Searches for exotic heavy quarks decaying into a W-Boson and a b-Quark with the ATLAS experiment

Nektarijevic, Snezana

UNIVERSITÉ DE GENÈVE
Département de physique nucléaire et corpusculaire

FACULTÉ DES SCIENCES
Professeur Allan Clark

Searches for Exotic Heavy Quarks Decaying into a W-Boson and a b-Quark with the ATLAS Experiment

THÈSE

présentée à la Faculté des sciences de l'Université de Genève
pour obtenir le grade de Docteur ès sciences, mention physique

par

Snežana Nektarijević

de Serbie

Thèse N° 4757

GENÈVE
Atelier d'impression ReproMail
2015

UNIVERSITÉ DE GENÈVE

FACULTÉ DES SCIENCES

Doctorat ès sciences
Mention physique

Thèse de *Madame Snezana NEKTARIJEVIC*

intitulée :

"Searches for Exotic Heavy Quarks Decaying into a W-Boson and a b-Quark with the ATLAS Experiment"

La Faculté des sciences, sur le préavis de Monsieur A. CLARK, professeur honoraire et directeur de thèse (Département de physique nucléaire et corpusculaire), Monsieur X. WU, professeur associé et codirecteur de thèse (Département de physique nucléaire et corpusculaire), Monsieur G. IACOBUCCI, professeur ordinaire (Département de physique nucléaire et corpusculaire), et Monsieur G. MIKENGERG, professeur (Department of Particle Physics, Weizmann Institute of Science, Rehovot, Israël), autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 15 janvier 2015

Thèse - 4757 -

Le Doyen

N.B. - La thèse doit porter la déclaration précédente et remplir les conditions énumérées dans les "Informations relatives aux thèses de doctorat à l'Université de Genève".

Physics isn't the most important thing. Love is.
Richard P. Feynman

Abstract

In this thesis two searches for the pair production of hypothetical heavy quarks decaying into a W boson and a b quark are presented. Both searches are performed in final states with one electron or muon, at least three jets, and significant missing transverse energy. The analyses are based on proton-proton collision data at 7 TeV centre-of-mass energy collected by the ATLAS experiment at CERN's Large Hadron Collider in 2011.

The first search, performed with a data set corresponding to an integrated luminosity of 1.04 fb^{-1}, uses a chiral fourth generation up-type quark t' as signal benchmark. In this analysis, the event reconstruction was performed using a kinematic likelihood fit to assign the reconstructed physics objects in the event to the initial quark and antiquark. The fit configuration was optimised to provide the best resolution of the reconstructed t' quark mass. The optimisation procedure comprised a detailed study of the performance of the fit with various choices of reconstructed jets and nominal t' quark masses as input to the fit, as well as different treatments of b-tagged jets. The study played an important role in the analysis, as they provided a powerful final discriminant, the reconstructed t' quark mass, which allowed for constraining some of the systematic uncertainties by profiting from using the full kinematic distribution in the statistical interpretation. In the absence of any excess of data events above the Standard Model expectation, an upper limit on the $t'\bar{t}'$ production cross-section was derived, which translates into an observed (expected) lower bound on the t' quark mass of 404 (394) GeV.

The second analysis, a search for a vector-like T, was performed with a data set corresponding to an integrated luminosity of 4.7 fb^{-1}, using a multi-variate technique. The T quark reconstruction was based on the identification of all participants of the hard process by making use of b-tagging probabilities, W-boson mass constraints for the determination of the longitudinal momentum of the neutrino, and the minimisation of the difference between the reconstructed masses of the T and $\bar{T}$ quarks. With all elements of the event identified, a set of variables was derived, exploiting the differences in the boost of the reconstructed W bosons and T quarks between signal and background. These variables were used as input to a neural network training, which provided a discriminator on which the final selection requirement was imposed. The neural network training was optimised by testing the effects of various machine learning algorithms, training functions, and a number of other training settings. The selection requirement on the neural network discriminator was optimised simultaneously with the binning of the final discriminant, the reconstructed T quark mass, to achieve the best expected limit on the $T\bar{T}$ production cross section. In the absence of any excess of data events over the Standard Model expectation, lower bounds on the T-quark mass of 618.4 (703.5) GeV observed (expected) were derived. Compared to an equivalent cut-based analysis performed with the same observables an improvement of $\sim$ 60 GeV on the expected lower bound of the T-quark mass was found, which translates into a $\sim$9 % improvement of the excluded production cross section.

Résumé

Dans cette thèse, deux recherches pour des quarks lourds hypothétiques, produits en paires et se désintégrant en un boson W et un quark b, sont présentées. Les deux recherches sont effectuées pour les états finaux avec un électron ou muon, au moins trois gerbes, et une quantité importante d'énergie transverse manquante. Les analyses sont réalisées avec les données des collisions proton-proton avec une énergie de centre de masse de 7 TeV, enregistrées par l'expérience ATLAS au Grand Collisionneur de Hadrons du CERN en 2011.

La première recherche, réalisée avec un ensemble de données correspondant à une luminosité intégrée d'environ $1,04$ fb^{-1}, utilise le quark de type up de la quatrième génération chirale, t', comme signal. Dans cette analyse, la reconstruction d'événements a été réalisée en utilisant une régression basée sur une fonction de vraisemblance cinématique servant à assigner les objets physiques reconstruits au quark et à l'antiquark initiaux. La configuration de la régression a été optimisée pour fournir la meilleure résolution de la masse reconstruite du quark t'. Le procédure d'optimisation comprend une étude détaillée de la performance de la régression avec des ajustements divers, comprenant le choix des gerbes reconstruites et la masse nominale du quark t' comme paramètres de la régression, ainsi que les différents traitements des gerbes attribuées à un quark b. Cette étude a joué un rôle important dans l'analyse, car elle offre un discriminant final puissant, la masse reconstruite du quark t', ce qui a permis de contraindre certaines des incertitudes systématiques, en tirant profit de l'utilisation de la distribution cinématique complète dans l'interprétation statistique. En l'absence de tout excès d'événements de données au-delà des prédictions du Modèle Standard, une limite supérieure sur la section efficace de la production de $t'\bar{t}'$ a été dérivée, ce qui se traduit par une limite inférieure observée (prédite) de la masse du quark t' de 404 (394) GeV.

La deuxième analyse, la recherche d'un quark T ayant une structure vectorielle, a été réalisée avec un ensemble de données correspondant à une luminosité intégrée de 4.7 fb^{-1}, en utilisant une technique à plusieurs variables. La reconstruction du quark T a été basée sur l'identification de tous les participants de l'interaction la plus énergétique en faisant usage de la probabilité des gerbes d'être attribuées à un quark b, de la contrainte de masse du boson W pour la détermination de la quantité de mouvement longitudinale du neutrino, et de la minimisation de la différence entre les masses reconstruites des quarks T et $\bar{T}$. Avec tous les éléments de l'événement identifiés, un ensemble de variables a été dérivées, exploitant les différences dans la poussée cinématique des bosons W et des quarks T reconstruits, entre le signal et le bruit de fond. Ces variables ont été utilisées comme paramètres pour l'entraînement d'un réseau de neurones, qui a fourni un discriminateur sur lequel le critère de sélection final a été imposé. L'entraînement du réseau de neurones a été optimisé en testant les effets des algorithmes d'apprentissage automatique et des fonctions d'activation différentes, ainsi que d'un certain nombre d'autres paramètres d'entraînement. Le critère de sélection sur le discriminateur du réseau de neurones a été optimisé simultanément avec le choix d'intervalle du discriminant final, la masse reconstruite du quark T, pour atteindre le meilleure limite prédite sur la section efficace de la production de $T\bar{T}$. En l'absence de tout excès d'événements de données sur la prédiction du Modèle Standard, des limites inférieures sur la masse observée (prédite) du quark T de $618,4\,(703,5)$ GeV ont été obtenues. Par rapport à une analyse similaire basée sur des coupures sur les mêmes observables, une amélioration de ~ 60 GeV sur la limite inférieure prédite de la masse du quark T a été obtenue, ce qui se traduit par une amélioration de $\sim 9\%$ de la section efficace exclue pour la production de $T\bar{T}$.

Contents

Chapter 1

Introduction

Simplicity is the ultimate sophistication.
Leonardo da Vinci

The continuous quest for the very origin and nature of matter has created a need for precise mathematical and experimental tools. This need has driven the remarkable advancement of theoretical frameworks as well as experimental facilities and techniques over the past century. The milestones of modern particle physics had been laid by Maxwell's gauge field theory of electromagnetism, which paved the road to the development of the quantum field theories, and the Rutherford experiment, by which scattering experiments were revolutionised as a method of studying the structure of matter. Since then, the field of particle physics has blossomed, pushing the limits of the rapidly expanding knowledge, and strengthening the demand on the power and accuracy of scientific tools. This impressive scientific path has resulted in the formulation of the Standard Model (SM) of particle physics, the extraordinarily simple and elegant theory based on the framework of relativistic quantum field theories. The SM explains all present experimental data with a remarkable accuracy. It was recently crowned by the discovery of a new scalar (Higgs) boson, the final missing piece of the theory. The story of modern particle physics could thus be summarised by saying that the quest for the ultimate simplicity has sparked the advancement of science to an astonishing sophistication.

Although describing the current particle physics collider data with a high accuracy, the current formulation of the SM cannot be the ultimate theory of matter. It does not provide any description of gravity, and fails to explain some astrophysical observations, such as the presence of dark matter. Over the past four decades, a large effort has been invested into the development of theoretical extensions of the SM, both through precise examination of the SM predictions, and the development of new theoretical concepts. The recent discovery of the Higgs boson has strengthened the significance of the SM and placed stringent constraints on numerous new physics models. Experimental signatures of the new models in the allowed phase space regions are extensively sought for in the particle physics laboratories around the world. The currently most extensive source of experimental data for such searches is provided by CERN's main collider, the Large Hadron Collider (LHC), with its four large particle detectors. This machine, unrivalled in its energy and intensity, has the world leading potential for revealing new physics phenomena. In three years of its operation, between November 2009 and December 2012, the LHC and its detectors have produced many significant results, including the re-observation of

most SM particles, many precision measurements of SM processes in a new collision energy regime, and most remarkably the discovery of the Higgs boson. The largest among the LHC detectors, the ATLAS detector, which played a key role in this discovery, will continue being of central importance for the new physics searches when the LHC starts producing collision data again in 2015.

The early measurements performed by the ATLAS detector were focused on processes with high production rates, which were relatively simple to extract from data. However, new physics signals can also arise at much lower production rates than the corresponding SM backgrounds, so that more sophisticated analysis techniques are required for their extraction. In this thesis, two searches for new heavy quarks are presented, both relying on advanced analysis methods. Both searches were performed using the pp collision data delivered by the LHC at a centre-of-mass energy ($\sqrt{s}$) of 7 TeV and collected with the ATLAS detector in 2011. Searches were performed on events with one electron or muon, at least three jets, and a significant missing transverse energy due to the presence of a neutrino in the final state.

The first search presented in this thesis is a search for a chiral fourth generation top-like quark, t'. At the time of the search, i.e. before the discovery of the new scalar boson, the existence of the fourth generation was still compatible with electroweak data. This simplest SM extension was primarily expected to shed some light on the matter-antimatter asymmetry in the Universe as well as the forward-backward asymmetry in the production of top quark-antiquark pairs. The event reconstruction in this analysis was performed by employing a kinematic fit to best assign the reconstructed objects to the initial quark and anti-quark. The main contribution of the author to this analysis was a set of studies performed in order to optimise the configuration of the fitter that gave the best resolution of the reconstructed t'-quark mass.

The second analysis of this thesis is a search for heavy quarks with vector-like couplings, decaying into a W-boson and a b-quark. Vector-like quarks (VLQs) are predicted by some of the models aiming to solve the so-called hierarchy problem, that is attempting to provide an explanation for the mechanism which keeps the mass of the SM Higgs boson at the electroweak scale. This search was based on a multi-variate event selection based on a neural network discriminant, developed by the author using an event reconstruction and a set of kinematic variables adopted from an existing analysis. The analysis potential of this method was quantified as the sensitivity improvement achieved by this analysis compared to a similar analysis based on the selection requirements on each variable independently. This analysis and comparison were carried out by the author.

This thesis is organised as follows. The theoretical foundation for the SM and the searches presented is set in Chapter 2. The experimental setup, including the descriptions of relevant aspects of the LHC and the ATLAS detector, is described in Chapter 3. The simulation of pp collision events is explained in Chapter 4, and the reconstruction of physics objects used in the analyses in Chapter 5. In Chapter 6, the multi-variate and reconstruction tools used in the searches are described. In Chapter 7, common elements of the two searches are explained, followed by the presentation of the t' search in Chapter 8, and the VLQ search in Chapter 9. Short conclusions of these two searches are presented in Chapter 10. A detector related technical work carried out by the author as a so-called qualification task is presented in Appendix A.

Chapter 2

Theoretical Background

The current understanding of the fundamental structure of matter and its interactions is summarised in the Standard Model (SM) of particle physics. According to the SM, the whole known Universe is built up from a dozen of elementary matter particles, steered by three fundamental forces, explained by means of relativistic quantum field theory in the Lagrangian formalism[1]. In the SM, the elementary matter particles are described as excitations of fermionic (spin 1/2) quantum fields, obeying the Fermi-Dirac statistics [1,2]. The fundamental interactions between matter particles are represented as an exchange of a force carrier particle, a quantum of a bosonic (spin 1) field, following the Bose-Einstein statistics [3].

According to their physical properties, the known elementary matter particles are divided into *quarks* and *leptons*, with six particles in each group. The quarks have rather high masses and are subdivided according to their electric charge (Q) into the so-called *up-type* quarks with $Q = +2/3$ and *down-type* quarks with $Q = -1/3$. The up-type quarks are called *up* (u), *charm* (c) and *top* (t) quark in the increasing mass order, while the down-type quarks are called *down* (d), *strange* (s) and *bottom/beauty* (b) quark. Each up-type quark is grouped with one down-type quark to form a so-called *generation*, resulting in three quark generations called the first (I), second (II) and third (III), ordered according to the increasing quark masses. Leptons are subdivided into light massive *charged leptons* with $Q = -1$, and their almost massless and neutral partners called *neutrinos*. The charged leptons are called *electron* (e), *muon* (μ) and *tau* lepton (τ) ordered according to growing masses, while their neutral partners are called *electron neutrino* (ν_e), *muon neutrino* (ν_μ) and *tau neutrino* (ν_τ). A charged lepton and a corresponding neutrino build a lepton generation. One quark generation and the lepton generation with the same position in the mass ordered sequence form a SM generation. The SM matter particle generations are shown in Table 2.1. Each elementary fermion has a corresponding anti-particle with the same spin and mass, but opposite electric charge. This makes a total of 24 known fundamental fermions in the SM. The number of the SM particle generations as well as the mass spectrum of its constituents are neither predicted nor limited by the SM.

Most of the visible matter in the Universe is built from the first SM generation. Protons and neutrons, the building blocks of atom nuclei, are built from the u- and d-quarks. Atoms of all chemical elements consist of nuclei surrounded by a cloud of electrons. All other fundamental

[1]The fourth elementary force, the gravity, does not influence the structure of the matter, as it is 10^{32} weaker than the other elementary forces. It is also not included in the SM because there is no quantum field theory of gravity yet.

particles and the bound states they form have short life times, and eventually decay into the first generation particles or photons, which are explained below.

	I	II	III	Q
Q	$\begin{pmatrix} u \\ d \end{pmatrix}$ 2.3 MeV 4.8 MeV	$\begin{pmatrix} c \\ s \end{pmatrix}$ 1.275 GeV 95 MeV	$\begin{pmatrix} t \\ b \end{pmatrix}$ 173.07 GeV 4.66 GeV	$+2/3$ $-1/3$
L	$\begin{pmatrix} \nu_e \\ e \end{pmatrix}$ < 0.225 keV 0.511 MeV	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}$ < 0.19 MeV 105.66 MeV	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}$ < 18.2 MeV 1.78 GeV	0 -1

Table 2.1: The SM particle generations with the masses taken from [4]. The errors on quoted masses can be found in the cited document.

There are four known fundamental interactions in Nature: the gravitational, electromagnetic, weak and strong force. All these forces other than gravitation are explained by the SM in the Lagrangian mechanism constructed to depict the symmetries of the interactions. The symmetries are considered to be local, which requires the introduction of gauge fields in order to reinforce the invariance of the Lagrangian describing the interaction under the symmetry transformation. The interactions are thus described as gauge theories. Among the SM interactions, only the electromagnetic interaction is abelian. It is carried by the electrically neutral spin-1 *photon* field, which couples to the electric charge and the spin of the particles. Being electrically neutral itself, the photon field does not exhibit self-interaction. The weak interaction is mediated by the electrically charged W^+ and W^-, and neutral Z bosons. The strong interaction is transmitted by eight massless *gluon* fields. The symmetry groups containing the symmetry transformations of the weak and the strong interaction are non-abelian and thus explained by the Yang-Mills theory. There is a common mathematical framework for the electromagnetic and weak interaction according to which above a certain energy scale they form one and the same interaction. After Gerard t'Hofft had shown that this so-called electroweak theory was renormalisable, it had been established as a solid theory. Since it has been experimentally found that the gauge bosons of the weak interaction are massive, the symmetry governing this interaction must be spontaneously broken. This spontaneous symmetry breaking induces the only spin-0 boson of the SM, the Higgs boson. The masses of all elementary fermions can also be explained by the interaction of the corresponding (fermionic) fields with the Higgs field. A new boson with several preliminarily measured properties compatible with the SM Higgs boson, has been discovered at CERN by the ATLAS [5] and CMS [6] collaborations in 2012. However, the studies of the nature of this new boson are still in an early stage.

