024, 2014. [Erratum: JHEP 02, 102 (2014)].

[81] Combination of ATLAS and CMS top quark pair cross section measurements in the $e\mu$ final state using proton-proton collisions at $\sqrt{s}$= 8 TeV. Technical report, CERN, Geneva, 2014. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2014-054.

[82] Roel Aaij et al. First observation of top quark production in the forward region. *Phys. Rev. Lett.*, 115(11):112001, 2015.

[83] Morad Aaboud et al. Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b-tagged jets in pp collisions at $\sqrt{s}$=13 TeV with the ATLAS detector. *Phys. Lett. B*, 761:136–157, 2016. [Erratum: Phys.Lett.B 772, 879–879 (2017)].

[84] Albert M Sirunyan et al. Measurement of the $\mathrm{t\bar{t}}$ production cross section, the top quark mass, and the strong coupling constant using dilepton events in pp collisions at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 79(5):368, 2019.

[85] A. M. Sirunyan et al. Measurement of the inclusive $\mathrm{t\bar{t}}$ cross section in pp collisions at $\sqrt{s} = 5.02$ TeV using final states with at least one charged lepton. *JHEP*, 03:115, 2018.

[86] M. Tanabashi et al. Review of Particle Physics. *Phys. Rev. D*, 98(3):030001, 2018.

[87] Georges Aad et al. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. *Eur. Phys. J. C*, 77:490, 2017.

[88] Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV. Technical report, CERN, Geneva, 2015. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-015.

[89] Morad Aaboud et al. Determination of jet calibration and energy resolution in proton-proton collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector. *Eur. Phys. J. C*, 80(12):1104, 2020.

[90] Morad Aaboud et al. Jet reconstruction and performance using particle flow with the ATLAS Detector. *Eur. Phys. J. C*, 77(7):466, 2017.

[91] Georges Aad et al. Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 81(8):689, 2021.

[92] M. Aaboud et al. Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. D*, 96(7):072002, 2017.

[93] Georges Aad et al. ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 79(11):970, 2019.

[94] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.

[95] Duhr C. FeynRules Implementation of Abelian Higgs Model. 2011. `https://feynrules.irmp.ucl.ac.be/wiki/HiddenAbelianHiggsModel`.

[96] Neil D. Christensen and Claude Duhr. FeynRules - Feynman rules made easy. *Comput. Phys. Commun.*, 180:1614–1641, 2009.

[97] James D. Wells. How to Find a Hidden World at the Large Hadron Collider. 2008.

[98] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi. DELPHES 3, A modular framework for fast simulation of a generic collider experiment. *JHEP*, 02:057, 2014.

[99] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet User Manual. *Eur. Phys. J.*, C72:1896, 2012.

[100] R. Brun and F. Rademakers. ROOT: An object oriented data analysis framework. *Nucl. Instrum. Meth. A*, 389:81–86, 1997.

[101] Benjamin Nachman, Pascal Nef, Ariel Schwartzman, Maximilian Swiatlowski, and Chaowaroj Wanotayaroj. Jets from Jets: Re-clustering as a tool for large radius jet reconstruction and grooming at the LHC. *JHEP*, 02:075, 2015.

[102] John Salvatier, Thomas V Wiecki, and Christopher Fonnesbeck. Probabilistic programming in python using pymc3. *PeerJ Computer Science*, 2:e55, 2016.

[103] D. Gerbaudo. Implementation of the Fully Bayesian Unfolding. 2019. `https://github.com/gerbaudo/fbu`.

[104] Andreas Höcker and Vakhtang Kartvelishvili. Svd approach to data unfolding. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 372(3):469–481, Apr 1996.

[105] G. D'Agostini. Improved iterative bayesian unfolding. `https://arxiv.org/abs/1010.0632`, 2010.

[106] Georges Aad et al. Measurements of normalized differential cross sections for $t\bar{t}$ production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector. *Phys. Rev. D*, 90(7):072004, 2014.

[107] M. Aaboud, G. Aad, B. Abbott, J. Abdallah, O. Abdinov, Baptiste Abeloos, S. Abidi, O. AbouZeid, Nicola Abraham, H. Abramowicz, Heather Abreu, Ricardo Abreu, Y. Abulaiti, Basanta Acharya, S. Adachi, L. Adamczyk, J. Adelman, Michael Adersberger, T. Adye, and L. Zwalinski. Measurements of top-quark pair differential cross-sections in the lepton+jets channel in *pp* collisions at $\sqrt{s} = 13$ tev using the atlas detector. *Journal of High Energy Physics*, 2017, 11 2017.

[108] Georges Aad et al. Differential top-antitop cross-section measurements as a function of observables constructed from final-state particles using pp collisions at $\sqrt{s} = 7$ TeV in the ATLAS detector. *JHEP*, 06:100, 2015.

[109] J. Pácalt and J. Kvita. Study of semi-boosted top quark reconstruction performance on the line shape of a $t\bar{t}$ resonance. `https://arxiv.org/abs/2106.01243`, 2021.

Appendices

Appendix A

Top Anti-top Quark Pair Invariant Mass Control Plots

The combined spectra of the reconstructed $t\bar{t}$ pair invariant mass ($m_{t\bar{t}}$) are shown in this appendix for samples with different hypothetical particles mass and different topologies are shown for the illustration of the private analysis performance on other samples. Spectra are weighted to the luminosity of the $t\bar{t}$ sample, see details in section 4.1, with additional signal scaling as described in section 6.2 and the signal presence corresponds to the fractions shown in Fig. 5.27. The spectra for the mass of the hypothetical particle $M_{Z'/y_0} = 500$ GeV are shown in Fig. A.1 for the resolved and semi-boosted mixed topology and in Fig. A.2 for the semi-boosted and boosted topology. Similarly, spectra for the mass $M_{Z'/y_0} = 600$ GeV are shown in Fig. A.3 and in Fig. A.4, for the mass $M_{Z'/y_0} = 700$ GeV in Fig. A.5 and in Fig. A.6, for the mass $M_{Z'/y_0} = 800$ GeV in Fig. A.7 and in Fig. A.8, for the mass $M_{Z'/y_0} = 900$ GeV in Fig. A.9 and in Fig. A.10 and for the mass $M_{Z'/y_0} = 1000$ GeV in Fig. A.11 and in Fig. A.12.