This chapter is organised as follows. The mathematical foundation of the SM is introduced in Section 2.1. Then, the structure of the SM, the experimental tests of its implications, as well as some of the limitations of the theory are described in Section 2.2. Finally, the benchmark models on which the searches presented in this thesis are based are explained in Section 2.3.

2.1 Mathematical Foundation

2.1.1 Dynamics of Fermionic Fields

The SM of particle physics describes the fundamental matter constituents and interactions within the framework of relativistic quantum field theories (QFT) based on the gauge invariance

principle. In QFT, particles are represented by the excitations of quantum fields. The fields are represented as quantised objects, describing a many-body system by stating the number of identical particles occupying a given state. This description is known as the second quantisation. The states are described by a set of variables $\vec{\alpha} = (\alpha_1, \ldots, \alpha_k)$ (e.g. spin, energy etc.). A field describing N_s possible states would have a form:

$$\left| n_{\vec{\alpha}_1}, \ldots, n_{\vec{\alpha}_{N_s}} \right\rangle . \tag{2.1}$$

For bosonic (integer spin) fields, each $n_{\vec{\alpha}}$ can be any non-negative integer number and the field obeys the Bose-Einstein statistics. Fermionic (half-integer spin) fields, however, have to follow the Pauli's exclusion principle, which allows the $n_{\vec{\alpha}}$ to be either 0 or 1, and accordingly, the fields follow the Fermi-Dirac statistics.

In the SM framework (for a pedagogical introduction see e.g. [7]), all elementary matter particles are described as the excitations of fermionic quantum fields with spin-1/2, represented in the form of spinors with four space-time dependent components $\psi(x^\mu) = (\psi_1(x^\mu), \ldots, \psi_4(x^\mu))$, with the space-time coordinates given by $x^\mu \equiv (x^0, x^1, x^2, x^3,) \equiv (t, x, y, z)$[2]. The equation of motion of such a free fermionic field ψ with a mass m is formulated by Dirac as:

$$(i\gamma^\mu \partial_\mu - m) = 0, \tag{2.2}$$

using the covariant derivative:

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} \equiv \left(\frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3} \right), \tag{2.3}$$

and the so-called Dirac matrices defined as:

$$\gamma^0 = \begin{bmatrix} \mathbb{1}_2 & 0 \\ 0 & \mathbb{1}_2 \end{bmatrix}, \gamma^k = \begin{bmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{bmatrix}, k = 1, 2, 3, \tag{2.4}$$

using the Pauli matrices:

$$\sigma^1 = \mathbb{1}_2, \sigma^1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \sigma^2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \sigma^3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{2.5}$$

The fundamental interactions are derived from the symmetry considerations using the Lagrange formalism. Dynamics of a Dirac spinor is given by the Lagrangian density $\mathcal{L}$, which is a function of the field and its covariant derivatives:

$$\mathcal{L} = \mathcal{L}(\psi, \partial_\mu \psi). \tag{2.6}$$

[2]The correct definition would be $x^\mu \equiv (ct, x, y, z)$. However, c is dropped as the natural units are used, in which $c = \hbar = 1$.

The system needs to satisfy the principle of the least action, with the action S defined as:

$$S = \int d^4x\mathcal{L}. \tag{2.7}$$

This principle translates into a $\delta S = 0$ requirement, which leads to the relativistic Euler-Lagrange equation, applicable to any component of the wave function:

$$\frac{\partial\mathcal{L}}{\partial\psi_i} - \partial_\mu\left(\frac{\mathcal{L}}{\partial(\partial_\mu\psi_i)}\right) = 0. \tag{2.8}$$

The principle of the least action remains valid in case a field transformation is applied, under which the dynamics of the system is invariant. For an infinitesimal space-time transformation:

$$x^\mu \to x'^\mu = x^\mu + \delta x^\mu, \tag{2.9}$$

the field transformation has a form:

$$\psi(x) \to \psi'(x') = \psi(x) + \delta\psi(x), \tag{2.10}$$

with:

$$\delta\psi(x) = \delta x^\mu\partial_\mu\psi_i + \delta_\psi\psi(x). \tag{2.11}$$

If the transformation is represented by n parameters α_i, the invariance of the system under this transformation translates into $\partial S/\partial\alpha_i = 0$ for $i = 1, \ldots, n$. Applying Equation 2.8, this expression transforms into:

$$\frac{\partial S}{\partial\alpha_i} = \int d^4x\partial_\mu\left(\frac{\delta x^\mu}{\delta\alpha_i}\mathcal{L} + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\frac{\delta_\psi\psi}{\delta\alpha_i}\right) = 0. \tag{2.12}$$

Defining the field current j_i^μ corresponding to the invariance under the α_i component of the transformation as:

$$j_i^\mu \equiv \left(\frac{\delta x^\mu}{\delta\alpha_i} + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\frac{\delta_\psi\psi}{\delta\alpha_i}\right), \tag{2.13}$$

Equation 2.12 provides:

$$\partial_\mu j_i^\mu = 0. \tag{2.14}$$

In other words, invariance of the system under a given transformation has lead to the existence of a *conserved current* of the field. This dynamics is predicted by the Noether theorem [8]. The space integral of the time-like component of this current, j_i^0, which reads as:

$$Q_i \equiv \int d^3x j_i^0 = \int d^3x\left(\frac{\delta x^0}{\delta\alpha_i} + \frac{\partial\mathcal{L}}{\partial(\partial_0\psi)}\frac{\delta_\psi\psi}{\delta\alpha_i}\right), \tag{2.15}$$

is the *conserved charge* of the field, since it is time independent:

$$\frac{dQ_i}{dt} = \int d^3x \partial_0 j_i^0 = -\int d^3x \left(\boldsymbol{\nabla} \cdot \mathbf{j}_i\right) = 0. \tag{2.16}$$

Conserved field currents and charges also result from the invariance under an internal (phase) transformations of the field, such as rotations of the field of the form:

$$\psi(x) \to \psi'(x) = e^{-i\alpha}\psi(x), \tag{2.17}$$

in which the transformation parameter α is global, i.e. space-time independent.

2.1.2 Symmetries and the Elementary Interactions

The picture described in the previous section changes when the invariance of the system under a local, i.e. space-time dependent ($\alpha = \alpha(x)$) phase transformations is demanded. The Lagrangian density of a free Dirac field has the form:

$$\mathcal{L} = \bar{\psi} i\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi. \tag{2.18}$$

The first term is the kinetic term and the second one is the mass term of the field. The following discussion shows that the form of a Lagrangian of a given interaction is determined by the group of local field transformations under which the interaction is invariant.

Under a local unitary transformation $U(x)$, a Dirac field $\psi(x)$ transforms as:

$$\psi(x) \to \psi'(x) = U(x)\,\psi(x). \tag{2.19}$$

Unitary transformations of an N-dimensional space of states form a symmetry group $U(N)$. In particle physics, for the case $N = 1$, as is the case for the electromagnetic interaction explained in Section 2.2.1, the rotations of the complex internal phase are described by the unitary group U(1). In case of $N > 1$, as for the weak and strong interactions, the symmetries of the interactions are explained by the special unitary groups $SU(N)$. The $SU(N)$ transformations are represented by $N \times N$ unitary matrices with determinant 1. The dimension of an $SU(N)$ group is then $N^2 - 1$, which is also the number of the group generators needed for a representation.

Any unitary transformation has a form:

$$U(x) = e^{i\vec{\alpha}(x)\vec{T}}, \tag{2.20}$$

where $\vec{T}$ is the vector of the transformation group generators T^a, which can be represented by $n \times n$ matrices, and $\vec{\alpha}(x)$ the vector of the corresponding space-time dependent phase rotations α^a [9]. Assuming that $U(x)$ is an infinitesimally small transformation, it can be expanded as:

$$U(x) = 1 + i\alpha^a(x)\,T^a + \mathcal{O}\left(\alpha^2\right). \tag{2.21}$$

The unitarity of $U(x)$ implies that the T^as must be traceless hermitian matrices, i.e. that $T^a = T^{a\dagger}$ holds.

Such a transformation leaves the mass term of the expression 2.18 invariant, as the field ψ is invariant under this local transformation by definition. On the contrary, the field derivative is not preserved, since it contains an additional term proportional to $\partial_\mu \alpha^a(x)$. In order to preserve the invariance of the kinetic term of the Lagrangian it is necessary to extend the definition of the derivative by introducing one gauge field A^a_μ transforming as:

$$A^a_\mu \to A^{a\prime}_\mu = A^a_\mu - \frac{1}{g}\partial_\mu \alpha^a(x), \tag{2.22}$$

for each generator T^a. Here, g represents a coupling constant. With the gauge field introduced, the covariant derivative reads as:

$$D_\mu = \partial_\mu - igA^a_\mu T^a. \tag{2.23}$$

Replacing ∂_μ by D_μ in the Lagrangian assures the form invariance of the kinetic term $\bar{\psi} i\gamma^\mu D_\mu \psi$ of the field ψ. However, since a new field is introduced, the Lagrangian needs to be completed by the kinetic and the mass term for the gauge field A^a_μ. The kinetic term is given by means of the field tensor $F^a_{\mu\nu}$, which is defined by the commutator of the covariant derivatives:

$$igF^a_{\mu\nu}T^a = [D_\mu, D_\nu]. \tag{2.24}$$

Using the definition of the structure constants f^{abc} of the symmetry group:

$$\left[T^a, T^b\right] = if^{abc}T^c, \tag{2.25}$$

the explicit expression for the *field strength tensor* $F^a_{\mu\nu}$ reads as:

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + gf^{abc}A^b_\mu A^c_\nu. \tag{2.26}$$

In case all structure constants f^{abc} are equal to 0, the symmetry group is called *abelian*. If at least one f^{abc} is non-zero, the group becomes *non-abelian*. This point makes an important difference. The kinetic term of the gauge field is given by $1/4F^a_{\mu\nu}F^{a\,\mu\nu}$. In case of a non-abelian symmetry group, described by the Yang-Mills theories [10], the kinetic term contains terms with three or four gauge fields multiplied. These terms correspond to the interactions among the gauge fields. In the abelian case, however, these terms are not present, suggesting that the gauge fields do not interact among themselves.

The expression 2.26 and its derivatives are locally invariant, and thus also the kinetic term of the gauge field A^a_μ. On the contrary, the mass term $m_A A^a_\mu A^{a\,\mu}$ corresponding to the gauge field is not locally invariant. Thus, in order to preserve the invariance of the total Lagrangian (which includes kinetic and mass terms of both the Dirac field and the gauge field), the gauge field must be massless. It can also be shown that any gauge field transforms as a vector under the Lorentz transformations, and has a spin-1. In other words, the gauge fields are in general

massless vector bosons. This holds for the electromagnetic and strong interaction, but runs into problems for the weak interaction, as explained in Section 2.2.2.

After the invariance under the continuous local transformations is ensured, it is furthermore required that the interaction is also invariant under the discrete space-time coordinate transformations: parity (i.e. the inversion of the algebraic sign of the space coordinates) and time inversion. Finally, the form of the Lagrangian of a free Dirac field, which is invariant under parity, time inversion and local transformation $U(x)$ is:

$$\mathcal{L} = \bar{\psi}\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi + g\bar{\psi}\gamma^{\mu}A_{\mu}^{a}\psi - \frac{1}{4}F_{\mu\nu}^{a}F^{a\,\mu\nu}. \tag{2.27}$$

These symmetry considerations have extensive implications, which could be summarized as follows: the requirement of the invariance of an interaction under a local gauge transformation induces a massless gauge field. All elementary interactions described by the SM can be induced in this way, as explained in the following sections.

2.1.3 Spontaneous Breaking of a Local $SU(2)$ Symmetry

In Section 2.2.2, it will be shown that the gauge bosons of the weak interaction cannot be massless. Their massiveness suggests that the symmetry inducing them, the local $SU(2)$ symmetry, must be broken. The spontaneous breaking of this symmetry is explained in this section.

Let the field $\phi(x)$ be an $SU(2)$ doublet of complex scalar fields:

$$\phi(x) = \begin{pmatrix} \phi_{\alpha}(x) \\ \phi_{\beta}(x) \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} \phi_1(x) + i\phi_2(x) \\ \phi_3(x) + i\phi_4(x) \end{pmatrix}. \tag{2.28}$$

The corresponding Lagrangian (with the x dependence valid but not explicitly written) has the form:

$$\mathcal{L} = (\partial_{\mu}\phi)^{\dagger}(\partial^{\mu}\phi) - \mu^2\phi^{\dagger}\phi - \lambda\left(\phi^{\dagger}\phi\right)^2 = (\partial_{\mu}\phi)^{\dagger}(\partial^{\mu}\phi) - V(\phi), \tag{2.29}$$

with a rotationally symmetric potential:

$$V(\phi) = \mu^2\phi^{\dagger}\phi + \lambda\left(\phi^{\dagger}\phi\right)^2. \tag{2.30}$$

Conservation of the invariance of the Lagrangian 2.29 under the (non-abelian) local $SU(2)$ transformations, induces three gauge fields $W_{\mu}^{a}(x)$ with $a = 1, 2, 3$, with which the covariant derivative becomes:

$$D_{\mu} = \partial_{\mu} + ig\frac{1}{2}\sigma_a W_{\mu}^{a}, \tag{2.31}$$

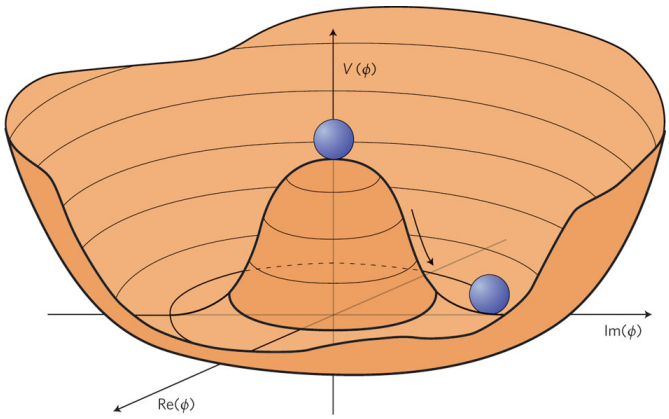


Figure 2.1: The Higgs potential (Equation 2.30) for $\mu^2 < 0$. Taken from [11].

and the Lagrangian reads as:

$$\mathcal{L} = \left(\partial_\mu \phi + ig\frac{1}{2}\boldsymbol{\sigma}\cdot\mathbf{W}_\mu\phi\right)^\dagger \left(\partial^\mu\phi + ig\frac{1}{2}\boldsymbol{\sigma}\cdot\mathbf{W}^\mu\right) - V(\phi) - \frac{1}{4}W_{\mu\nu}W^{\mu\nu}, \tag{2.32}$$

with $\boldsymbol{\sigma}$ being the vector of the Pauli matrices, and $\mathbf{W}_\mu$ being the vector of gauge fields. In case $\mu^2 > 0$, this Lagrangian describes four scalar particles ϕ_i of mass μ interacting with three massless gauge bosons W^a_μ. In the case $\mu^2 < 0$ and $\lambda > 0$, however, the potential $V(\phi)$ becomes the "mexican hat" shape (Figure 2.1) with its minimum at:

$$\phi^\dagger\phi = \frac{1}{2}\left(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2\right) = -\frac{\mu^2}{2\lambda}, \tag{2.33}$$

which represents a sphere in the space of the ϕ_i fields. However, in order to accommodate the sequence expansion of an infinitesimal $SU(2)$ transformation of the filled ϕ, it is necessary to select one particular point on this sphere as the origin of the expansion. It is convenient to select the point characterised by:

$$\phi_1 = \phi_2 = \phi_4 = 0, \quad \phi_3^2 = -\frac{\mu^2}{\lambda} \equiv v^2, \tag{2.34}$$

with the *vacuum expectation value* (vev) $v = \sqrt{-\mu^2/\lambda}$ defined in this way, so that the vacuum becomes:

$$\phi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix}. \tag{2.35}$$

An infinitesimal $SU(2)$ transformation around the vacuum can then be expressed as:

$$\phi'(x) = \frac{e^{i\boldsymbol{\sigma}\cdot\boldsymbol{\theta}(x)/v}}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \tag{2.36}$$

with a real field $h(x)$ (the so-called Higgs field) and three real fields θ_1, θ_2 and θ_3. The θ_i fields can, however, be absorbed by the gauge fields W^i_μ, and thus do not induce any Goldstone bosons.

With this choice of the vacuum and the gauge, the Lagrangian of the transformed field reads:

$$\mathcal{L}' = (\partial_\mu h)^\dagger(\partial^\mu h) - \frac{1}{2}(-2\mu^2)h^2 - \lambda v h^3 - \frac{1}{4}\lambda h^4. \tag{2.37}$$

The second term or this expression corresponds to the mass term of the $h(x)$ field. In other words, the spontaneous breaking of the local $SU(2)$ symmetry has induced a particle of the mass:

$$m_h = \sqrt{-2\mu^2}, \tag{2.38}$$

with $\mu^2 < 0$. In addition, it can also be deduced that the Higgs field undergoes self-interactions represented by the latter two terms of the Lagrangian 2.37.