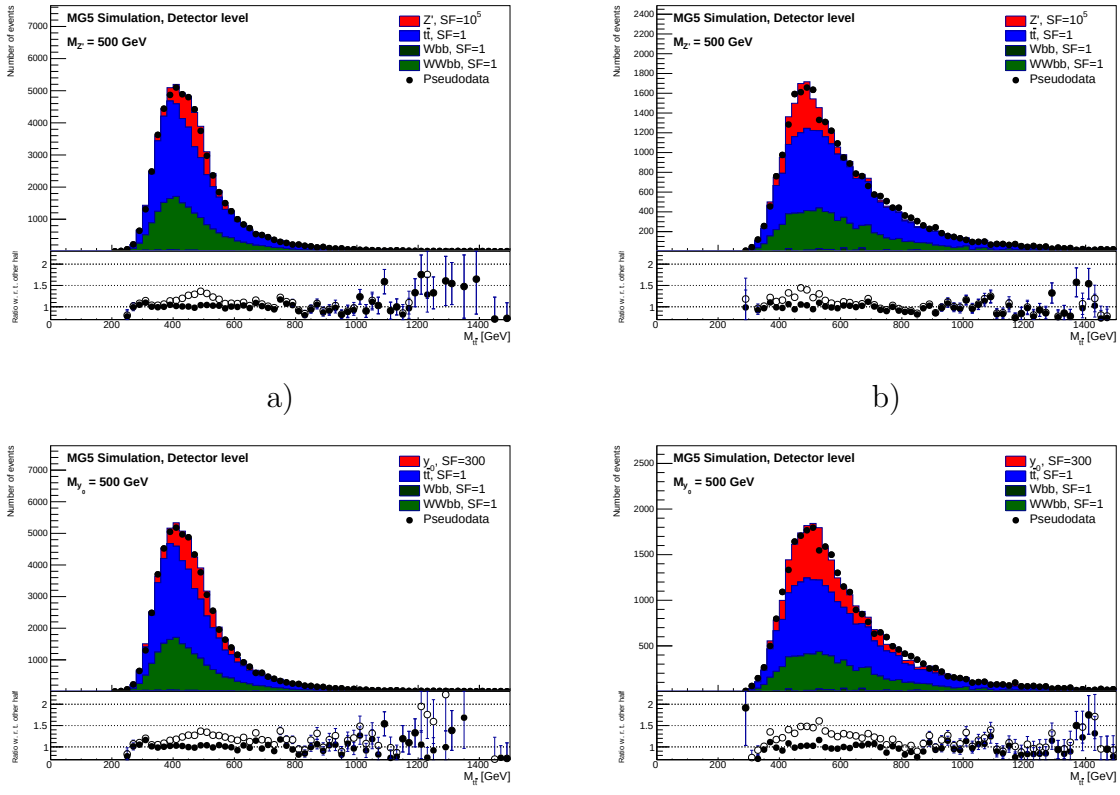


Figure A.1: Combined spectrum of the reconstructed invariant $t\bar{t}$ pair mass spectra ($m_{\mathrm{t\bar{t}}}$) at the detector level for resolved (a), c)) and semi-boosted mixed topology (b), d)) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{\mathrm{Z'}} = 500$ GeV (top row) and with $M_{\mathrm{y_0}} = 500$ GeV (bottom row), all samples are weighted on the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

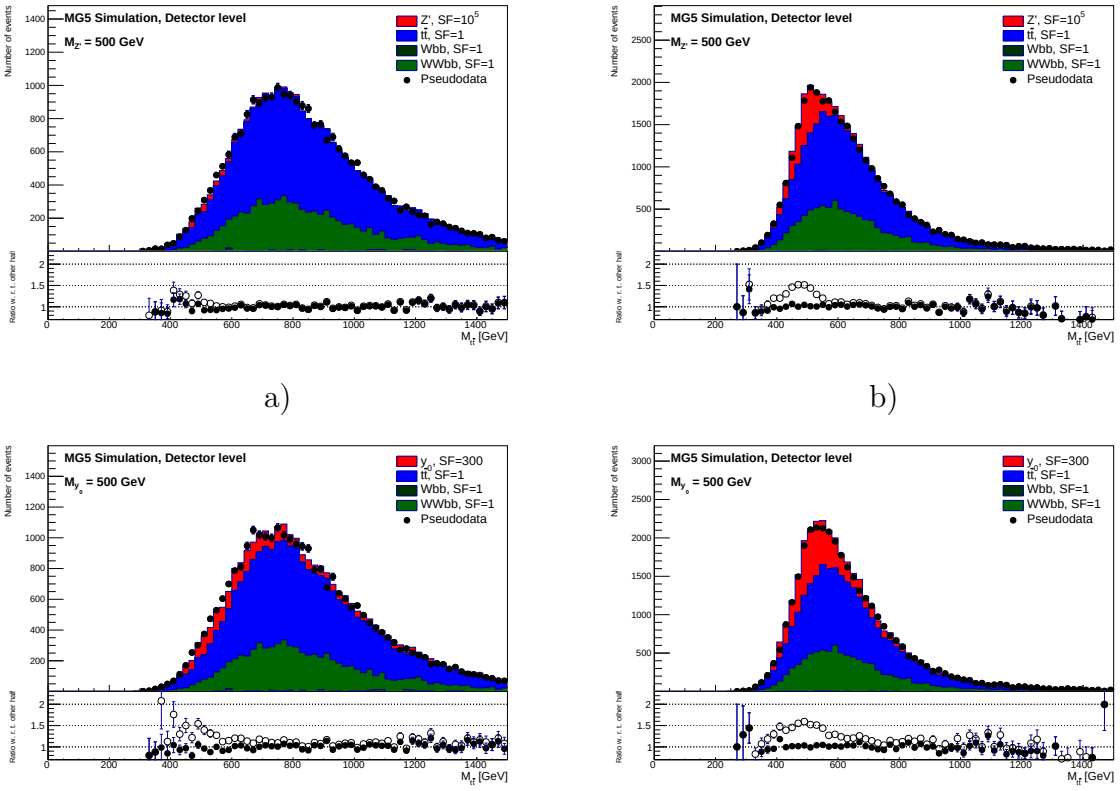


Figure A.2: Combined spectrum of the reconstructed invariant $t\bar{t}$ pair mass spectra ($m_{\mathrm{t\bar{t}}}$) at the detector level for semi-boosted (a), c)) and boosted topology (b), d)) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{\mathrm{Z'}} = 500$ GeV (top row) and with $M_{\mathrm{y_0}} = 500$ GeV (bottom row), all samples are weighted on the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