Furthermore, this procedure has also generated the masses of the gauge bosons. This becomes obvious from the development of $|D_\mu \phi'|^2$, which provides terms of the form:

$$\frac{g^2}{8}\begin{pmatrix} 0 & v \end{pmatrix}\sigma_i\sigma^j\begin{pmatrix} 0 \\ v \end{pmatrix}W^i_\mu W^{j\,\mu}, \tag{2.39}$$

which yield the mass terms of the gauge bosons:

$$\mathcal{L}_i^{mass} = \frac{g^2 v^2}{8}W^i_\mu W^{i\,\mu} \equiv \frac{1}{2}m_W^2 W^i_\mu W^{i\,\mu}. \tag{2.40}$$

This expression implies that all three gauge bosons have the same masses given by:

$$m_W = \frac{gv}{2}. \tag{2.41}$$

To summarise, the discussion of this section has shown that spontaneous breaking of the local $SU(2)$ symmetry has provided one massive self-interacting scalar boson.

2.2 The Standard Model of Particle Physics

2.2.1 Electromagnetic Interaction

The simplest SM interaction is the (abelian) electromagnetic interaction, which is described by quantum electrodynamics (QED). The gauge invariance of electromagnetism was already introduced in Maxwell's formulation of electrodynamics [12]. It was later extended to the invariance under rotations of the complex internal phase, which accommodated the explanation of the interaction in terms of QFT.

In the language of the discussion from the previous section, the group of the symmetry transformations of the electromagnetic interaction is the $U(1)$ group, with the identity $\mathbb{I}$ as the only generator. The local transformations of this group have the form:

$$\psi(x) \to \psi'(x) = e^{i\alpha(x)}\psi(x). \tag{2.42}$$

The gauge field introduced to reinforce the invariance of the Lagrangian under such a rotation is the massless photon field A_μ, leading to the formulation of covariant derivative as $D_\mu = \partial_\mu - ieA_\mu$, with the coupling constant e corresponding to the elementary electric charge (i.e. the electic charge of an electron). Furthermore, since $[\mathbb{I}, \mathbb{I}] = 0$, all the structure constants are equal to zero and the group is abelian. Therefore, the field tensor:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{2.43}$$

contains no self-interaction terms of the gauge field.

The final QED Lagrangian for a free Dirac field then reads as:

$$\mathcal{L}_{QED} = \bar{\psi}\left(i\gamma^\mu\partial_\mu - m\right)\psi - e\bar{\psi}\gamma^\mu A_\mu\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{2.44}$$

2.2.2 Weak Interaction

The weak interaction was first identified in the β-decays of atomic nuclei in the late 19th century. The experiments regarding these decays also lead to the postulation of neutrinos by Wolfgang Pauli [13], introduced to ensure the energy, momentum and spin conservation in these processes. After the parity violation of the weak processes was confirmed in the experiment conducted by Wu [14], the Fermi formulation of the weak interaction theory as a 4-body contact interaction [15] was abandoned, and a new gauge theory formulation was constructed. The Wu experiment showed that the decay particles were traveling in a direction opposite to their spin, which suggested that the weak interaction had a form of a difference of a vector (representing the particle momentum) and an axial vector (representing the particle spin), i.e. a V-A form. The weak interaction is thus not the same for a particle and its mirror symmetric partner. This feature is referred to as the *parity violation* (P-violation) of the weak interaction. Antiparticles, however, were later shown to have their momenta aligned with their spins, meaning that the parity transformation (P) combined with the charge conjugation (C) is a symmetry of the weak interaction - CP-conservation. This inspired the introduction of the co-called *chirality operators* P_L and P_R, which project a spinor into its V-A component, called *left-handed*, and V+A, called *right-handed*, components, in the following way:

$$P_L\psi = \frac{1}{2}\left(1 - \gamma^5\right)\psi, \quad P_R\psi = \frac{1}{2}\left(1 + \gamma^5\right)\psi, \tag{2.45}$$

with the matrix $\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3$.

The first formulation of the β-decay theory in analogy to electrodynamics was provided by Enrico Fermi in 1934. The process which occurs in β-decay is a decay of a neutron into a proton, electron and an electron anti-neutrino: $n \to pe^-\bar{\nu}_e$. According to the quark model [16–18], the

neutron and proton are baryons, i.e. bound states consisting of three quarks. As a neutron has the configuration $n = (udd)$, and a proton $p = (uud)$, on the quark level the β-decay process reads as $d \to ue^-\bar{\nu}_e$. Fermi described this process as an interaction of two currents: $(j^\mu)_u$ describing the $d \to u$ transition, and $(j^\mu)_e$ describing the $\nu_e \to e^-$ transition, where the final state anti-neutrino was treated as an initial state neutrino. These currents are referred to as *charged currents*, as they modify the electric charge of the initial state. This process can be understood as an exchange of a boson with $Q = -1$. Similar processes carried by a boson with $Q = +1$ have also been observed, suggesting that the weak interaction must be carried by two gauge bosons: the W^+ with $Q = +1$, and the W^- with $Q = -1$.

The propagator of a weak gauge boson was first formulated analogously to the (massless) photon propagator as:

$$iG^{\mu\nu}(q^2) = \frac{-ig^2 g^{\mu\nu}}{q^2}, \tag{2.46}$$

with the weak coupling constant g and the metric tensor $g^{\mu\nu}$ defined as $g^{00} = 1$, $g^{ii} = -1$ for $i \neq 0$, and $g^{ij} = 0$ for $i \neq j$. However, the fact that the weak interaction was observed only as a short range interaction suggested that it should be mediated by a massive particle. With this assumption, the propagator becomes:

$$iG^{\mu\nu}(q^2) = \frac{-ig^2\left(g^{\mu\nu} + \frac{q^\mu q^\nu}{M_W^2}\right)}{q^2 - M_W^2}. \tag{2.47}$$

In the low momentum limit, where $q^2 \ll M_W^2$, this propagator reduces to $-ig^2/M_W^2$, transforming the weak current interaction into a point interaction, explaining the weakness of the interaction by the high mass of the $W^\pm$ bosons.

Early experiments have shown that charge currents conserve the so called *lepton number* defined as:

$$L_l = \begin{cases} +1, & \text{for } l^-,\ \nu_l \\ -1, & \text{for } l^+,\ \bar{\nu}_l \\ 0, & \text{for } l' \end{cases} \tag{2.48}$$

,

where l stands for the lepton flavour in question, while l' stands for other lepton flavours. This depicts the experimental fund that the weak interaction never mixes the lepton flavour, assuming massless neutrinos. However, more recent experiments have confirmed the lepton number violation in observed neutrino flavour oscillations [19]. This observation cannot be explain by the SM, as it implies non-zero masses of neutrinos, and as no mechanism leading to such oscillations is present in the SM.

On the other hand, in the quark sector, after the discovery of the second and third generation of quarks, it has been observed that the charged current interactions between the generations was possible, but suppressed compared to the interaction within the generation. This phenomenon has been mathematically suppressed by introducing the weak interaction eigenstates d', s' and b' for the down-type ($Q = -1/3$) quarks, mixing via weak charged currents with the mass

eigenstates of the up-type ($Q = +2/3$) quarks u, c and t respectively. The d', s' and b' states represent linear combinations of the mass eigenstates d, s and b, which can be expressed using the so-called CKM[3] matrix [20, 21] U:

$$\begin{bmatrix} d' \\ s' \\ b' \end{bmatrix} = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{tb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} \begin{bmatrix} d \\ s \\ b \end{bmatrix} . \tag{2.49}$$

The magnitudes of the CKM matrix elements are measured experimentally. The averaged values combining the currently best measurements are [4]:

$$\begin{bmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{bmatrix} = \begin{bmatrix} 0.97425 \pm 0.00022 & 0.2252 \pm 0.0009 & (4.15 \pm 0.49) \times 10^{-3} \\ 0.230 \pm 0.011 & 1.006 \pm 0.023 & (40.9 \pm 1.1) \times 10^{-3} \\ (8.4 \pm 0.6) \times 10^{-3} & (42.9 \pm 2.16) \times 10^{-3} & 0.89 \pm 0.07 \end{bmatrix} \tag{2.50}$$

Following the symmetry of quarks under internal phase transformations and requiring the unitarity of the CKM matrix, it can be shown that a $N \times N$ CKM matrix describing N quark generations has $(N-1)(N-2)/2$ complex phases. In case of the SM, where three generations are known, this implies the existence of one complex phase. It can be shown that due to this complex phase the weak interaction does not remain exactly invariant under the CP-transformation, as previously believed. This phenomenon, called CP-violation of the weak interaction, was first observed in the decays of the neutral kaons [22].

Following this idea of the mixing between the generations and looking back at the lepton sector, the absence of the observed flavour mixing in the charged lepton sector can be explained by very low neutrino masses, which make linear combinations of the neutrino mass eigenstates very hard to observe.

Regarding their participation in weak charged currents, fermions are arranged in doublets of left-handed particles, transforming in each other through weak charged currents, and right-handed singlets, composed of particles that do not undergo weak charged currents (Table 2.2). Right-handed neutrinos have never been observed. The structure of the left-handed doublets can mathematically be represented in the same way as the spin-1/2 particles. Thus, these multiplets are assigned a quantum number called *weak isospin* I, which is equal 1/2 for doublets and 0 for singlets. The particles within a doublet differ in the third component of the weak isospin, I_3, which is equal $+1/2$ for the neutrinos and the up-type quarks, and $-1/2$ for charged leptons and down-type quarks. The rotations within a left-handed doublet can thus be mathematically expressed by a rotation symmetry group of a spin-1/2 particle, i.e. by the $SU(2)$ symmetry group, called $SU(2)_L$ in this case, to emphasise that it acts only on left-handed states. In this sense, the charge-rising and charge-lowering currents can be written as the step-up (τ_+) and step-down (τ_-) operators defined as $\tau_\pm = 1/2(\sigma_1 \pm \sigma_2)$, using the Pauli matrices. However, the $SU(2)$ group has a dimension 3, so it needs a third generator along with $\tau_\pm$, as well as a third gauge boson.

In 1973, neutral currents were observed in the processes of neutrino scattering off an electron or a nucleus. This brought a different kind of interaction to the particle physics scene. This

[3] CKM stands for the initials of the three physicists, Cabibbo, Kobayashi and Maskawa, who developed the concept of the mixing between the weak interaction eigenstates and mass eigenstates in the quark sector.

interaction could not be carried by photons, which do not couple to neutral particles. It was also found that the event rates of neutrino scatterings were very similar to those of the weak interaction. This led to an introduction of a neutral weak boson, given the name W^0, a candidate to complete the fundamental triplet of weak gauge bosons. However, it didn't completely respect the $SU(2)_L$. Contrary to the pure $V-A$ nature of the charged currents, the neutral currents were found to also have a small right-handed component, which increases the similarity of the weak neutral currents and electromagnetic currents. In addition, in the limit of very high momentum transfer q^2, the mass effect in the propagator $G^{\mu\nu}_{W^0}$ (Equation 2.47) vanishes, so that the propagator becomes analogous to that of the photon (Equation 2.46). It has further been found experimentally that the event rates of weak neutral and electromagnetic currents are comparable in the high q^2 limit, suggesting that the electromagnetic and weak coupling constants should be similar. This strengthens the idea of constructing a common theory of the electromagnetic and weak interaction, which is discussed in the following section.

	Fermion multiplets			I	I_3	Y	Q
Leptons	$\begin{pmatrix}\nu_e\\ e\end{pmatrix}_L$	$\begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix}_L$	$\begin{pmatrix}\nu_\tau\\ \tau\end{pmatrix}_L$	$1/2$	$+1/2$ $-1/2$	-1	0 -1
	e_R	μ_R	τ_R	0	0	-2	-1
Quarks	$\begin{pmatrix}u\\ d'\end{pmatrix}_L$	$\begin{pmatrix}c\\ s'\end{pmatrix}_L$	$\begin{pmatrix}t\\ b'\end{pmatrix}_L$	$1/2$	$+1/2$ $-1/2$	$+1/3$	$+2/3$ $-1/3$
	u_R	c_R	t_R	0	0	$+4/3$	$+2/3$
	d_R	s_R	b_R	0	0	$-2/3$	$-1/3$

Table 2.2: Multiplets of the electroweak interaction [7].

2.2.3 Electroweak Unification

To mathematically unify the weak and the electromagnetic properties of the weak neutral currents, a new quantity called *hypercharge* Y was defined, connecting the electromagnetic charge Q and the third component of the weak isospin I_3 by the Gelmann-Nishijima [23, 24] equation:

$$Q = I_3 + \frac{Y}{2}. \tag{2.51}$$

In analogy to the electromagnetic symmetry group $U(1)$ generated by Q, the hypercharge Y generates the symmetry group $U(1)_Y$, under which the hypercharge current remains invariant. The symmetry group governing the unified interaction called *electroweak interaction* becomes therefore $SU\left(2\right)_L \times U\left(1\right)_Y$. To each subgroup corresponds a separate coupling constants: g for $SU\left(2\right)_L$ and g' for $U\left(1\right)_Y$.

With the symmetry and the coupling constants at hand, a gauge theory can be developed. A fermionic field $\psi\left(x\right)$ transforms then as:

$$\psi\left(x\right) \rightarrow \psi'\left(x\right) = e^{\frac{i}{2}\vec{\sigma}\cdot\vec{\alpha}} e^{i\frac{Y}{2}\beta} \psi\left(x\right). \tag{2.52}$$

The $SU(2)_L \times U(1)$ covariant derivative becomes:

$$D_\mu = \left(\partial_\mu - i\frac{g}{2}\sigma \cdot \mathbf{W}_\mu - i\frac{g'Y}{2}B_\mu\right), \tag{2.53}$$

requiring the introduction of the gauge fields:

- $\vec{W}_\mu = \left(W_\mu^1, W_\mu^2, W_\mu^3\right)$, as the $SU(2)_L$-gauge field, transforming as $\mathbf{W}_\mu \to \mathbf{W}_\mu - \frac{1}{g}\partial_\mu \boldsymbol{\alpha} - \boldsymbol{\alpha} \times \mathbf{W}_\mu$, corresponding to the interaction with the weak coupling constant g and the field strength tensor $W = \partial_\mu \mathbf{W}_\nu - \partial_\nu \mathbf{W}_\mu - g\mathbf{W}_\mu \times \mathbf{W}_\mu$[4],
- B_μ, as the $U(1)_Y$-gauge boson, transforming as $B_\mu \to B_\mu - \frac{1}{g'}\partial_\mu \alpha'$, corresponding to the interaction with the weak coupling constant g', and the field strength tensor $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$.

With the covariant derivative 2.53, the Lagrangian of the electroweak interaction becomes:

$$\mathcal{L}_{elw} = \mathcal{L}_{kin} + \mathcal{L}_L + \mathcal{L}_R, \tag{2.54}$$

with:

$$\mathcal{L}_{kin} = -\frac{1}{4}W_{\mu\nu} \cdot W^{\mu\nu} - \frac{1}{4}B_{\mu\nu} \cdot B^{\mu\nu}, \tag{2.55}$$

$$\mathcal{L}_L = \bar{\psi}_L \left(i\partial_\mu - i\frac{g}{2}\sigma \cdot \mathbf{W}_\mu - i\frac{g'Y}{2}B_\mu\right)\psi_L, \tag{2.56}$$

$$\mathcal{L}_R = \bar{\psi}_R \left(i\partial_\mu - i\frac{g'Y}{2}B_\mu\right)\psi_R. \tag{2.57}$$

The terms $\mathcal{L}_L$ and $\mathcal{L}_R$ describe the dynamics of the left-handed and right-handed Dirac fields respectively. The $\mathcal{L}_L$ term can be split into a charged $\mathcal{L}_L^{CC}$ component, composed of the terms containing W_μ^1 and W_μ^1, and a neutral component $\mathcal{L}_L^{NC}$, combining the terms containing W_μ^3 and B_μ. Defining the positive and negative weak gauge bosons as:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp iW_\mu^2\right), \tag{2.58}$$

and writing an arbitrary $SU(2)_L$ multiplet as:

$$\psi_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \tag{2.59}$$

[4]This expression follows from the equation 2.26 using the structure constants given by the three dimensional Levi-Civita Tensor:

$$\epsilon_{ijk} = \begin{cases} +1 & \text{if } ijk \text{ is a cyclical permutation} \\ -1 & \text{if } ijk \text{ is a anti-cyclical permutation} \\ 0 & \text{otherwise} \end{cases},$$

deduced from the commutation rules for the Puli matrices.

the $\mathcal{L}_L^{CC}$ terms obtains the shape:

$$\mathcal{L}_L^{CC} = -\frac{1}{\sqrt{2}} g \left(\bar{u}_L \gamma^\mu W_\mu^+ d_L + \bar{d}_L \gamma^\mu W_\mu^- u_L \right). \tag{2.60}$$

The remaining neutral component $\mathcal{L}_L^{NC}$ represents a mixture of the physical Z_μ field, carrying the weak neutral currents, and the photon field A_μ. As discussed in the previous section, the Z_μ field cannot be interpreted solely by W_μ^3. Thus, the fields Z_μ and A_μ are defined as linear combinations of W_μ^3 and B_μ given by:

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{bmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{bmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}, \tag{2.61}$$

using the so-called *Weinberg angle* θ_W. The neutral component of $\mathcal{L}_L$ then reads as:

$$\begin{aligned} \mathcal{L}_L^{NC} &= -\bar{\psi}_L \gamma^\mu (g I_3 W_\mu^3 + g' \frac{Y}{2} B_\mu) \psi_L \\ &= -\bar{\psi}_L \gamma^\mu (g \sin\theta_W I_3 + g' \frac{Y}{2} \cos\theta_W) A_\mu \psi_L \\ &\quad -\bar{\psi}_L \gamma^\mu (g \cos\theta_W I_3 + g' \frac{Y}{2} \sin\theta_W) Z_\mu \psi_L. \end{aligned} \tag{2.62}$$

From the last equation, the Weinberg angle θ_W can be calculated. Demanding the coupling constant corresponding to the A_μ field to be equal to the electric charge Qe of the ψ_R field, and using Equation 2.51, the following relations imply:

$$g \sin\theta_W = g' \cos\theta_W = e. \tag{2.63}$$

This shows that the gauge theory based on the symmetry group $SU(2)_L \times U(1)_Y$ explains all physical gauge boson fields of the electromagnetic and weak interactions, and is thus a valid theory that unifies these two interactions. However, to conserve the gauge invariance of the Lagrangian $\mathcal{L}_{elw}$ of the unified theory, these gauge bosons are required to be massless. They can, nonetheless, obtain masses through the spontaneous symmetry breaking, as explained in Section 2.1.3. The mechanism that generates masses of the electroweak gauge bosons is described in the next section.

2.2.4 The BEH Mechanism

In Section 2.1.3 it has been shown that by introducing a symmetry breaking mechanism, the local gauge invariance of the Lagrangian can be conserved while at the same time the gauge bosons can be given masses. The broken symmetry of the SM is $SU(2)_L \times U(1)_Y$, i.e. that governing the weak interaction, which is transmitted with massive gauge bosons. This mechanism is referred to as Brout-Englert-Higgs (BEH) mechanism [25–30]. The spontaneous symmetry breaking is induced by introducing a scalar Higgs field. This field, however, must have the form of a $SU(2)_L \times U(1)_Y$ multiplet in order to satisfy the SM gauge invariance. In the so

called "minimal" Higgs sector, the Higgs field is an $SU(2)_L \times U(1)_Y$-isodoublet with the isospin $I = 1/2$ and the hypercharge $Y = 1$:

$$\phi = \begin{pmatrix} \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \\ \frac{1}{\sqrt{2}}(\phi_3 + i\phi_4) \end{pmatrix} \equiv \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \tag{2.64}$$

containing one positively charged component ϕ^+ and one neutral component ϕ^0. The dynamics of this field is given by the Lagrangian 2.29, with the Higgs potential defined in Equation 2.30 with the minimum defined by Equation 2.33. Any solution of this equation which breaks the symmetry operation inevitably leads to mass generation of the corresponding gauge boson. However, the solution given by Equation 2.35 is convenient, as it does not break the $SU(2)_L \times U(1)_Y$ symmetry completely. While it does break the $SU(2)_L$ and the $U(1)_Y$, given that ϕ_0 is neutral, the $U(1)_{em}$ remains a symmetry of the SM, and the corresponding gauge boson, photon, remains massless. It is thus said that the $SU(2)_L \times U(1)_Y$ symmetry is broken down to $U(1)_{em}$.

2.2.4.1 Gauge boson masses

The mass terms of the bosons can be obtained from:

$$\begin{aligned} \left| \left(\partial_\mu - i\frac{g}{2}\sigma \cdot \mathbf{W}_\mu - i\frac{g'Y}{2}B_\mu \right) \phi_0 \right|^2 &= \frac{1}{8} \left| \begin{pmatrix} gW_\mu^3 + g'B_\mu & g(W_\mu^1 - iW_\mu^2) \\ g(W_\mu^1 + iW_\mu^2) & -gW_\mu^3 + g'B_\mu \end{pmatrix} \right|^2 = \\ &= \frac{1}{8}v^2g^2(W_\mu^1)^2 + \frac{1}{8}v^2g^2(W_\mu^1)^2 + \\ &+ \frac{1}{8}v^2[gW_\mu^3 - g'B_\mu]^2 + 0[g'W_\mu^3 + gB_\mu]^2. \end{aligned} \tag{2.65}$$

With the definitions from Equations 2.58 and 2.61, and using the relations from Equation 2.63, the mass terms of Equation 2.66 correspond to the physical bosons $W_\mu^\pm$, Z_μ and A_μ with the masses given by:

$$M_W = \frac{1}{2}vg, \tag{2.66}$$

$$M_Z = \frac{1}{2}v\sqrt{g^2 + g'^2}, \tag{2.67}$$

$$M_A = 0. \tag{2.68}$$

In this form the fields correspond to the physical gauge fields. $W_\mu^\pm$ correspond to $W^\pm$ bosons, Z_μ to the Z-boson and A_μ to the photon. These formulas imply that the Weinberg angle can be determined as:

$$\cos\theta_W = \frac{M_W}{M_Z}. \tag{2.69}$$

Using this, the vev of the Higgs field was determined as:

$$v = \frac{2M_W}{g} \approx 246 \ \ \text{GeV}. \tag{2.70}$$

2.2.4.2 Fermion masses

In the BEH mechanism, the fermion masses are generated by coupling the left- and right-handed representation of one family via the Higgs field through the so-called Yukawa interaction. The masses of all massive fermions are generated in the analogous way.

The Lagrangian mass term for the lepton l can be derived from:

$$\begin{aligned} \mathcal{L}_l &= -y_l \left(\begin{pmatrix} \bar{\nu}_l & \bar{l} \end{pmatrix}_L \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v+h \end{pmatrix} l_R + \bar{l}_R \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & v+h \end{pmatrix} \begin{pmatrix} \nu_l \\ l \end{pmatrix}_L \right) \\ &= -\frac{y_l}{\sqrt{2}} v \left(\bar{l}_L l_R + \bar{l}_R l_L \right) - \frac{y_l}{\sqrt{2}} \left(\bar{l}_L e_R + \bar{l}_R l_L \right) h \\ &= -m_l \bar{l} l - \frac{m_l}{v} \bar{l} l h, \end{aligned} \tag{2.71}$$

with the Yukawa coupling strength y_l of the lepton l to the Higgs field, with which the lepton mass becomes:

$$m_l = \frac{y_l v}{\sqrt{2}}. \tag{2.72}$$

The first term of Equation 2.72 is the mass term corresponding to the lepton l, while the second term describes its interaction with the Higgs field. Quark masses and interactions with the Higgs field can be derived in an analogous way.

While the gauge boson masses can be determined from the known values of the coupling constants g and g', the fermion masses are the free parameters of the SM, since their Higgs Yukawa couplings y_f (f stands for any massive elementary fermion) are not predicted by the SM.

2.2.5 Strong Interaction

The necessity to introduce the strong interaction arose from the need to explain what keeps protons together in nuclei and quarks in hadrons. The postulation of colour as the elementary strong charges was motivated by the need to acquire the accordance of the Δ^{++} baryon with the Fermi statistics. Being a spin-3/2 fermion composed of three identical quarks (uuu), this state needed an additional feature which would make it asymmetric under a rotation of the quarks. At the same time, an explanation was needed for the experimental finding that only the states of the form qqq, $\bar{q}\bar{q}\bar{q}$ (baryons) and $q\bar{q}$ (mesons) had been observed[5], but no free quarks nor qq states. The proposed solution was the existence of three colours: red (R), blue (B) and green (G). A quark can carry one colour at the time. Anti-quarks carry anti-colours. Only

[5]The LHCb collaboration has recently reported the observation of a tetra-quark resonance [31]

colour singlet states, i.e. such carrying either all three colours (as baryons) or a colour-anti-colour combination (as mesons), exist as bound states. The carriers of the strong interaction, called *gluons*, are massless spin-1 bosons, carrying one of eight bi-colour combinations, i.e. a superposition of colour-anti-colour states, but no electric charge. Carrying the strong charge, gluons can interact among themselves, which makes the strong interaction non-abelian.

The strong interaction is described by quantum chromodynamics (QCD). The base states, R, G and B, are represented as vectors:

$$\psi_r = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \psi_b = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \psi_g = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{2.73}$$

The group of rotations in the colour space is the symmetry group $SU(3)_C$, generated by $3^2 - 1 = 8$ Gell-Mann matrices λ^a[6], which are 3×3 generalizations of the Pauli matrices. The commutation law for Gell-Mann matrices is $\left[\frac{\lambda^a}{2}, \frac{\lambda^b}{2}\right] = if^{abc}\frac{\lambda^c}{2}$, with the $SU(3)$ group structure constants:

$$\begin{aligned} f^{123} &= 1, \\ f^{458} &= f^{678} = \sqrt{3}/2, \\ f^{147} &= f^{165} = f^{246} = f^{257} = f^{345} = f^{376} = \frac{1}{2}. \end{aligned} \tag{2.74}$$

QCD describes three quark colour fields q_j. A colour field transforms under a $SU(3)$-transformation as:

$$\psi(x) \to \psi'(x) = e^{\frac{i}{2}\alpha_a(x)\lambda^a}\psi(x) \tag{2.75}$$

The preservation of local $SU(3)$ symmetry requires the introduction of eight gauge fields G^a_μ which transform as $G^a_\mu \to G^{a\,\prime}_\mu = G^a_\mu - \frac{1}{g_s}\partial_\mu\alpha_a$, with the strong coupling constant g_s. These eight fields G^a_μ correspond to eight gluon colour-states.

The induced covariant derivative is then:

$$D_\mu \equiv \partial_\mu + \frac{i}{2}g_s\lambda^a G^a_\mu - f^{abc}\alpha_b G^c_\mu. \tag{2.76}$$

The colour field tensor, which also contains the gauge field self-interaction term, has the form:

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f^{abc} G^b_\mu G^c_\nu, \tag{2.77}$$

[6]The Gell-Mann matrices:

$$\lambda_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \lambda_2 = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \lambda_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \lambda_4 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\lambda_5 = \begin{bmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{bmatrix}, \lambda_6 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \lambda_7 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix}, \lambda_8 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

with which the gauge invariant QCD Lagrangian for a quark field of the flavour q with the mass m_q becomes:

$$\mathcal{L}_{QCD} = \bar{\psi}\left(i\gamma^\mu \partial_\mu - m_q\right)\psi - \frac{g_s}{2}\bar{\psi}\gamma^\mu \lambda^a G^a_\mu \psi - \frac{1}{4} G^a_{\mu\nu} G^{a\,\mu\nu}. \tag{2.78}$$

2.2.6 Renormalisation and Running Coupling Constants

Computation of physical observables of a theory at tree level, i.e. including only bare initial and final state and propagators, does not reflect the reality accurately, as each field can exhibit virtual fluctuations. According the the Ward identity [32], the fluctuations of the initial and final state fields cancel out with the loop fluctuations between different fields (so-called vertex corrections), so that only the loop fluctuations of the propagator enter the computation. As the virtual particles in the loop can have any momentum, loop contributions need to be integrated over all possible momenta, from zero to infinity. This brings divergences into the computation of the observables, which is not physical. It can be shown that the loop corrections can be absorbed in the redefined coupling constant $\alpha(\mu_R)$ of the theory, and an arbitrary observable can be developed as a power series in α.

The problem of divergences is solved by introducing a cutoff, i.e. an energy scale up to which the observables are calculated. The observables then become finite, but dependent on that cutoff. This procedure is referred to as *renormalisation* and the cutoff is called *renormalisation scale* μ_R. As physical observables other than the coupling constant do not seem to depend on any unphysical scale, the so called *Renormalisation Group Equation* (RGE):

$$\left(\mu_R^2 \frac{\partial}{\partial \mu_R^2} + \beta(g)\frac{\partial}{\partial g}\right) R = 0 \tag{2.79}$$

needs to hold for an arbitrary observable R, with g being the charge of the theory. The β-function can be expanded as:

$$\beta(\alpha) = -\beta_0 \alpha^2 - \beta_1 \alpha^3 - \ldots \tag{2.80}$$

The coefficient β_i depend on the theory. Using the RGE and considering only 1-loop corrections (i.e. the first order of the expansion), the coupling constant α at an arbitrary scale Q^2 becomes:

$$\alpha(Q^2) = \frac{\alpha(\mu_R^2)}{1 + \alpha(\mu_R^2)\beta_0 \ln \frac{Q^2}{\mu_R^2}} \tag{2.81}$$

for a given renormalisation scale μ_R. This particular equation is the place where the qualitative differences between QED and QCD become obvious. With the corresponding β_0 coefficients, the QED coupling constant α_{em} as a function of the transferred momentum Q^2 becomes:

$$\alpha_{em}(Q^2) = \frac{\alpha_{em}(\mu_R^2)}{1 - \frac{\alpha_{em}(\mu_R^2)}{3\pi} \ln \frac{Q^2}{\mu_R^2}}. \tag{2.82}$$

This shows that the α_{em} slowly grows with Q^2. For all currently accessible energies, however, this growth appears negligible.

In the QCD case, the situation changes dramatically. With the QCD β function, the development of the strong coupling constant α_s reads as:

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu_R^2)}{1 + \frac{\alpha_s(\mu_R^2)}{12\pi}(33 - 2n_f)\ln\frac{Q^2}{\mu_R^2}}, \tag{2.83}$$

with n_f quark flavours lighter than μ_R. The sign in the denominator remains positive up to $n_f = 16$, which is far beyond the number of the currently known SM quarks. Thus, contrary to α_{em}, α_s decreases with increasing Q^2, and vanishes for $Q^2 \to \infty$ (i.e. very short distances). This phenomenon is known as the *asymptotic freedom* and directly follows from the self-interaction of gluons.

Equation 2.83 splits the Q^2 range in two parts. Below a certain scale Λ defined as:

$$\Lambda^2 = \mu_R^2 \exp\left(\frac{-12\pi}{(33 - 2n_f)\alpha_s(\mu_R^2)}\right), \tag{2.84}$$

the strong coupling constant α_s is too high for expressing physical observables as power expansions in α_s. This value of the Λ scale, experimentally estimated to be ~ 200 MeV, splits the QCD into a perturbative regime above Λ, and non-perturbative regime below it. Λ is not predicted by theory, so it is a free parameter of QCD.

In the perturbative regime quarks and gluons can be considered independent from each other. The physics of this regime is described by the Parton Model [33]. On the other hand, at a larger distance, where the strong coupling becomes larger, it is energetically more favorable for quarks to stay bounded within composite states. Thus, quarks cannot be found as free particles, which is known as *confinement*. The non-perturbative regime or QCD is described by phenomenological models, as explained in the next chapter.

2.2.7 The Full Standard Model Lagrangian

Combining the results of the previous two sections, the SM emerges as a theory combining the electroweak and strong interactions through a gauge theory with the underlying symmetry group $SU(3)_c \times SU(2)_L \times U(1)_Y$. The total Lagrangian describing the dynamics of a fermionic field ψ with the mass m under the SM interactions is:

$$
\begin{aligned}
\mathcal{L}_{SM} &= -\frac{1}{4}\sum_{i=1}^{3} W^i_{\mu\nu} W^{i\,\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4}\sum_{i=1}^{8} G^i_{\mu\nu} G^{i\,\mu\nu} && \left\{ \begin{array}{l} \text{gauge fields} \\ \text{kinematics} \end{array} \right. \\
&+ \bar{\psi}_L \left(i\partial_\mu - i\frac{g}{2}\sum_{i=1}^{3} \sigma^i \cdot W^i_\mu - i\frac{g'Y}{2} B_\mu \right) \psi_L && \left\{ \begin{array}{l} \text{left-handed} \\ \text{weak currents} \end{array} \right. \\
&+ \bar{\psi}_R \left(i\partial_\mu - i\frac{g'Y}{2} B_\mu \right) \psi_R && \left\{ \begin{array}{l} \text{left-handed} \\ \text{weak currents} \end{array} \right. \\
&- \frac{g_s}{2}\sum_{a=1}^{8} \bar{\psi}\gamma^\mu \lambda^a G^a_\mu \psi && \{\text{strong interaction} \\
&- \frac{1}{2} M_H^2 (v+h)^2 - \frac{M_H^2}{v^2}(v+h)^4 && \left\{ \begin{array}{l} \text{Higgs mass and} \\ \text{self-interaction} \end{array} \right. \\
&+ \left| \left(\partial_\mu - i\frac{g}{2}\sum_{i=1}^{3} \sigma^i \cdot W^i_\mu - i\frac{g'Y}{2} B_\mu \right) (v+h) \right|^2 && \left\{ \begin{array}{l} \text{gauge fields mass,} \\ \text{self-interaction and} \\ \text{Higgs coupling} \end{array} \right. \\
&- m\bar{\psi}\psi - \frac{m}{v}\bar{\psi}\psi h && \left\{ \begin{array}{l} \text{fermion mass and} \\ \text{Higgs coupling} \end{array} \right.
\end{aligned}
\tag{2.85}
$$

Among the SM fermions, only quarks exhibit all interactions presented in Equation 2.86. Charged leptons undergo only the electromagnetic and the weak interactions, while neutrinos only the weak interaction.