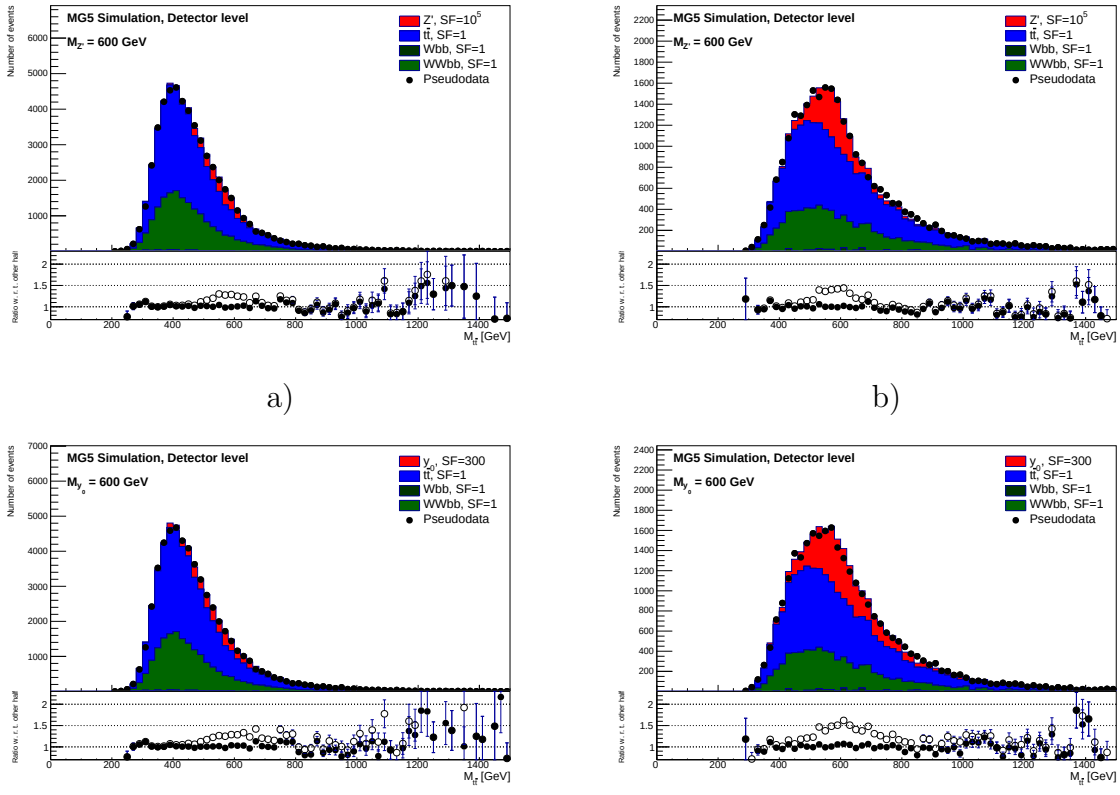


Figure A.3: Combined spectrum of the reconstructed invariant $t\bar{t}$ pair mass spectra ($m_{\mathrm{t\bar{t}}}$) at the detector level for resolved (a), c)) and semi-boosted mixed topology (b), d)) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{\mathrm{Z'}} = 600$ GeV (top row) and with $M_{\mathrm{y_0}} = 600$ GeV (bottom row), all samples are weighted on the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

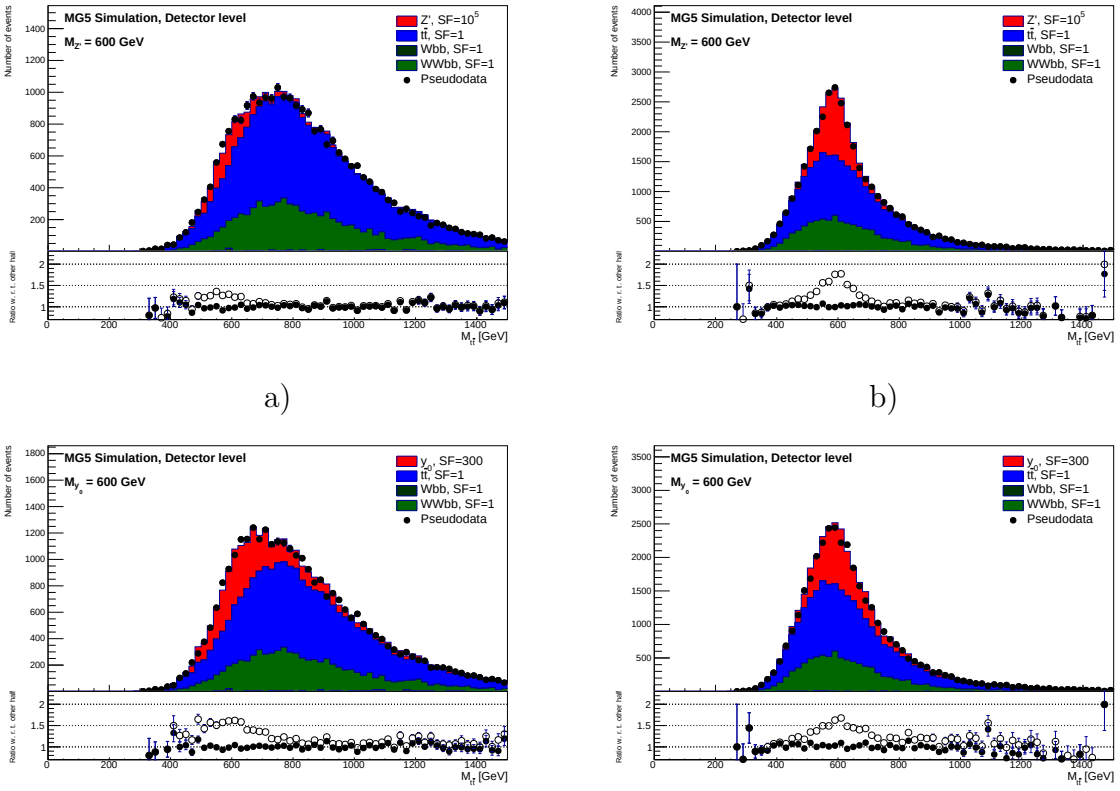


Figure A.4: Combined spectrum of the reconstructed invariant $t\bar{t}$ pair mass spectra ($m_{\mathrm{t\bar{t}}}$) at the detector level for semi-boosted (a), c)) and boosted topology (b), d)) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{\mathrm{Z'}} = 600$ GeV (top row) and with $M_{\mathrm{y_0}} = 600$ GeV (bottom row), all samples are weighted on the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

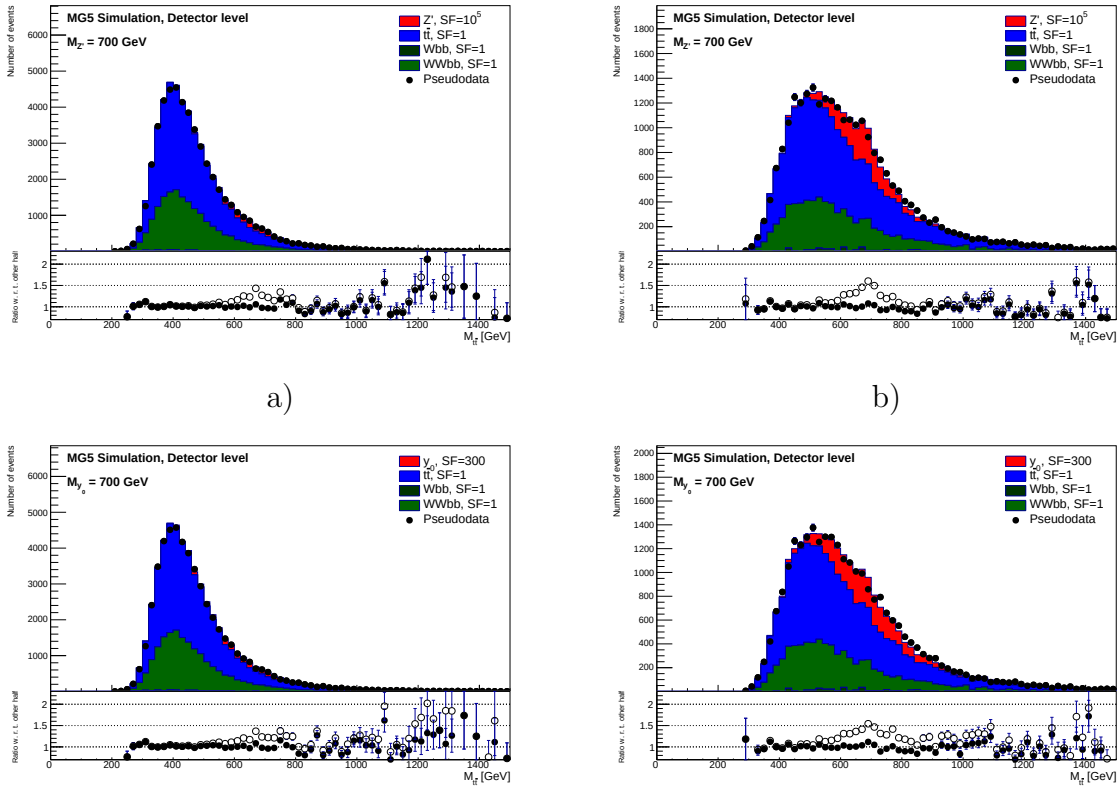


Figure A.5: Combined spectrum of the reconstructed invariant $t\bar{t}$ pair mass spectra ($m_{\mathrm{t\bar{t}}}$) at the detector level for resolved (a), c)) and semi-boosted mixed topology (b), d)) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{\mathrm{Z'}} = 700$ GeV (top row) and with $M_{\mathrm{y_0}} = 700$ GeV (bottom row), all samples are weighted on the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