This lengthy but elegant formula, able to explain all fundamental interactions of fermionic fields except for gravity, and describe all particle physics data remarkably well, is, however, determined by 19 parameters, which are not predicted by the SM. The parameters comprise [34]:

- three coupling parameters, usually expressed in terms of the electromagnetic and strong coupling constants α_{em} and α_s, and the weak mixing angle $\sin^2\theta_W$,
- two parameters defining the Higgs potential, μ^2 and λ, or equivalently m_h and m_W or m_Z,
- six quark masses, or equivalently six Yukawa couplings of the quarks to the Higgs field,
- three masses of charged leptons, or the corresponding Yukawa couplings,
- four parameters for the CKM matrix: three mixing angles and one CP-violating phase,
- one parameter to accommodate non-perturbative CP violation in QCD.

To accommodate newly established neutrino masses [35], an additional three parameters for the masses and four for the mixings [36] are required.

This non-negligible number of parameters suggests that there might be a more fundamental theory explaining these parameters and embracing the SM as a special case. This idea is further elaborated in the next section.

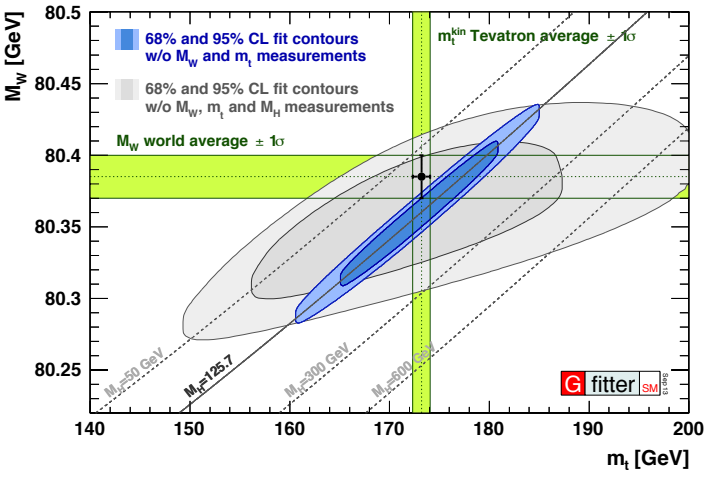


Figure 2.2: Results of the global electroweak fit compared to the measurements. Contours of 68% and 95% confidence level obtained from scans of fits with fixed variable pairs M_W (W-boson mass) vs. m_t (top quark mass). The narrower blue and larger grey allowed regions are the results of the fit including and excluding the M_H (Higgs boson mass) measurements, respectively. The horizontal bands indicate the 1σ regions of the M_W and m_t measurements (world averages). [48].

2.2.8 Experimental tests and limitations of the Standard Model

The SM has so far shown a remarkable success. It is a mathematically consistent theory accommodating most of the experimental findings (see Figure 2.2 for an illustration), with an excellent prediction power due to its renormalisability (calculability). As an example, the existence of the top quark was predicted, and its mass inferred from the radiative corrections to the masses of the weak gauge bosons [37] long before its discovery. One of the most impressive predictions was that of the Higgs boson mass obtained by combining the LEP and Tevatron exclusions with electroweak precision data from 2011, prior to the discovery [38]:

$$M_H = 120^{+12}_{-5} \text{ GeV}. \tag{2.86}$$

The discovery of a Higgs boson with the mass close to 125 GeV by the ATLAS [5] and CMS [6] collaborations which followed in 2012 was undoubtedly the crown of the experimental establishment of the SM. The measurements of the Higgs boson mass [39, 40], its couplings to W- and Z-bosons [40–43], spin-parity quantum numbers [40, 41, 43, 44], as well as the recent evidences of the decays into fermions [45–47], gradually strengthened the compatibility of the observed new boson with the expected agent of the electroweak symmetry breaking. However, at the time of writing this thesis, studies of the nature of the Higgs boson are still at an early stage.

Despite its tremendous success, a handful of theoretical arguments suggests that the SM is not the final theory of the matter in the Universe, but more likely a low energy manifestation, i.e. an effective theory, of a more general theory.

The SM has no explanation of why there are exactly three generations of chiral fermions and why their masses are so diverse, i.e. which mechanism generates their Yukawa couplings. It also does not contain any mechanism for generating the neutrino masses. Furthermore, it provides no explanation of the astronomical observations such as the presence of the dark matter or the baryon (i.e. matter-antimatter) asymmetry in the Universe. In addition, the success of the electroweak unification rises the question whether the electroweak and strong interactions perhaps represent the low energy regime of a single interaction, described by a Grand Unified Theory (GUT), which acts as unique above a certain energy scale, the so-called GUT scale, far beyond the current energy frontiers (estimated to $\sim 10^{15}$ GeV). Finally, an ultimate

theory of interactions, the so-called Theory of Everything, would have to incorporate gravity as well. It would include an explanation of the gravity in a quantum field theory framework. Some potential solutions to the latter issue are proposed by the string theory [49] and quantum gravity [49]. These theories are expected to gain importance around the so-called Planck scale[7], $\sim 10^{19}$ GeV, where the quantum effects of gravity are expected.

The Higgs mechanism, believed with increasingly strong evidences to govern the spontaneous breaking of the $SU(2)_L \times U(1)_Y$ symmetry, has itself several serious shortcomings [51]. In particular, the SM leaves unexplained how the Higgs boson mass remains close to the electroweak scale despite the quantum corrections ought to drive it up to GUT or the Planck scale. Moreover, the Higgs field, which pervades all of space according to this mechanism, generates a vacuum energy density far higher than what is observed, which brings up the question of the Higgs field vacuum stability.

In the rest of this section the open issues of the SM which are attempted to be addressed by the searches presented in this thesis are discussed more closely.

2.2.8.1 Baryon Asymmetry Problem

According to the Standard Models of particle physics and cosmology, the Universe is expected to consist of an equal number of particles and anti-particles produced out of energy throughout its evolution. However, the observations show that in the observable Universe matter prevades anti-matter by far. This effect can be quantified in terms of the *baryon-to-photon ratio* η, defined as [34]:

$$\eta = \frac{n_b - \bar{n}_b}{n_\gamma} \simeq \frac{n_b}{n_\gamma}, \tag{2.87}$$

with the number densities of baryons, anti-baryons and photons in the Universe n_b, $\bar{n}_b$, and n_γ respectively. This ratio has been estimated from the measurement of the acoustic peaks in the cosmic microwave background (CMB) by the WMAP experiment [52] to be $\sim 6.1 \cdot 10^{-10}$. This finding contradicts the zero expectation for a Universe which has had an inflationary epoch with a conventional thermal equilibrium in its cosmological history. According to Sakharov, to achieve such a particle-antiparticle asymmetry, referred to as the *baryon asymmetry* of the Universe (BAU), three conditions are necessary [53]: violation of the baryon number, microscopic violation of the charge (C) and charge times parity (CP) symmetry, and loss of thermal equilibrium. The latter can naturally occur through the expansion of the Universe, the baryon number violation is possible in the SM only at the non-perturbative level, while the CP violation enters the SM through the weak phase of the CKM matrix (Equation 2.49). While some recent models support the view that the CP violation incorporated in the SM can be sufficient to explain the BAU [54], the conventional view is that it fails to accommodate the observed asymmetry by several orders of magnitude [55], so that new physics is needed to address this problem.

[7] The Planck scale is given by the Planck mass M_{Pl}, at which the gravitational coupling strength becomes ~ 1, i.e. comparable to those of the gauge couplings [50].

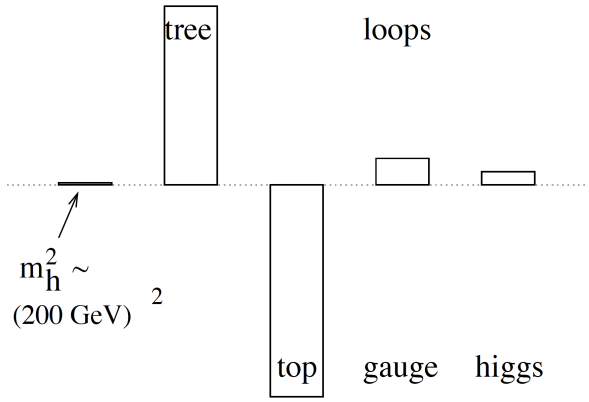


Figure 2.3: A graphical representation of the observed Higgs mass as well as the top quark, gauge and Higgs boson loop correction to it. The horizontal line represents the zero point of the m_h^2 axis growing in the bottom-top direction. The box labeled as "tree" represents the magnitude of a one-loop correction needed to bring the sum of the loop corrections to the observed m_h^2. Taken from [58], which was released before the Higgs boson discovery.

2.2.8.2 Hierarchy Problem, Naturalness and Fine Tuning

The hierarchy problem refers to the concern regarding the huge imbalance between the coupling strengths of the weak interaction (i.e. the weakest gauge coupling) and gravity, expressed as $G_F \gg G_N$, with the Fermi coupling constant G_F and the Newton constant G_N [51]. As $G_F \sim m_W^2$ and $G_N \sim M_{Pl}^2$, with the W-boson mass m_W and the Planck mass $M_{Pl} \sim 10^{19}$ GeV, the hierarchy problem can equivalently be expressed as the problem of $M_{Pl} \gg m_W$. As m_W is determined by the Higgs field vev v, the problem is equivalent to a concern why the Higgs boson mass m_H is so much smaller than M_{Pl}, especially in the face of quadratically divergent quantum loop corrections. Taking into account the corrections from the scalar (spin-0), fermionic (spin-1/2) and vector boson (spin-1) loops, the M_H^2 at a given experimentally achievable momentum scale p^2 reads as:

$$M_H^2(p^2) = M_H^2(\Lambda^2) + \mathcal{C} \int_{p^2}^{\Lambda^2} dk^2 = M_H^2(\Lambda^2) + \frac{\mathcal{C}}{16\pi^2}\Lambda^2, \tag{2.88}$$

with a model dependent coefficient function $\mathcal{C}$, and the reference scale Λ [56]. Considering only the 1-loop contributions, the Higgs boson mass correction becomes determined by the 1-loop coefficient function $\mathcal{C}_1$, which comprises the dominant contributions originating from the virtual $t\bar{t}$ pair production, gauge boson loops and Higgs loops [57]:

$$\mathcal{C}_1 = \lambda^2 + \frac{9}{4}g^2 + \frac{3}{4}g'^2 - 6y_t^2 = \frac{3}{v^2}(M_H^2 + M_Z^2 + 2M_W^2 - 4m_t^2), \tag{2.89}$$

with the Higgs quartic coupling constant λ (from Equation 2.30), the weak coupling constants g and g', the top-Higgs Yukawa coupling y_t, and the corresponding observed particle masses M_H, M_Z, M_W and m_t. The top quark loop makes the largest contribution, since $\lambda_t \sim 1$. As an illustration, for assumed $\Lambda \sim 10$ TeV, the top loop would contribute with $-(2 \text{ TeV})^2$, while the gauge and Higgs boson loops would contribute with $(0.7 \text{ TeV})^2$ and $(0.5 \text{ TeV})^2$ respectively [56], which would leave a negative imbalance of over 3 TeV^2, as shown in Figure 2.3.

The situation gets even more dramatic when moving Λ towards higher scales. For a system described by the SM, a natural reference scale would be the Planck mass, i.e. $\Lambda \sim 10^{19}$ GeV. If a GUT theory was established, the reference scale would be the GUT scale, i.e. $\Lambda \sim 10^{15}$ GeV. In both cases the range of integration is huge, leading to large loop contributions scaling as $\sim \Lambda^2$. However, the experimentally measured Higgs mass is ~ 125 GeV [39, 40], meaning that

either the integrand must be limited, or that $M_H(\Lambda^2)$ must be finely tuned to add up with the integrand of the order of $\sim 10^{38}$ GeV^2 to the observed $\sim 10^4$ GeV^2. The latter option appears rather unnatural, as it assumes a fine tuning of the order 10^{-34}, although fine tuning examples are not foreign to the Nature[8]. The former option requires the introduction of a new reference scale at the TeV scale corresponding to some kind of new physics in order to save the *naturalness* of the theory. Some proposed approaches to solving this problem are presented in the next section.

2.2.8.3 Electroweak Vacuum Stability Problem

Not only does the low Higgs boson mass cause the hierarchy problem discussed in the previous section, it also raises the question of the stability of the Higgs potential at high energy scales. Assuming that the newly discovered scalar boson really is the SM Higgs boson, now that its mass has been measured, the parameters of the effective Higgs potential (Equation 2.30), μ and λ, are fixed:

$$\mu^2_{exp} \approx (89 \text{ GeV})^2 \tag{2.90}$$

$$\lambda_{exp} \approx 0.13. \tag{2.91}$$

These parameters are, however, not necessarily global. They might be energy scale dependent. In such a case, they could be extrapolated to higher energies through RGE running. Being the weight of the fourth power of the Higgs field, the quartic coupling λ has a stronger effect than μ at the large values of the higgs field ($h \gg v$), and has thus been the subject of numerous theoretical studies recently [60–62]. Studying the renormalisation group evolution of λ with two-loop QCD and three-loop Yukawa corrections, it has been found that it diminishes with the growing scale, reaching zero at some critical point and turning negative above it [60]. For an illustration and a further explanation see Figure 2.4.

This critical scale vitally depends on the Higgs boson mass. Requirement of the absolute stability of the Higgs potential up to the Planck scale can be translated into a lower bound on the Higgs boson mass which assures it. This limit has been found to be $M_H > 129.4 \pm 1.8$ GeV, which implies that with the observed Higgs boson mass, the stability of the electroweak vacuum up to the Planck scale is excluded at the confidence level of 98% (2 standard deviations) [60]. Moreover, it has been estimated that data implies the vacuum instability at $10^{11\pm1}$ GeV. These theoretical findings have far reaching implications. Presence of some new physics might rescue the electroweak vacuum ensuring the absolute stability up to the Planck scales. If, however, the vacuum is confirmed to be unstable or metastable, it could imply existence of very exotic forms of physics, such as multiverses or the Big Crunch, i.e. the collapse of the Universe (Figure 2.4).

2.3 BSM Models with Heavy Quarks

Although some recent works, e.g. [57], argue that the SM can be sufficient to explain most of the burning questions of today's particle physics and cosmology, the widely spread belief is

[8] "The Sun and the Moon have radius and distance from the Earth 'tuned' to appear equal in the sky (with a precision of about 5%), for no better reason than producing rare and spectacular eclipses (and permitting us to test general relativity)" - Gian Giudice [59]

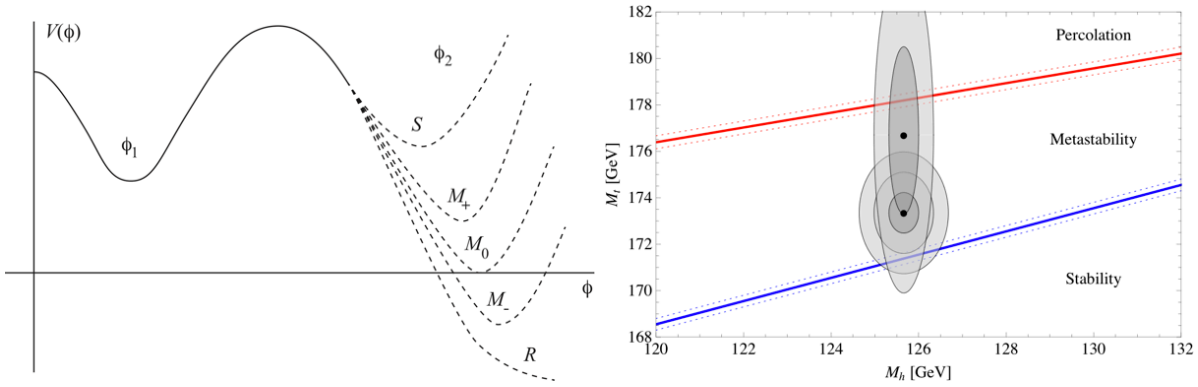


Figure 2.4: Left: a 2-dimensional projection of the Higgs potential with various scenarios for the high ϕ range. The current known Higgs vacuum position is ϕ_1 and the hypothetical second vacuum position is ϕ_2. If $V(|\phi_1|) > V(|\phi_2|)$, the Universe lives in a stable vacuum. If $V(|\phi_1|) < V(|\phi_2|)$ but $V(|\phi_2|) > 0$, the vacuum ϕ_1 is metastable, meaning that the Universe could tunnel into the false vacuum at $V(|\phi_2|)$. This scenario is most compatible with the data. Otherwise, the vacuum ϕ_1 is unstable. Right: stability regions of the electroweak vacuum as a function of the Higgs and top quark masses. The Higgs mass used in the plot is the average of the ATLAS [39] and CMS [40] measurements. The top quark mass used in the smaller contours comes is taken from the Tevatron measurements [63], and that in larger contours from the CMS top pole mass measurement [64]. Taken from [65].

that these questions need to be addressed by extensions of the SM, i.e. the so-called beyond SM (BSM) models. In the past decades, numerous BSM models have been constructed with various approaches to solving the open questions of the SM. The observation of a light Higgs boson, however, has disfavoured many of these models[9], and put very stringent constraints on the others.

The searches presented in this thesis concern hypothetical fourth generation t' and vector-like quarks (VLQs) decaying into a W-boson and a b-quark. These heavy quarks arise from the BSM models targeting the problems described in the previous section, which are (at the time of the search) allowed by the experimental constraints. The motivation and the phenomenology of these benchmark particles are presented below in this section.

Prior to the t' search presented in Chapter 8, the most stringent lower boundary on the mass of the t'-quark was set to $m_{t'} > 358$ GeV at 95% confidence level (C.L.) for the t' decay into W^+b, and $m_{t'} > 340$ GeV for the decay into W^+q, $(q = d, s, b)$. This result was provided by the CDF collaboration with 5.6 fb^{-1} of $p\bar{p}$ collision data [66]. Before the realisation of the search presented in Chapter 9, which started as another search for t'-quark before the discovery of the Higgs boson, and evolved to a VLQ search after the discovery, the most stringent limit was set to $m_{t'} > 557$ GeV by the CMS collaboration [67], while no searches for VLQs had been performed yet.

[9] "So far, though experiments at the LHC have discovered the Higgs boson, as yet they have found no direct hint of any new physics beyond the Standard Model such as supersymmetry or compositeness. The combination of these facts has caused a high mortality rate among theories, though not among theorists!" John Ellis [11]

2.3.1 Sequential Fourth Generation Quarks

2.3.1.1 Motivation

Although the SM with three generations of chiral fermions describes the Nature well, there is no particular reason why the number of generations would need to be exactly three. The number of quark generations has, however, an upper limit from theoretical considerations. In a SM with more than eight quark generations, the QCD would no longer exhibit the asymptotic freedom (Equation 2.83). The number of light neutrino generations (with a mass smaller than $M_Z/2$) was found to be three in the precision measurement of partial widths of the Z resonance at the LEP [68]. However, the observed neutrino mixing suggests that there might be further heavier neutrino generations. Prior to the discovery of the Higgs boson, a fourth generation of quarks containing a top-like t' and a bottom-like b' was still allowed according to the electroweak precision data and the measurements of the CKM matrix elements. A SM with four generations, referred to as the SM4 model, would mitigate some of the open questions of the SM (see [69] and references therein). The extended 4×4 CKM matrix would bring in additional source of CP violation and could shed some light on the baryon asymmetry problem. A chiral fourth generation of quarks could open up the possibility of introducing some further fermions, such as e.g. sterile neutrinos, which could be new dark matter candidates. As it is expected to couple dominantly to the third generation, it could have a contribution to the forward-backward asymmetry observed in $t\bar{t}$ production at Tevatron [70–72]. However, the fourth sequential generation of quarks is incompatible with the observed light Higgs boson, as a contribution of the fourth generation quark loops to the $gg \to H$ production mode would increase the Higgs boson production cross section in this channel by a factor of approximately 9. This contradicts the observed Higgs production cross section. Hence, the existence of the sequential fourth generation is now considered excluded [73].

2.3.1.2 Phenomenology

The global electroweak fit before the discovery of a Higgs boson required $m_{t'} - m_{b'} < m_W$ and $m_{b'} > m_t + m_W$, which combined with the preservation of the unitarity of the extended CKM matrix enforced the fourth generation quarks to predominantly mix with the third generation quarks. The predominant decay modes of the fourth generation quarks were thus expected to be $t' \to Wb$ and $b' \to Wt$. The fourth generation quarks can be produced either in pairs or in single production, as discussed below in Section 2.3.3.

2.3.2 Vector-Like Quarks

2.3.2.1 Motivation

While the observation of a light Higgs boson strengthens the evidence that the electroweak symmetry breaking is indeed governed by the BEH mechanism, it also increases the urge to resolve the weak points of the SM, especially the longstanding hierarchy problem. Various models with different approaches to this problem, such as little Higgs [74] or warped extra dimensions [75], all briefly summarised below, predict new non-chiral, i.e. vector-like quarks at the TeV scale. Apart from mitigating the hierarchy problem, VLQs play a role in the models

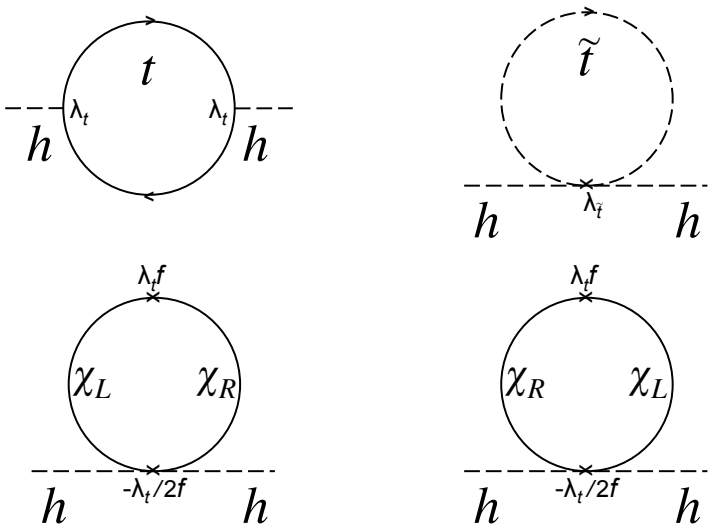


Figure 2.5: One loop corrections to the Higgs boson mass squared: top quark loop (top left), stop loop (top right), as well as two possible loops of the VLQ χ (bottom). The scalar fields (Higgs and stop) are drawn as dotted lines. Quarks are drawn as full lines. The left- and right-handed components of the VLQ are designated as $\psi_{L/R}$. The top (stop) Yukawa coupling to the Higgs boson is given by λ_t ($\lambda_{\tilde{t}}$). The coupling of the VLQ to the Higgs boson is determined as $-\lambda_t/2f$, with the top-Higgs Yukawa coupling λ_t and the compositeness scale f. Figures are adapted from [58]

targeting other open issues, such as the existence of dark matter [76], the observed cosmic baryon asymmetry [77], or the instability of the electroweak vacuum [78]. The mass range in which the VLQs are expected to arise was just about reached with the LHC data at $\sqrt{s} = 7$ and 8 TeV, and is expected to be extensively explored with the LHC data at $\sqrt{s} = 13$ and 14 TeV. This makes searches for VLQs particularly promising for revealing new physics knowledge in the near future.

Compositeness One of approaches to the hierarchy problem solution compatible with observations is based on the assumption that the Higgs boson is not a fundamental particle but rather a bound state of a new strong interaction. The so called *compositeness* models [74, 79] following this approach have a historical inspiration. Before it was established as a meson, i.e. a quark-antiquark bound state of the strong interaction, the neutral pion was considered an elementary particle, responsible for mediating the strong interaction. The low mass (~ 100 MeV) of the neutral pion could, however, not be explained without interpreting it as a meson. This in turn required new particles at the GeV scale, which were indeed found thereafter. Following this paradigm, numerous compositeness models have been developed, considering the Higgs boson as a state which appears elementary at the currently achievable energies, but whose substructure should be resolvable at some higher energy scale. In particular, in the compositeness models the Higgs boson represents a Nambu-Goldstone boson (NGB) of a new symmetry. In case of a spontaneous breaking of an exact symmetry, the NGB remain massless. If however the symmetry is not exact, or if it is broken both spontaneously and exactly, the NGBs obtain mass and become pseudo Nambu-Goldstone bosons (pNGB). This interpretation could render the Higgs boson light. The idea that the Higgs boson is a pNGB of some spontaneously broken

approximate global symmetry was first suggested in [80, 81] and later successfully constructed in [82]. This idea was then further developed by introducing collective symmetry breaking as in the Lillte Higgs (LH) models [56] or including extra dimensions [83–85].

Little Higgs Contrary to the initial idea of a new broken symmetry with compact nature, e.g. $SU(5)$, the LH models rely on the assumption that the new global symmetry is composed of at least two subgroups, e.g. $SU(3) \times SU(3)$, so that its breaking has a collective nature [58]. In other words, more than one coupling has to be turned on at the same time to enforce the symmetry breaking. This approach was motivated by the theoretical finding that if some global symmetries are simultaneously broken, there are no quadratic divergences in the mass squared of the mediator at the level of one loop corrections up to an energy scale of ~ 10 TeV. The loop divergences cancel between the same-spin particle loops, as shown in Figure 2.5, generating thereby little or no fine tuning. The spectrum of new particles varies among different LH models, but all of them predict the presence of at least one quark at the TeV scale with vector-like quantum numbers, along with additional scalars and gauge bosons.

Large Extradimensions The central concept of the extra dimensions (ED) models is that the SM fields and interactions are confined in the 4-dimensional space-time manifold, while gravity extends over additional dimensions. The weakness of the gravity is in such a case explained by the fact that the known Universe experiences only a projection into the 4-dimensional manifold. In the model proposed in [83], for $n \geqslant 2$ compact spatial dimensions large compared to the weak scale, in the sub-millimeter regime the gravitational force scales with the distance r as $1/r^{2+n}$, while at larger distances it has the Newtonian $1/r^2$ form. The ED extend over the reach of the energies present in the Universe and achievable with the experiments, which is why they are not observed. The ED models predict new particles, including quarks with vector-like couplings.

2.3.2.2 Phenomenology of Vector-Like Quarks

Vector-like quarks (VLQs) arising from various models are constructed as spin-1/2 fermions that transform as triplets under the $SU(3)_C$ gauge group, and whose left- and right-handed components carry the same colour and electroweak quantum numbers [86]. This implies that VLQs have no chirality, meaning that their left- and right-handed component transform the same way under the $SU(2)_L \times U(1)_Y$. Therefore, the mass term $m_f \bar{\psi}_f \psi_{\bar{f}}$ of a VLQ f is gauge invariant under this group, i.e. the mass m_f is generated without the Yukawa interaction with the Higgs boson. This is why the loop contributions to the $gg \to H$ process decouple faster than for chiral quarks. Thus, VLQs are less severely constrained by the current Higgs data, precision electroweak measurements, and direct searches than the fourth generation of chiral quarks. The left-right symmetry of VLQs allows for tree-level flavour changing neutral currents, which are their distinctive feature. The construction of the vector-like sector has been developed with various approaches. Some models introduce only the vector-like top partners [87], the others [88] develop the full VLQ sectors, while in some the vector-like families contain both quarks and leptons [89].

VLQs can mix with the SM quarks and thus contribute to their couplings to the Higgs and gauge bosons. The branching ratio of a decay of a VLQ with a mass of M_Q into a SM quark with a mass of m_q scales as $\sim m_q/M_Q$. Thus the vector-like quarks are expected to predominantly mix

with the third SM generation, while the other decays into the SM quarks are mass suppressed. The VLQ interactions involving the Higgs boson could have an impact on the Higgs branching fractions. Particularly, the apparent excess in the $H \to \gamma\gamma$ could have a VLQ contribution [42, 90].

The vector-like fields can have SM-like charges, like for $T_{2/3}$ and $B_{-1/3}$, or exotic charges, such as for $X_{5/3}$ and $Y_{-4/3}$. Given the symmetry structure of the SM, the VLQs can appear in only seven gauge-covariant multiplets of the $SU(3)_C \times SU(2)_L \times U(1)_Y$ symmetry group: singlets, doublets and triplets, all of them with both left- and right-handed component. The multiplets are summarised in Table 2.3 including the corresponding weak isospin and hypercharge values.

	Singlets		Doublets			Triplets	
Multiplet	T	B	$\begin{pmatrix} T \\ B \end{pmatrix}$	$\begin{pmatrix} X \\ T \end{pmatrix}$	$\begin{pmatrix} B \\ Y \end{pmatrix}$	$\begin{pmatrix} X \\ T \\ B \end{pmatrix}$	$\begin{pmatrix} T \\ B \\ Y \end{pmatrix}$
isospin	0	0	1/2	1/2	1/2	1	1
hypercharge	2/3	-1/3	1/6	7/6	-5/6	2/3	-1/3

Table 2.3: VLQ multiplets

The introduction of the new fields to the SM may lead to deviations of gauge couplings of the SM quarks. The couplings to the Z-boson are stringently constrained by the measurements of the relative partial particle width of the Z-boson into $b\bar{b}$ and $c\bar{c}$ pairs at LEP [4]. The constraints are the weakest for the top quark, which leaves a lot of room for accommodating a vector-like top partner. The impact of the VLQs can be expressed in terms of the Lagrangian terms of the gauge and the Higgs interaction of the SM and vector-like quarks, as well as the interaction between these two classes of quarks.

The Lagrangian terms of the gauge and Higgs Yukawa interaction of the third generation SM quarks with the couplings modified due to the presence of the positively and negatively charged vector-like quarks Q_t and Q_b respectively reads as:

$$
\begin{aligned}
\mathcal{L}_W &= -\frac{g}{\sqrt{2}}\bar{t}\gamma^\mu(V_{tb}^L P_L + V_{tb}^R P_R)bW_\mu^+ + h.c.\,, \\
\mathcal{L}_Z &= -\frac{g}{2c_W}\bar{t}\gamma^\mu(-X_{tt}^L P_L - X_{tt}^R P_R - 2Q_t s_W^2)tZ_\mu, \\
&\quad -\frac{g}{2c_W}\bar{b}\gamma^\mu(-X_{bb}^L P_L - X_{bb}^R P_R - 2Q_b s_W^2)bZ_\mu, \\
\mathcal{L}_H &= -\frac{gm_t}{2M_W}Y_{tt}\bar{t}tH - \frac{gm_b}{2M_W}Y_{bb}\bar{b}bH.
\end{aligned}
\tag{2.92}
$$

Here, the left- and right-handed projection operators $P_{L/R} = 1/2(1 \pm \gamma^5)$ are used. The factors c_W and s_W stand for the sine and cosine of the Weinberg angle. The left- and right handed charged current couplings are given by $V_{q_1 q_2}^{L/R}$ and, the neutral current couplings by $X_{qq}^{L/R}$, and the Yukawa couplings to the Higgs boson by Y_{qq}. These couplings vary between the singlet, doublet and triplet states and are summarised in the Appendix of [86]. The same applies to the couplings in the following two equation blocks. Apart from distorting the neutral currents and interactions with the Higgs boson, the presence of the VLQs introduces the right-handed

contribution to the charged currents not present in the SM. This contribution, however, has to be small to be consistent with data.

The Lagrangian terms for the gauge and Yukawa interactions of the vector-like quarks are:

$$\begin{aligned}
\mathcal{L}_W &= -\frac{g}{\sqrt{2}}\bar{Q}\gamma^\mu(V^L_{QQ'}P_L + V^R_{QQ'}P_R)Q'W^+_\mu + h.c.\,, \\
\mathcal{L}_Z &= -\frac{g}{2c_W}\bar{Q}\gamma^\mu(\pm X^L_{QQ}P_L \pm X^R_{QQ}P_R - 2Q_Q s^2_W)QZ_\mu, \\
\mathcal{L}_H &= -\frac{gm_Q}{2M_W}Y_{QQ}\bar{Q}QH,
\end{aligned} \tag{2.93}$$

with $Q, Q' = X, T, B, Y$. The $+$ $(-)$ sign in the $\mathcal{L}_Z$ term corresponds to X and T (B and Y). This set of equations describes the charged and neutral currents among the VLQs. Their realisation is, however, kinematically constrained by the mass splitting between the VLQs. In most of the current phenomenological models these decays are not yet considered.

The Lagrangian terms describing the interaction of the SM quarks with the vector-like quarks are:

$$\begin{aligned}
\mathcal{L}_W &= -\frac{g}{\sqrt{2}}\bar{Q}\gamma^\mu(V^L_{Qq}P_L + V^R_{Qq}P_R)qW^+_\mu + h.c. \\
&\quad -\frac{g}{\sqrt{2}}\bar{q}\gamma^\mu(V^L_{qQ}P_L + V^R_{Qq'}P_R)qW^+_\mu + h.c.\,, \\
\mathcal{L}_Z &= -\frac{g}{2c_W}\bar{q}\gamma^\mu(\pm X^L_{qQ}P_L \pm X^R_{qQ}P_R - 2Q_Q s^2_W)QZ_\mu + h.c.\,, \\
\mathcal{L}_H &= -\frac{gm_Q}{2M_W}\bar{q}(Y^L_{qQ}P_L + Y^R_{qQ}P_R)QH + h.c..
\end{aligned} \tag{2.94}$$

These interactions determine the decay modes of the VLQs into the third generation SM quarks. The most striking consequence of these interaction is the presence of the flavour changing neutral currents described by the $\mathcal{L}_Z$ term in the previous set of equations. Since the VLQ couplings differ between the singlet, doublet and triplet states, their decay modes differ correspondingly.

The singlet states can exhibit the following decays:

$$\begin{aligned}
T &\rightarrow W^+b/Zt/Ht, \\
B &\rightarrow W^-t/Zb/Hb, \\
X &\rightarrow W^+t, \\
Y &\rightarrow W^-b.
\end{aligned} \tag{2.95}$$

The branching ratios of these decay modes are not exactly known in any existing model (although they can be calculated under certain assumptions). However, the couplings are correlated in a way that the existence of the charged current interaction automatically implies the neutral and scalar interactions of comparable strengths.