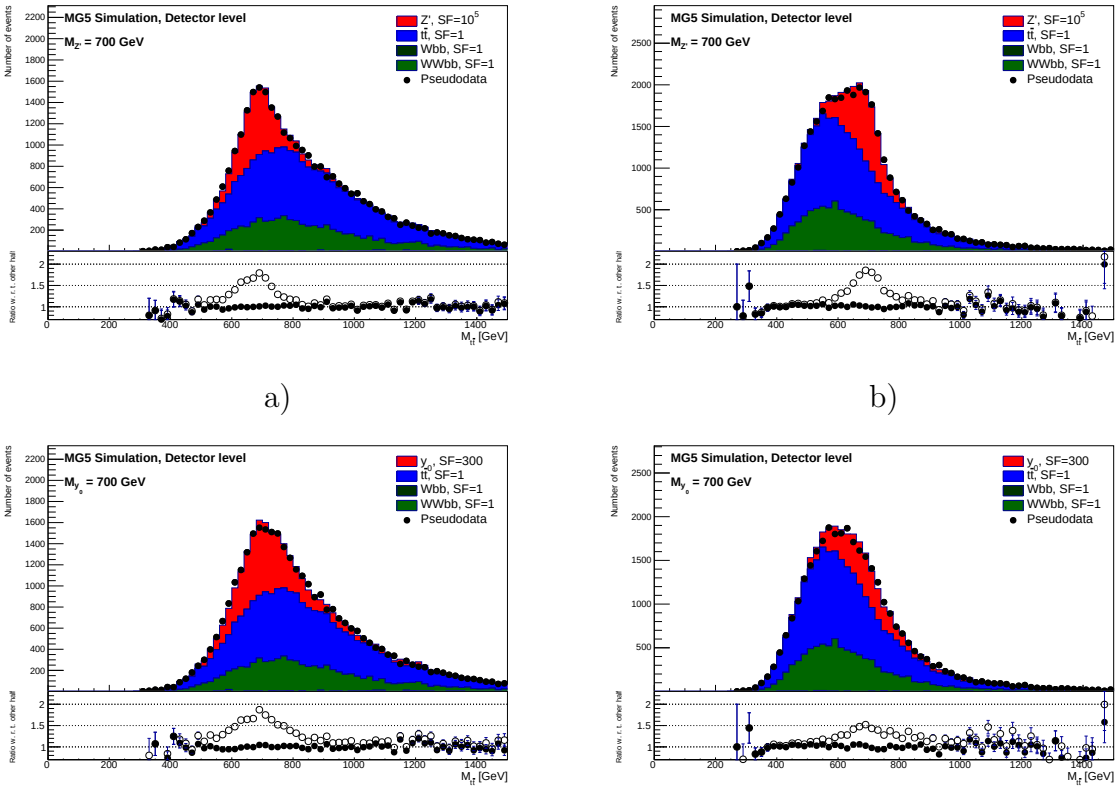


Figure A.6: Combined spectrum of the reconstructed invariant $t\bar{t}$ pair mass spectra ($m_{\mathrm{t\bar{t}}}$) at the detector level for semi-boosted (a), c)) and boosted topology (b), d)) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{\mathrm{Z'}} = 700$ GeV (top row) and with $M_{\mathrm{y_0}} = 700$ GeV (bottom row), all samples are weighted on the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

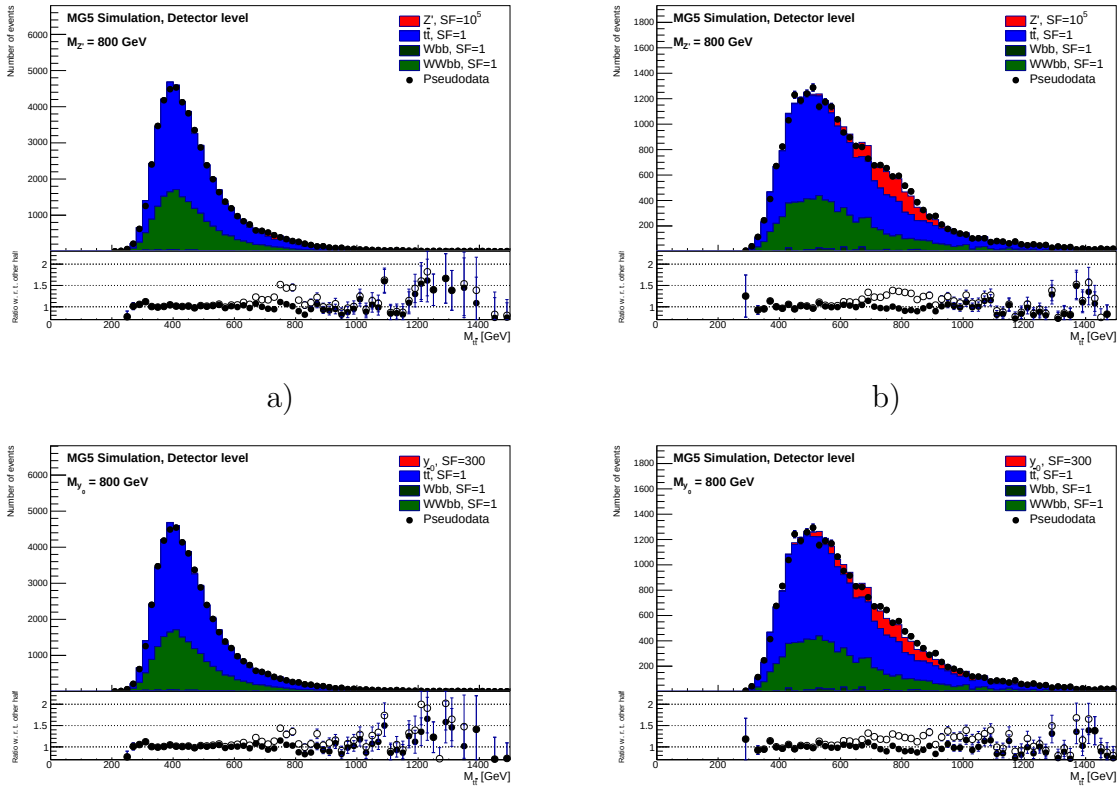


Figure A.7: Combined spectrum of the reconstructed invariant $t\bar{t}$ pair mass spectra ($m_{\mathrm{t\bar{t}}}$) at the detector level for resolved (a), c)) and semi-boosted mixed topology (b), d)) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{\mathrm{Z'}} = 800$ GeV (top row) and with $M_{\mathrm{y_0}} = 800$ GeV (bottom row), all samples are weighted on the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

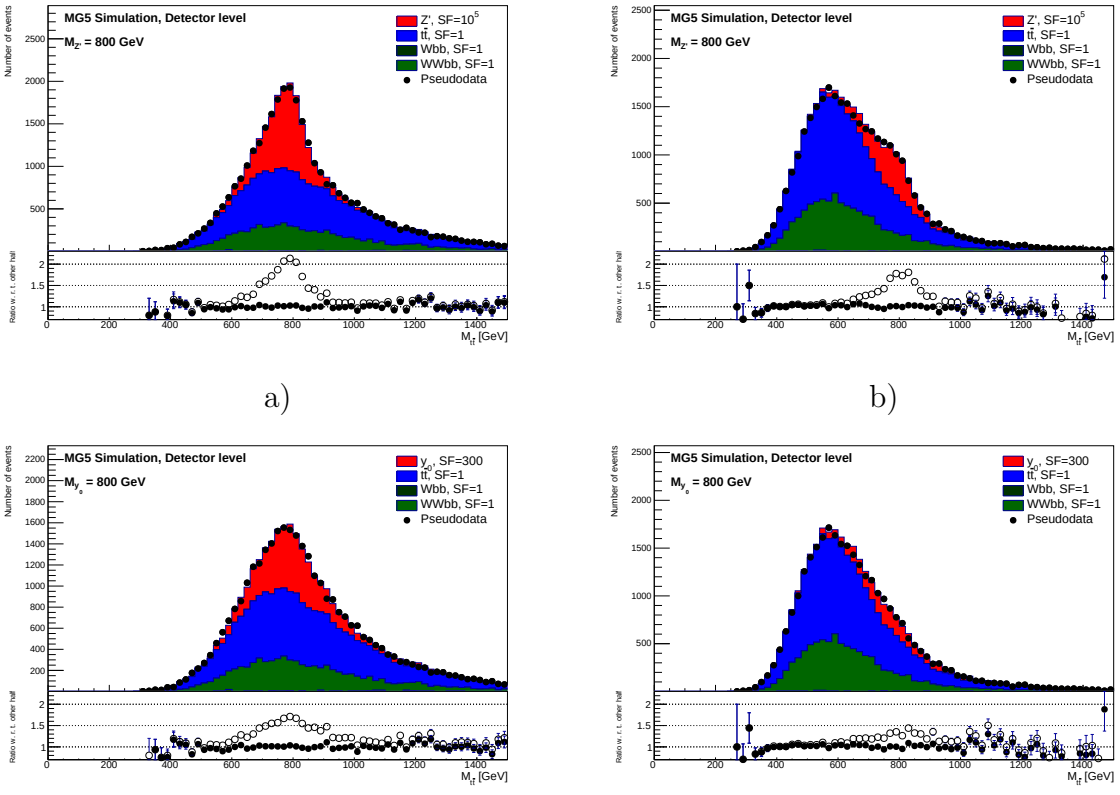


Figure A.8: Combined spectrum of the reconstructed invariant $t\bar{t}$ pair mass spectra ($m_{\mathrm{t\bar{t}}}$) at the detector level for semi-boosted (a), c)) and boosted topology (b), d)) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{\mathrm{Z'}} = 800$ GeV (top row) and with $M_{\mathrm{y_0}} = 800$ GeV (bottom row), all samples are weighted on the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

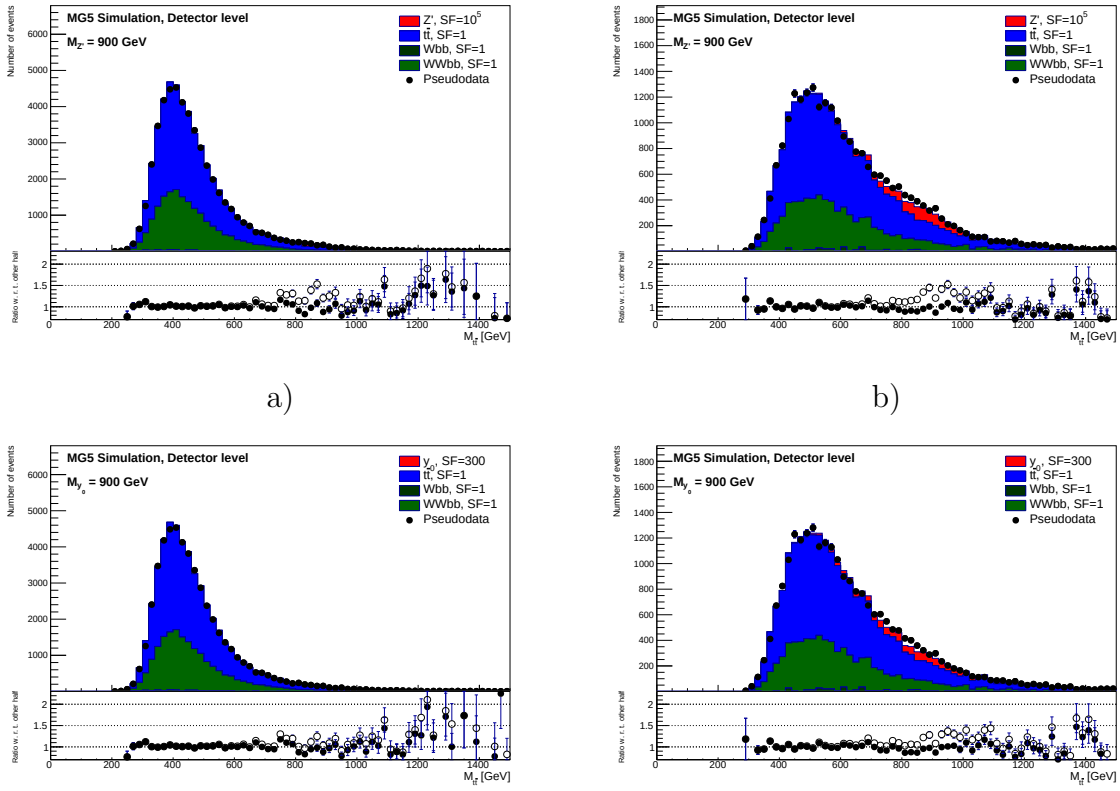


Figure A.9: Combined spectrum of the reconstructed invariant $t\bar{t}$ pair mass spectra ($m_{\mathrm{t\bar{t}}}$) at the detector level for resolved (a), c)) and semi-boosted mixed topology (b), d)) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{\mathrm{Z'}} = 900$ GeV (top row) and with $M_{\mathrm{y_0}} = 900$ GeV (bottom row), all samples are weighted on the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