In the $(T\,B)$ doublet states, the experimental constraints imply that the T- and B-quarks are almost degenerate in mass and that their couplings of to the b-quark are suppressed. In the $(X\,T)$ and $(B\,Y)$ doublets, two quarks are also expected to have very similar masses. In

addition, they do not exhibit charged current interactions at the leading order. This results in the decay modes:

$$\begin{aligned} \begin{pmatrix} T \\ B \end{pmatrix} &: \quad \begin{aligned} T &\to Zt/Ht \\ B &\to W^-t \end{aligned} \\ \begin{pmatrix} T \\ X \end{pmatrix} &: \quad \begin{aligned} T &\to Zt/Ht \\ X &\to W^+t \end{aligned} \\ \begin{pmatrix} B \\ Y \end{pmatrix} &: \quad \begin{aligned} B &\to Zb/Hb \\ Y &\to W^-b \end{aligned} \end{aligned} \tag{2.96}$$

It is worth noticing that in the case of a simple search without knowing the charge of the b-quark of the event, the experimental signature of the Y-quark in both the singlet and doublet state would be indistinguishable from that of the fourth generation t'-quark. The T-quark would also have the same signature assuming the $T \to W^+b$ as the only decay. This assumption can be used for searches based on simplified models, although it is not a realistic scenario in most of the existing models.

Vector-like quarks contribute to the Higgs production through gluon-gluon fusion (GGF) and to the Higgs decay into a pair of photons through loop diagrams. At the lowest order, the Higgs production cross section through GGF scales as [91]:

$$\sigma(gg \to H) \sim \left(\sum_q Y_{qq} F_{1/2}(\tau_q) \right)^2 , \tag{2.97}$$

with q running over the b, t and included vector-like quarks, Y_{qq} being the Yukawa couplings of the quark q to the Higgs boson, $\tau_q = M_H^2/4m_q^2$, and $F_{1/2}(\tau_q)$ defined in [91]. For $m_q \gg M_H$, the function $F_{1/2}(\tau_q)$ converges to 4/3 while for very small m_q it approaches zero. Taking into account the constraints on the mixing between the SM quarks and VLQs, the Higgs production cross section through GGF in a model which includes a VLQ extension of the SM deviates from that with no extensions by only a few percent [92]. In other words, the VLQs decouple from the GGF amplitude. Thus, contrary to the hypothetical sequential fourth generation quarks, the existence of VLQs is consistence with current Higgs data, while it accommodates the cancellation of quadratic divergences in m_H^2 development arising from the top quark loops.

2.3.3 Heavy Quark Production

In hadron collisions, vector-like quarks are produced either in pairs through QCD or as single quarks in association with bosons through the weak interaction. The production cross sections depend on the quark mass and the centre-of-mass energy $\sqrt{s}$, while the single production also depends on the corresponding couplings. In the mass range explored in this thesis, i.e. in the sub-TeV range, the pair production gains higher cross sections (Figure 2.6 left) and is thus the only process considered.

At tree level, i.e. not considering loop corrections, $t'\bar{t}'$- and $T\bar{T}$-pairs are produced through gluon-gluon fusion $gg \to t'\bar{t}'/T\bar{T}$, and through quark-antiquark annihilation $q\bar{q} \to t'\bar{t}'/T\bar{T}$ (Figure 2.6 right).

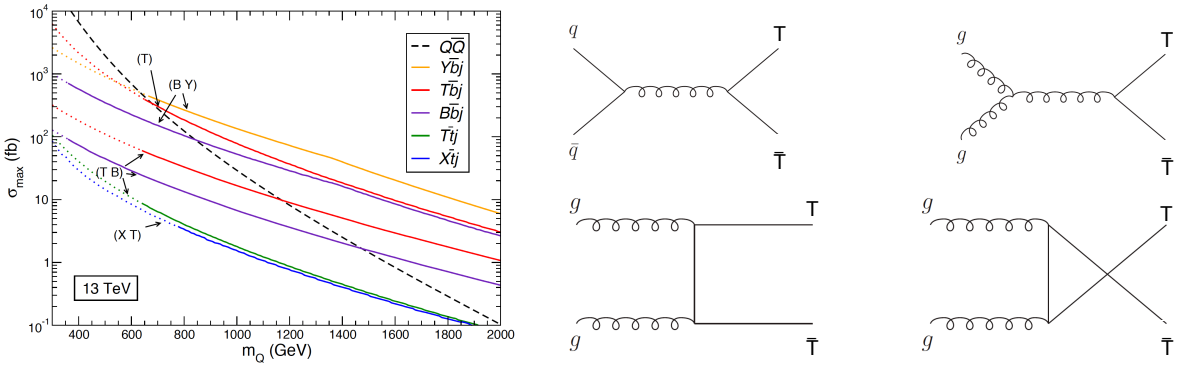


Figure 2.6: Left: heavy quark production cross sections through pair and single production for selected quark multiplets, at the LHC with 13 TeV. The dotted part of the lines indicate the range of masses already excluded by direct searches. The plot is taken from [86]. Right: leading order heavy quark pair-production modes ($q\bar{q}$-initiated production in the top left and three possibilities for the gg-initiated production in the remaining diagrams). The diagrams are taken from [93]

Chapter 3

Experimental Setup

Particle colliders have played a key role in particle physics research over the past decades. Tremendous developments of the collider accelerator and particle detector technologies have allowed a detailed examination of the Standard Model of Particle Physics, including discoveries of the W- and Z-bosons at European Organisation for Nuclear Research (CERN) with the Super-Proton Synchrotron (SPS) in 1983, the top-quark at Fermilab with the Tevatron in 1993 and finally the Higgs boson at CERN with the Large Hadron Collider (LHC) in 2012.

Particle colliders are mainly built with one of the two technologies: linear accelerators and circular synchrotron colliders. Linear particle accelerators have a rather simple architecture with a drawback that for a strong acceleration very long machines are needed. Circular colliders are in this sense advantageous as the particles can accumulate energy over numerous circulations along the same path. A drawback of circular architecture is the energy loss that charged particles experience in the form of synchrotron radiation ΔE_{syn} when changing the direction of flight. This energy loss is inversely proportional to the radius of the accelerator ring and to the fourth power of the mass of the particle used in the beam. For this reason synchrotron colliders with large circumference using heavy particle beams are favourable.

The two most important features of a particle accelerator are the provided $\sqrt{s}$ of the colliding particles and the instantaneous luminosity $\mathcal{L}$. These two quantities determine which scatter processes are possible and how many such interactions are expected in a given data taking time. The luminosity $\mathcal{L}$ achieved with two colliding proton beams is defined as the number of proton-proton collisions per unit area per unit of time. For two colliding beams with N_b particles per bunch, n_b bunches in each beam at a revolution frequency f_{rev}, relativistic factor γ_{r} and a beta function β^*, luminosity is computed as [94]:

$$\mathcal{L} = \frac{N_b^2 n_b f_{\text{rev}} \gamma_{\text{r}}}{4\pi\varepsilon_n \beta^*} F, \tag{3.1}$$

with the normalised transverse beam emittance ε_n and the geometrical reduction factor F due to the crossing angle of the beams at the interaction point (IP). The protons are assumed to be distributed according to two-dimensional Gaussian distribution in the transverse plane within bunches. The number of events of the process $X \to Y$ with a cross section $\sigma_{X\to Y}(\sqrt{s})$

(dependent on $\sqrt{s}$) occurring in the time interval Δt is given by:

$$N_{X\to Y}(\sqrt{s}, \mathcal{L}, \Delta t) = \int_{\Delta t} \sigma_{X\to Y}(\sqrt{s})\mathcal{L}dt = \sigma_{X\to Y}(\sqrt{s})\mathcal{L}_{int}, \tag{3.2}$$

with the integrated luminosity $\mathcal{L}_{int}$. The instantaneous luminosity $\mathcal{L}$ is time dependent, even for constant operational parameters, since it depends on the beam intensity, which decreases over time.

The experimental results presented in this thesis have been performed with proton-proton (pp) collision data at $\sqrt{s} = 7$ TeV. The collisions were supplied by CERN's main particle accelerator, the Large Hadron Collider (LHC), described in Section 3.1. The collisions were recorded by the largest LHC particle detector, ATLAS, described in Section 3.2.

3.1 Large Hadron Collider

The LHC [95–97] is a 2-ring circular synchrotron hadron collider with a circumference of about 27 km situated in the underground area near Geneva, Switzerland. It is installed in the tunnel of its dismounted predecessor, the Large Electron-Positron Collider (LEP), buried $45 - 170$ m below the surface. The LHC is currently the world's largest and the most powerful particle collider, primarily dedicated to the searches for the Higgs boson and new physics beyond the SM. It is designed to operate at $\sqrt{s} = 14$ TeV for pp-collisions and $\sqrt{s} = 2.76$ TeV per nucleon (575 TeV per ion) for lead ions collisions. The collisions provided by the LHC are recorded by seven particle detectors: ALICE [98], ATLAS [99], CMS [100], LHCb [101], LHCf [102], MoEDAL [103], and TOTEM [104]. ATLAS and CMS have a general purpose design, i.e. are aimed to a large variety of physics processes, while the other experiments are specialised for various physics fields.

The LHC operation involves several other machines from CERN's accelerator complex (Figure 3.1). Preparation of the LHC beams starts by ionising hydrogen atoms to produce protons. These protons are then organised into bunches, focused and accelerated to 750 keV in the Radio Frequency Quadrupole (RFQ). Such proton beams are then accelerated in the linear accelerator Linac2 to 50 MeV and passed to the Proton Synchrotron Booster (PSB), composed of four superimposed synchrotron rings, to get further energy increase up to 1.4 GeV. After the PSB, the beam is transferred to the Proton Synchrotron (PS), where the energy of the protons is raised to 25 GeV and the bunches grouped into so-called bunch trains. At the nominal running conditions each bunch contains about $1.15{\cdot}10^{11}$ protons and the bunches are separated by bunch spacing of 25 ns. These design conditions have, however, not yet been achieved (Section 3.1.1). The last pre-acceleration takes place in the Super Proton Synchrotron (SPS), where proton energy is increased to 450 GeV. From there, proton bunches are injected into the LHC in two opposite directions forming two beams which travel in separate vacuum beam pipes. The beam pipes are embedded in the superconducting dipole magnets cooled to 1.9 K by the cryogenic system using liquid helium. There are 1232 dipole magnets in total along the LHC ring of 26.7 km in circumference. The dipole magnets maintain the beams on a stable orbit with their dipole magnetic field with a designed value of 8.33 T, while additional 392 superconducting quadrupole magnets focus the beams in the transverse plane. The maximal number of bunches in the LHC is foreseen to be 2808.

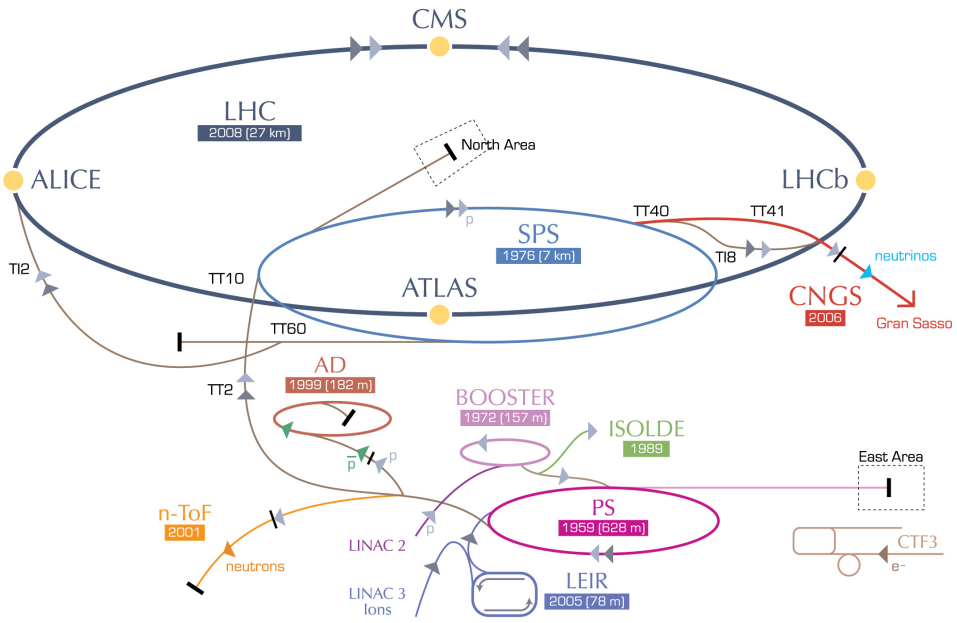


Figure 3.1: CERN's accelerator complex containing the LHC and its pre-accelerators Linac2, PSB, PS, and SPS.

Once the planned energy and appropriate focusing of the proton beams is reached, the beams are declared stable. At that point the experiments start taking data. During a run, beam intensity decreases gradually due to a proton loss through collisions with other protons or beam gas particles. When the intensity of the beams has dropped below a critical value, the beams are redirected outside of the accelerator into absorber material. The usual duration of data taking for one fill is several hours.

3.1.1 LHC Run I

After a decade of construction and commissioning, the LHC finally started operating in November 2009. With short breaks, the LHC continued producing pp- and P_bP_b-collisions until the end of 2012, when operation was terminated for the first long shut down (LS1), in which certain components need to be replaced in order to prepare for running at nominal conditions. Over this data taking period, referred to as the LHC Run I, the LHC beam conditions were varying. Starting from first collisions at 900 GeV, the LHC collision energy was gradually raised to reach 7 TeV in March 2010. This collision energy was maintained during the whole 2011, and was then raised to 8 TeV in 2012.

The luminosity was also increased over the run period, as shown in Figure 3.2. This was achieved by increasing the beam intensity, i.e. the number of protons per bunch $1.2 - 1.6 \cdot 10^{11}$ while increasing the number of bunches per beam from 348 to 1380, and thus reducing the bunch spacing from 150 ns to 50 ns.

With increasing instantaneous luminosity, the probability of multiple pp collisions per bunch crossing, referred to as in-time pile-up, also increases. In addition to the in-time pile-up, the out-of-time pile-up can also occur, in which a collision from previous bunch crossing is falsely assigned to the ongoing bunch crossing due to the relatively long integration time of the front-

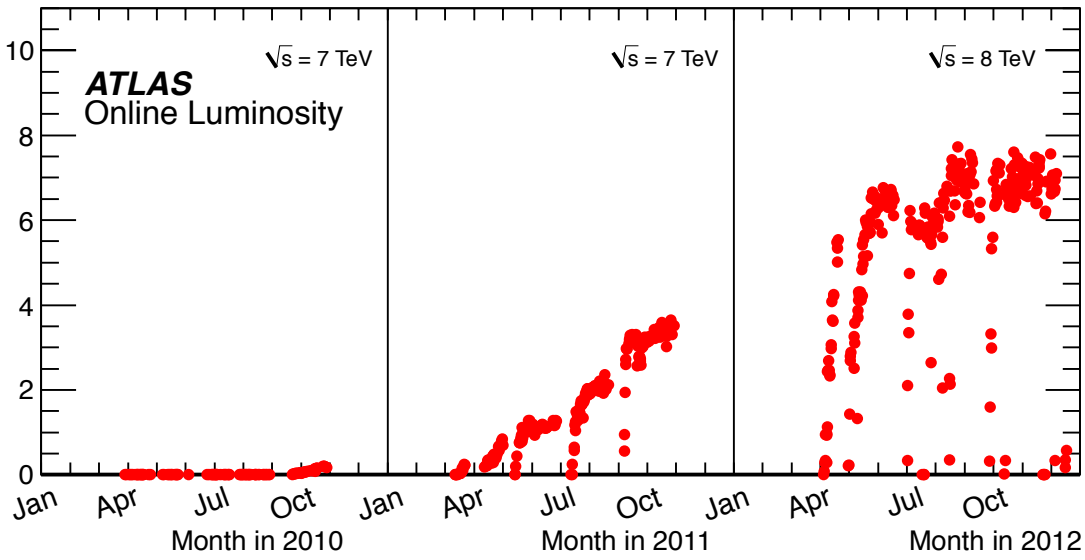


Figure 3.2: Evolution of the instantaneous luminosity in the LHC Run I. Taken from [105].