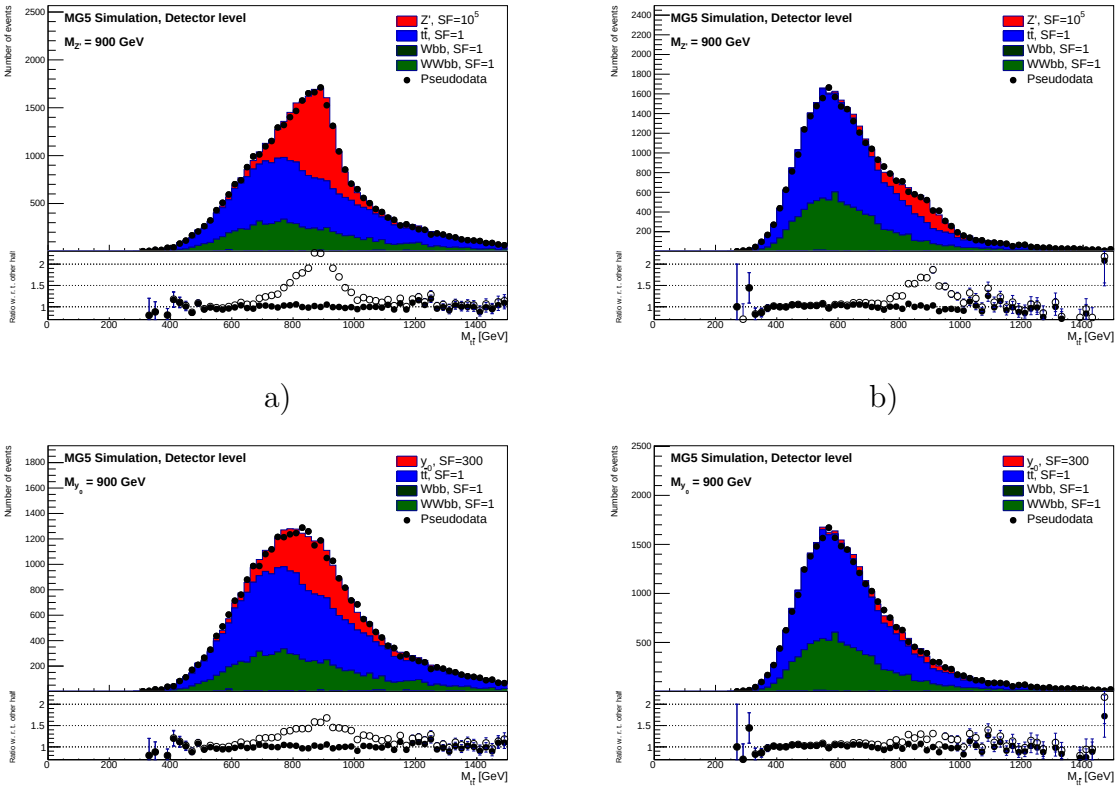


Figure A.10: Combined spectrum of the reconstructed invariant $t\bar{t}$ pair mass spectra ($m_{\mathrm{t\bar{t}}}$) at the detector level for semi-boosted (a), c)) and boosted topology (b), d)) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{\mathrm{Z'}} = 900$ GeV (top row) and with $M_{\mathrm{y_0}} = 900$ GeV (bottom row), all samples are weighted on the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

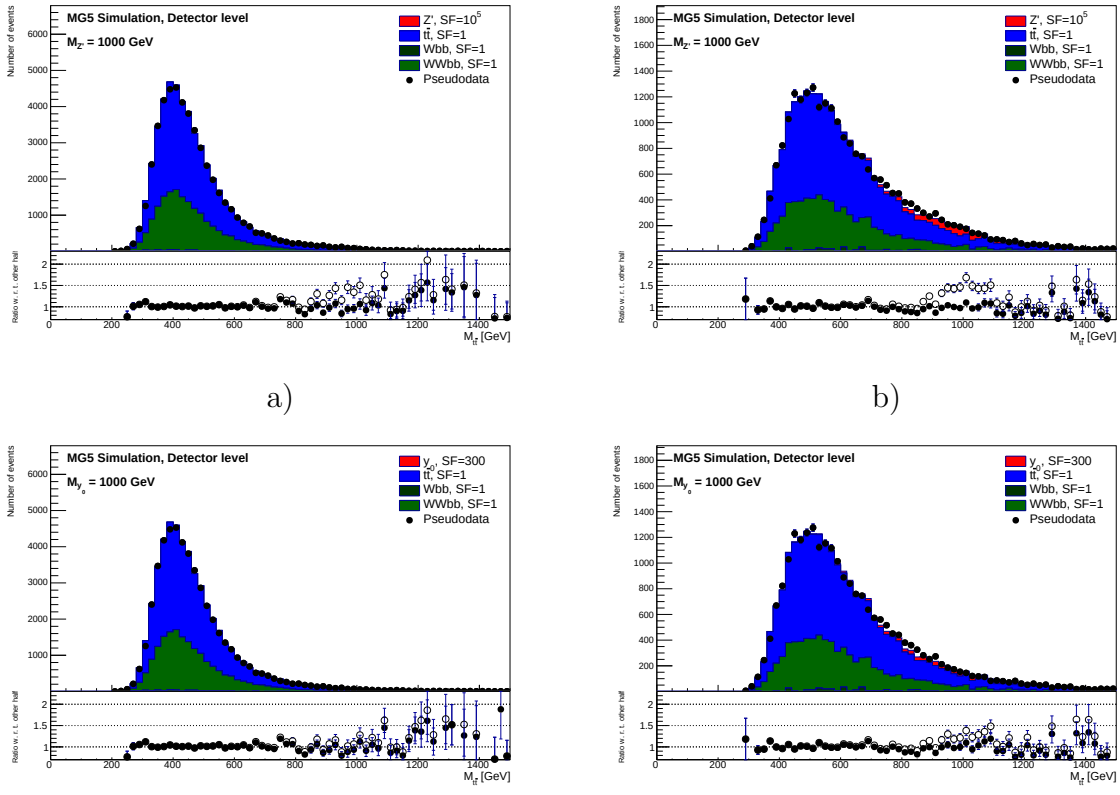


Figure A.11: Combined spectrum of the reconstructed invariant $t\bar{t}$ pair mass spectra ($m_{\mathrm{t\bar{t}}}$) at the detector level for resolved (a), c)) and semi-boosted mixed topology (b), d)) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{\mathrm{Z'}} = 1000$ GeV (top row) and with $M_{\mathrm{y_0}} = 1000$ GeV (bottom row), all samples are weighted on the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

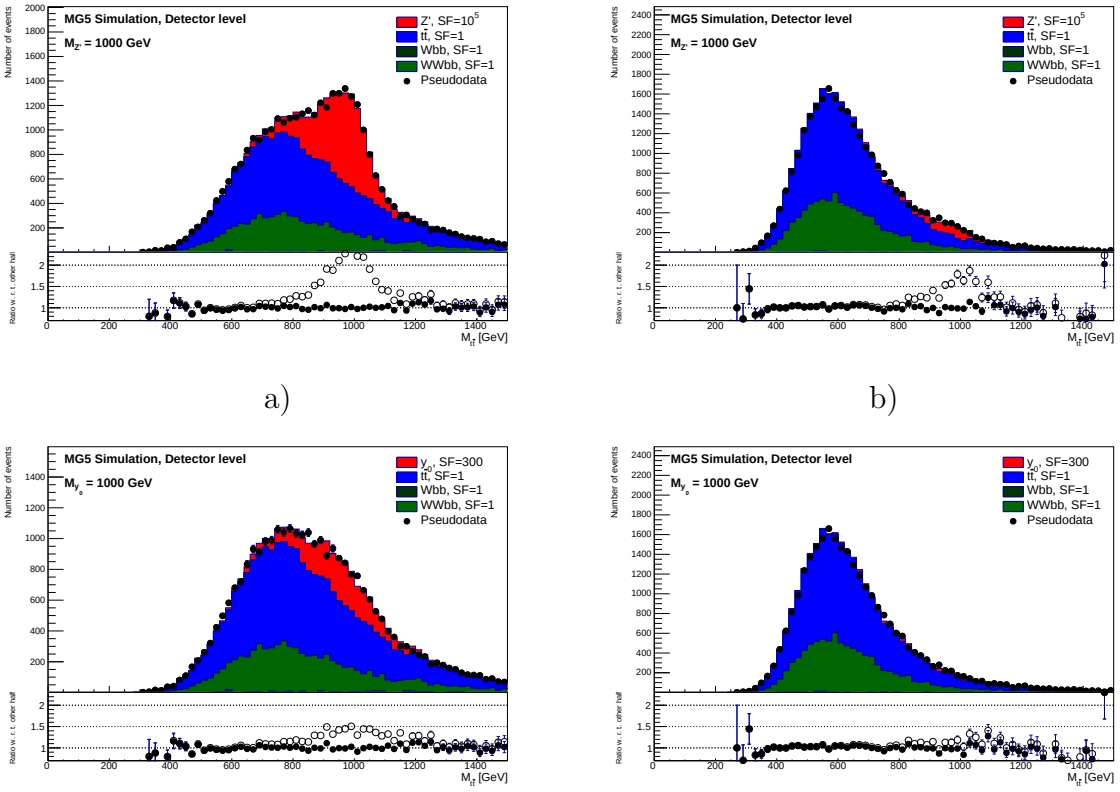


Figure A.12: Combined spectrum of the reconstructed invariant $t\bar{t}$ pair mass spectra ($m_{\mathrm{t\bar{t}}}$) at the detector level for semi-boosted (a), c)) and boosted topology (b), d)) and particle levels (bottom row) after the application of the selection conditions for the signal sample with $M_{\mathrm{Z'}} = 1000$ GeV (top row) and with $M_{\mathrm{y_0}} = 1000$ GeV (bottom row), all samples are weighted on the luminosity of the $t\bar{t}$ sample for comparison purposes. The pseudodata denotes the combined spectrum from statistically independent samples. Bottom part of each plot shows the ratio between pseudodata and combined spectrum (full markers) and ratio between combined spectrum with and without the addition of the hypothetical particle signal (open markers).

Appendix B

Top Anti-top Quark Pair Invariant Mass Significances

Significances of the hypothetical particles Z' and y_0 signal are presented here for the reconstructed $t\bar{t}$ invariant mass spectra. Significances corresponds to the unfolded spectra from Appendix A and illustrates the performance of the unfolding and its influence on the hypothetical particle signal strength. Resulting significances for hypothetical particle masses $M_{\mathrm{Z'/y_0}} = 500$ GeV and $M_{\mathrm{Z'/y_0}} = 600$ GeV are shown in Fig. B.1, for masses $M_{\mathrm{Z'/y_0}} = 700$ GeV and $M_{\mathrm{Z'/y_0}} = 800$ GeV in Fig. B.2 and for masses $M_{\mathrm{Z'/y_0}} = 900$ GeV and $M_{\mathrm{Z'/y_0}} = 1000$ GeV in Fig. B.2.

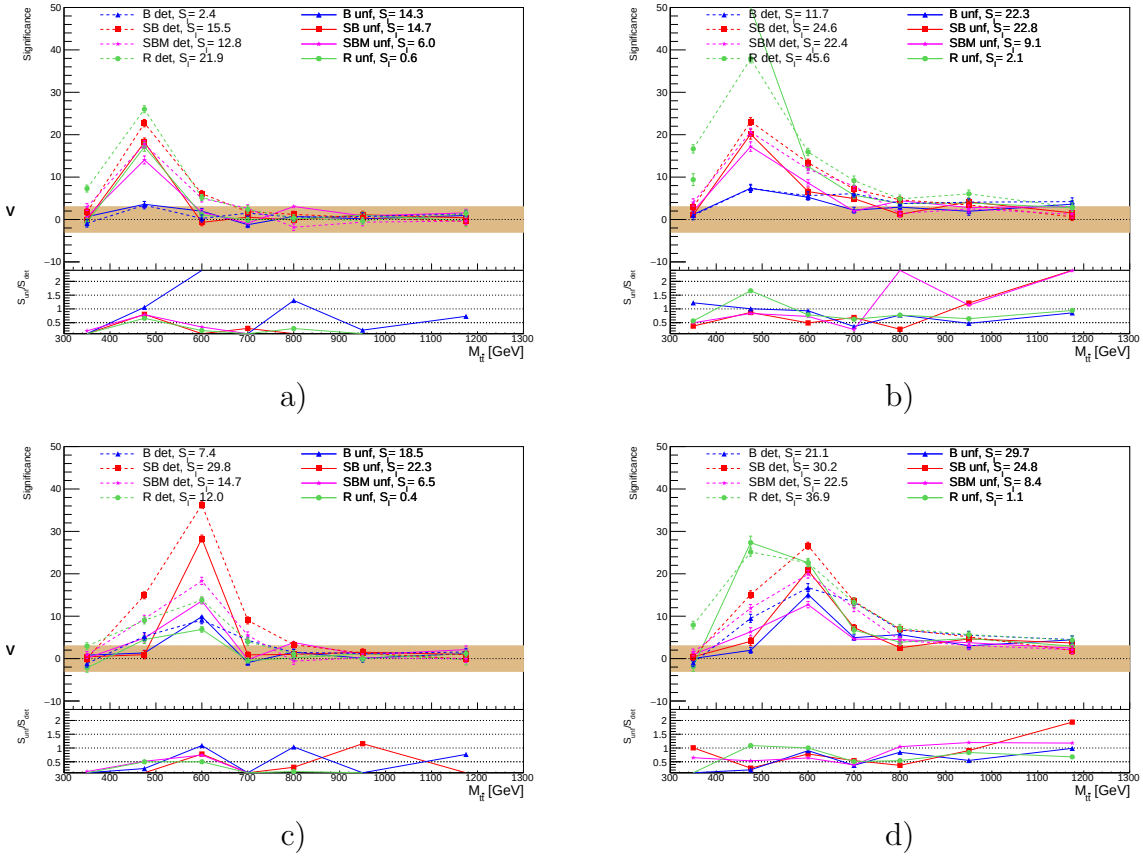


Figure B.1: The detector (open markers, dashed line) and unfolded (full markers, solid line) significances for the invariant mass of the reconstructed $t\bar{t}$ pair ($M_{\mathrm{t\bar{t}}}$) for samples with $M_{\mathrm{Z'}} = 500$ GeV a) and $M_{\mathrm{y_0}} = 500$ GeV b) and for samples with $M_{\mathrm{Z'}} = 600$ GeV c) and $M_{\mathrm{y_0}} = 600$ GeV d), plotted for all the topologies in each bin. The band defines the area where the absolute value of the significance is below three, corresponding to the 3-σ interval. The lower pads present the ratios of the unfolded over the detector level significances, without uncertainties which are highly correlated.

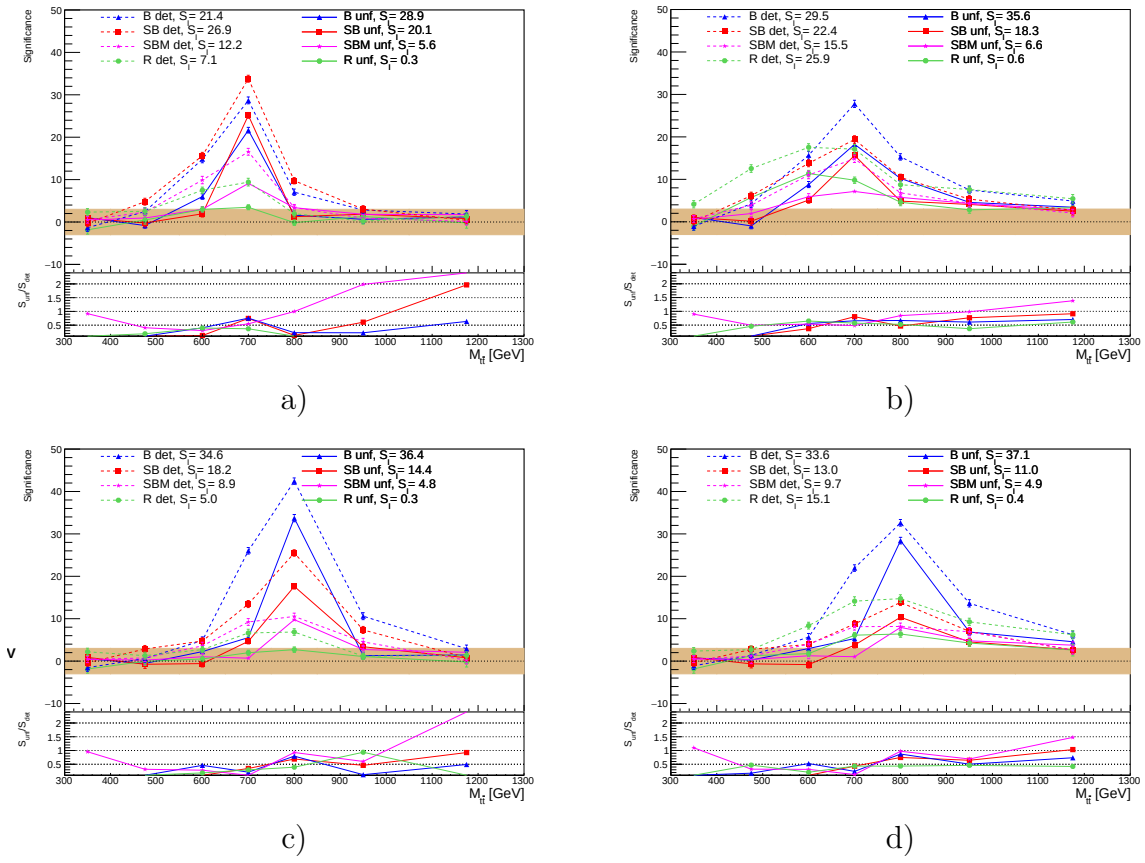


Figure B.2: The detector (open markers, dashed line) and unfolded (full markers, solid line) significances for the invariant mass of the reconstructed $t\bar{t}$ pair ($M_{\mathrm{t\bar{t}}}$) for samples with $M_{\mathrm{Z'}} = 700$ GeV a) and $M_{\mathrm{y_0}} = 700$ GeV b) and for samples with $M_{\mathrm{Z'}} = 800$ GeV c) and $M_{\mathrm{y_0}} = 800$ GeV d), plotted for all the topologies in each bin. The band defines the area where the absolute value of the significance is below three, corresponding to the 3-σ interval. The lower pads present the ratios of the unfolded over the detector level significances, without uncertainties which are highly correlated.

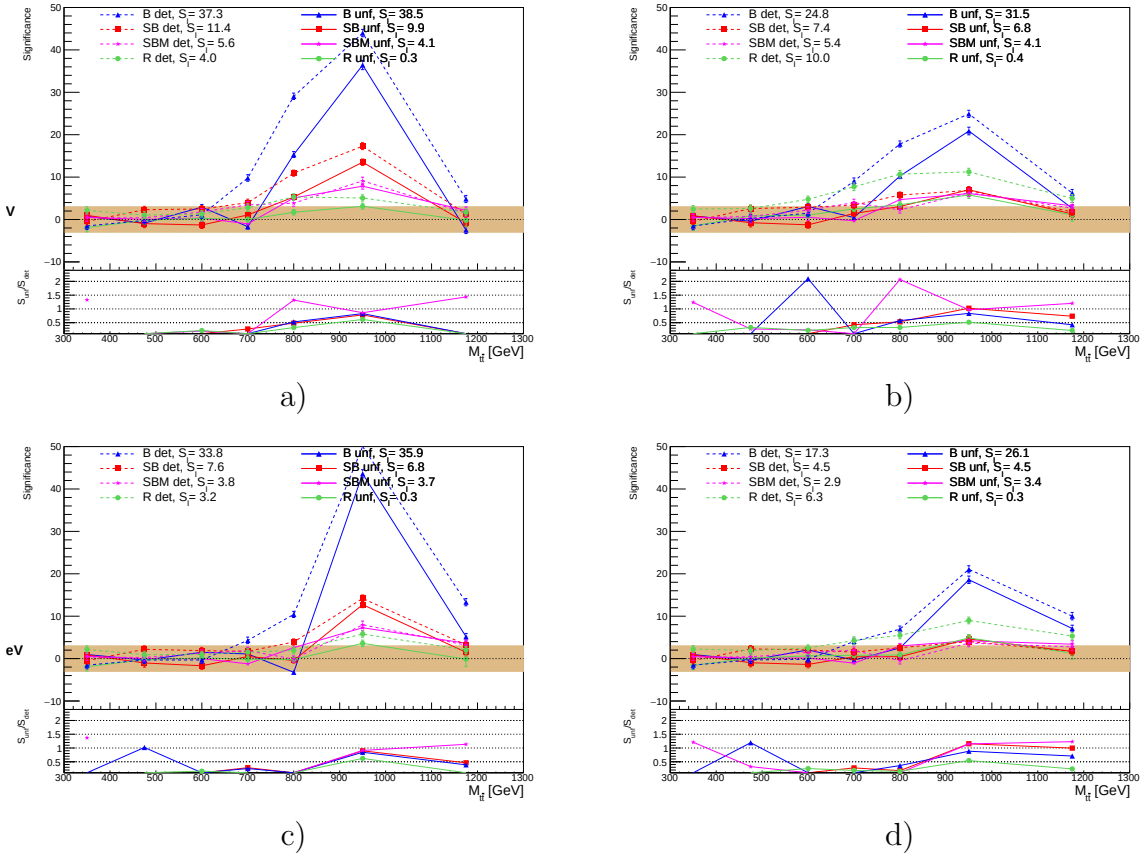


Figure B.3: The detector (open markers, dashed line) and unfolded (full markers, solid line) significances for the invariant mass of the reconstructed $t\bar{t}$ pair ($M_{\mathrm{t\bar{t}}}$) for samples with $M_{\mathrm{Z'}} = 900$ GeV a) and $M_{\mathrm{y_0}} = 900$ GeV b) and for samples with $M_{\mathrm{Z'}} = 1000$ GeV c) and $M_{\mathrm{y_0}} = 1000$ GeV d), plotted for all the topologies in each bin. The band defines the area where the absolute value of the significance is below three, corresponding to the 3-σ interval. The lower pads present the ratios of the unfolded over the detector level significances, without uncertainties which are highly correlated.