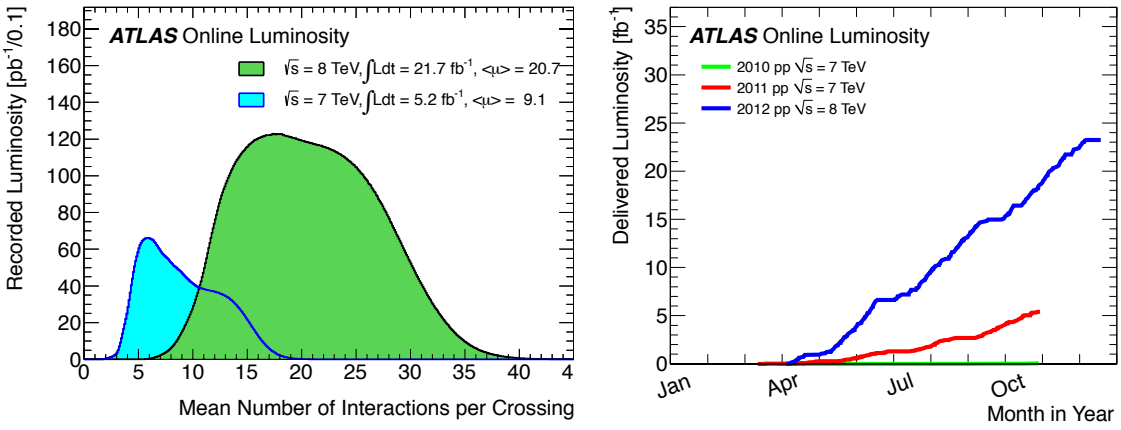


Figure 3.3: Left: Recorded luminosity as a function of the mean number of interactions per bunch crossing for the 2011 (2012) data taking campaign shown in cyan (green). Right: cumulation of the total delivered luminosity of *pp* collisions over the year shown for the 2010 (green line), 2011 (red line) and 2012 (blue line) data taking campaigns. Taken from [105].

end electronics. This kind of pile-up gains importance with short bunch spacing. The effect of the pile-up in a given run is expressed in terms of the average number of *pp* collisions per bunch crossing in that run, $< \mu >$. The average $< \mu >$ increased from 9.1 in 2011 to 20.7 in 2012 (Figure 3.3 left). The Run I data taking period resulted in the total integrated luminosity of 45.0 (48.1) pb^{-1} recorded by ATLAS (delivered by the LHC) in 2010, 5.08 (5.46) fb^{-1} in 2011, and 21.3 (22.8) fb^{-1} in 2012 (Figure 3.3 right).

3.2 ATLAS Detector

The design of the ATLAS (A Toroidal LHC ApparatuS) experiment [106] is targeted to reconstruct all elements of the final states of events involving the Higgs boson and the particles expected from physics beyond the SM. It is optimised to perform precise particle identification in almost the full space angle while sustaining a challenging radiation environment at the nominal instantaneous luminosity of 10^{34} cm^{-2}s^{-1}. The sensitive elements as well as the readout electronics were built from radiation-hard materials with fast response. The trigger system was designed to efficiently recognise the events with high transverse momentum (p_{T}) objects interesting for physics analysis, which are then recorded by the fast read-out system.

The ATLAS detector is situated in the Point 1 (P1) underground experimental hall on the LHC tunnel. The whole apparatus measures 44 m length and 25 m in diameter. It has a radially symmetric construction consisting of a set of cylindrical sub-detectors concentrically arranged around the beam line with the nominal interaction point as the origin. A schematic overview of the ATLAS detector is shown in Figure 3.4. Subsystems are arranged in a way that maximises the efficiency of particle identification as well as the position, momentum and/or energy measurements of the reconstructed objects. Figure 3.5 shows the response of individual detector layers to different particles. The particles relevant to the searches presented in this thesis are electrons, muons, hadrons and neutrinos. According to Figure 3.5, for a successful reconstruction of events containing these particles, all ATLAS subsystems are relevant.

The innermost part is the fine granularity Inner Detector (ID) immersed in a 2 Tesla solenoid magnetic field, dedicated to the reconstruction of charged particle tracks. It is explained in Section 3.2.2. The ID is surrounded by the Electromagnetic Calorimeter (ECAL), described in Section 3.2.3, dedicated to the energy measurement and identification of electrons and photons, and the succeeding Hadronic Calorimeter (HCAL) described in Section 3.2.3.2, dedicated to an accurate measurement of the energy of hadrons. Combined measurements from ECAL

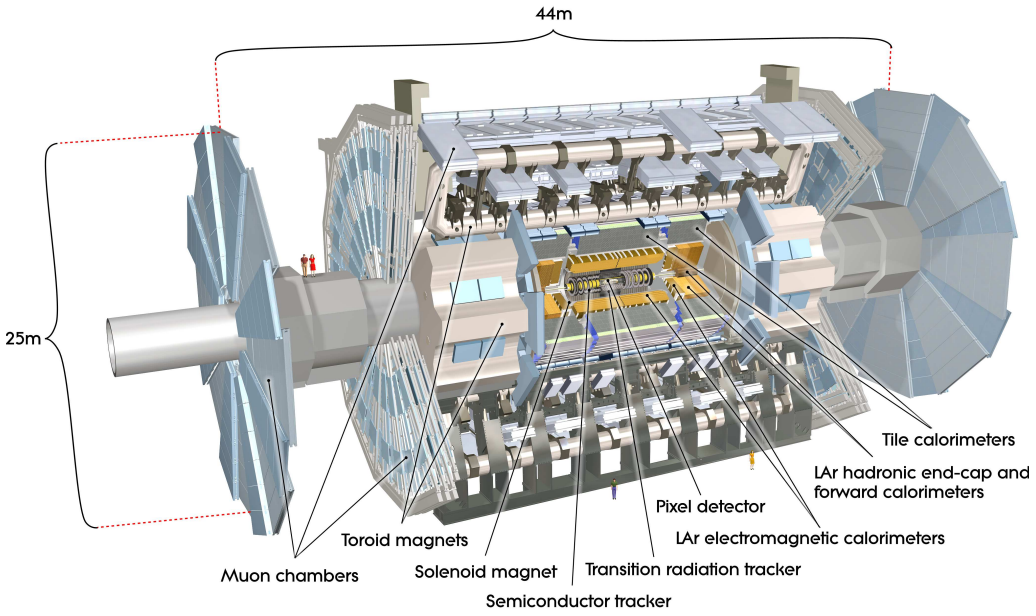


Figure 3.4: A schematic overview of the ATLAS detector showing all subsystems. Taken from [99].

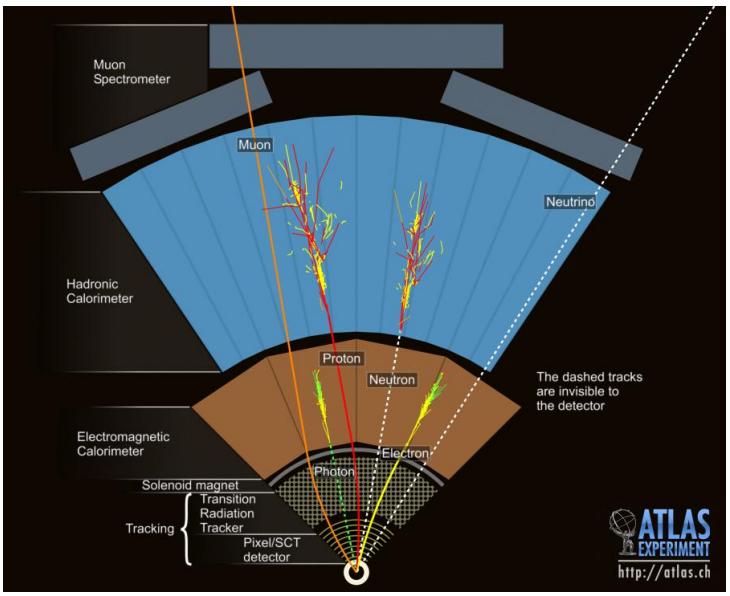


Figure 3.5: A slice of a transverse section of ATLAS showing the detector response to the various particles. Photons: energy deposit in the ECAL, no traces in the ID tracker. Electrons: energy deposit in the ECAL,a negatively bended trace in the ID. Neutrons: energy deposit in the HCAL, no trace in the ID and ECAL. Protons: energy deposit in the HCAL, small portion energy in the ECAL, a positively bended trace in the ID. Muons: negatively bended traces in the ID and the MS, almost no energy deposit. Neutrinos: no signal. Taken from [99].

and HCAL allows for an estimation of the missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) originating from neutrinos (see Section 5.6 for an explanation). The outermost ATLAS sub-detector is the Muon Spectrometer (MS), presented in Section 3.2.4, embedded in the toroidal magnetic field, designed to detect muon tracks from which muon momentum is determined. The solenoid and toroid magnets are explained in Section 3.2.1.

The ATLAS coordinates are expressed in terms of a right handed coordinate system with the origin at the nominal IP, the x-axis laying in the horizontal plane pointing towards the centre of the LHC, the y-axis pointing upwards and the z-axis laying along the beam line. Another coordinate system used in ATLAS is the cylindrical system around the z-axis with the polar angle θ measured from the positive z-axis and the azimuthal angle ϕ measured in the $x-y$-plane from the positive x-axis.

A particularly useful coordinate in the collider physics is the rapidity y, defined as:

$$y = \frac{1}{2} \ln \left(\frac{E + p_{\mathrm{z}}}{E - p_{\mathrm{z}}} \right), \tag{3.3}$$

with the particle energy E and the z-component of its momentum. The advantage of this variable is that the rapidity difference of two particles is a boost invariant quantity. However, this variable requires a simultaneous measurement of the energy and the momentum of the particle, which is particularly difficult for the particles with a high p_z. A more convenient variable which depends only on the magnitude of the particle momentum $|\mathbf{p}|$ and its longitudinal component p_{z} is the pseudo-rapidity η defined as:

$$\eta = \frac{1}{2} \ln \left(\frac{|\mathbf{p}| + p_{\mathrm{z}}}{|\mathbf{p}| - p_{\mathrm{z}}} \right). \tag{3.4}$$

In the limit where the particle speed is very close to the speed of light in the vacuum, or if the particle mass is nearly zero, this quantity is an approximation of the rapidity y. It is related to the polar angle θ as $\eta = -\ln(\tan(\theta/2))$.

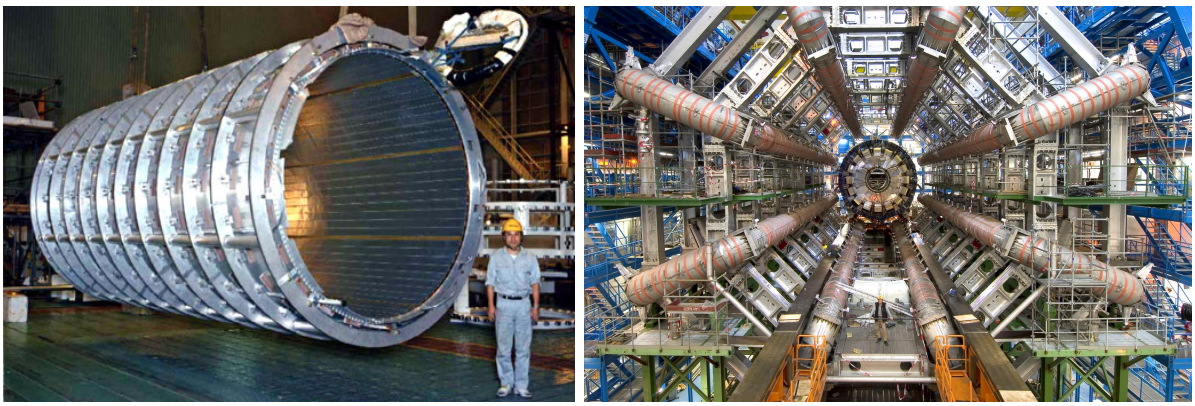


Figure 3.6: The ATLAS solenoid (left) and the ATLAS barrel toroid as installed the cavern [99]

3.2.1 Magnetic System

The ATLAS detector has two tracking detectors, the ID and the MS, with a calorimeter system in between. Each tracking system is immersed in a magnetic field, to provide a momentum measurement.

The ID is situated inside a superconducting solenoid, aligned along the beam axis, providing an axial magnetic field of 2 T at the IP, which drops steeply from ~ 1.8 T at $|z| = 1.7$ m to ~ 0.9 T at the end of the ID cavity. The field is produced by an operational current of 7.7 kA and stores an energy of 40 MJ. The solenoid (Figure 3.6 left) is made of a single-layer high-strength Al-stabilised NbTi conductor, designed to achieve a high field while minimising the radiative thickness upstream of the calorimeter. It is contained inside a 12 mm thick Al alloy support cylinder. The total thickness of the solenoid and the container is equivalent to ≈ 0.66 radiation lengths. The inner diameter of the solenoid is 2.46 m, and its axial length is 5.8 m. The HCAL steel and girder structure serve as the return for the magnetic flux. It takes 30 minutes to charge or discharge the solenoid.

The MS is embedded within a system composed of a barrel and two endcap toroids. The barrel toroid consists of eight coils placed in individual vacuum vessels. The system of coils is supported by inner and outer rings of struts (eight of each). The barrel toroid (Figure 3.6 right) is 25.3 m long and has the inner and outer diameters of 9.4 m and 20.1 m respectively. The endcap toroid magnets are each composed of 8 flat, squared coil units, glued together to form a rigid structure capable of withstanding the Lorentz force. Each endcap toroid has an inner and outer diameter of 1.65 m and 10.7 m respectively and a length of 5 m. The endcap toroid coils are rotated by 22.5° with respect to the barrel toroid coils to provide an optimal bending power at the overlapping region between the two coil systems. In the range $|\eta| < 1.4$, magnetic field is provided by the barrel magnet, providing an R and ϕ dependent field varying between 0.15 T and 2.5 T, and a bending power of 1.5 Tm to 5.5 Tm. The range $1.6 < |\eta| < 2.7$ is covered by the endcap magnets producing a field of 0.2 T to 3.5 T, and a bending power of 1 Tm to 7.5 Tm. In the region $1.4 < |\eta| < 1.6$ the barrel and the endcap magnetic fields merge and the bending power is reduced compared to the other two regions. Together with the supporting structure and the vacuum vessels, the barrel toroid weighs 1318 tons, while two endcaps each weigh 400 tons.

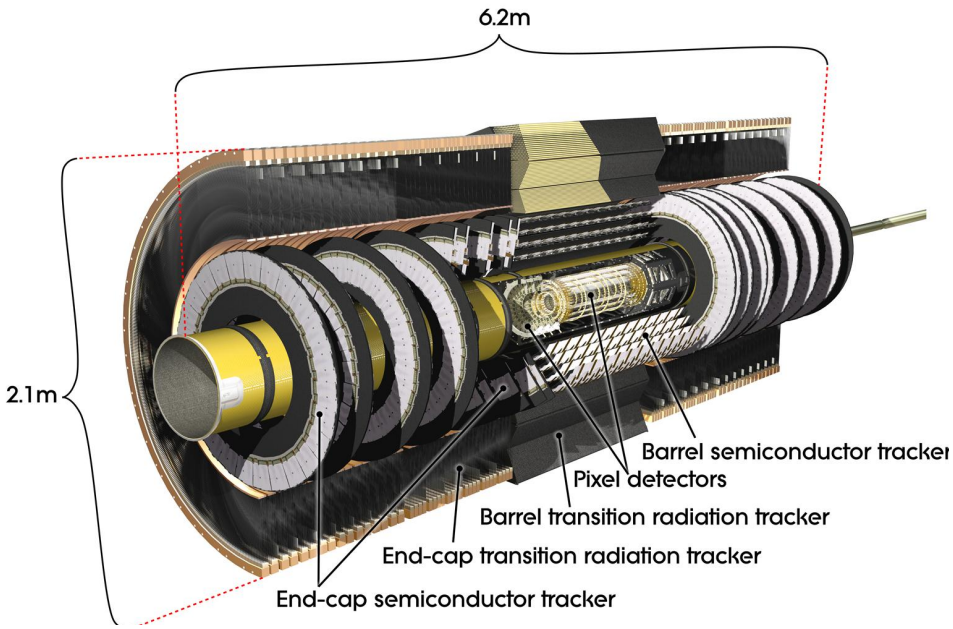


Figure 3.7: Schematic view of the ATLAS ID [99].

3.2.2 Inner Detector

The Inner Detector (ID) is the innermost subsystem of the ATLAS detector. It is specialised for: efficient tracking, which is necessary for the momentum measurements of high p_{T} leptons; vertexing, which is crucial for the reconstruction of the IP, as well as the decay vertices of produced particles; the identification of photons, electrons, τ-leptons and heavy-flavour hadrons. Charged particles traversing the ID material leave their traces in a form of electron-hole pairs in the silicon or ionised gas. Being immersed in the solenoid field, these traces bend, allowing for the momentum determination from the measured radius of the trace curvature. At low luminosity, the ID is capable of carrying out the full event reconstruction.

The ID has a cylindrical shape with a length of 7 m, spreading between the two endcap calorimeters, with an inner radius of 4.55 cm, corresponding to the innermost layer of the Pixel detector, and an outer radius of 110.6 cm, corresponding to the outermost ring of the TRT endcap (Figure 3.7). The ID consists of three independent and complementary sub-detectors: the Pixel detector, the Semi-conductor Tracker (SCT) and the Transition Radiation Tracker (TRT) described below. The ATLAS ID is designed to perform transverse momentum measurement with a resolution of $\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}} = 0.05\%\ p_{\mathrm{T}} \oplus 1\%$. An $y-z$-plane quarter-section of the ID showing the spacial coverage of all ID subsystems is shown in Figure 3.8.

3.2.2.1 Pixel Detector

The Pixel detector is the innermost ATLAS sub-detector. The purpose of the Pixel detector is to accurately resolve the dense charged track occupation in the proximity to the IP. The signals in the Pixel detector and especially those in its innermost layer, called b-layer, play a key role in vertex reconstruction and in the identification of the jets originating from b-quarks (b-tagging). The Pixel detector has an active radius in the range $45.5 < r < 242$ mm, and a length of

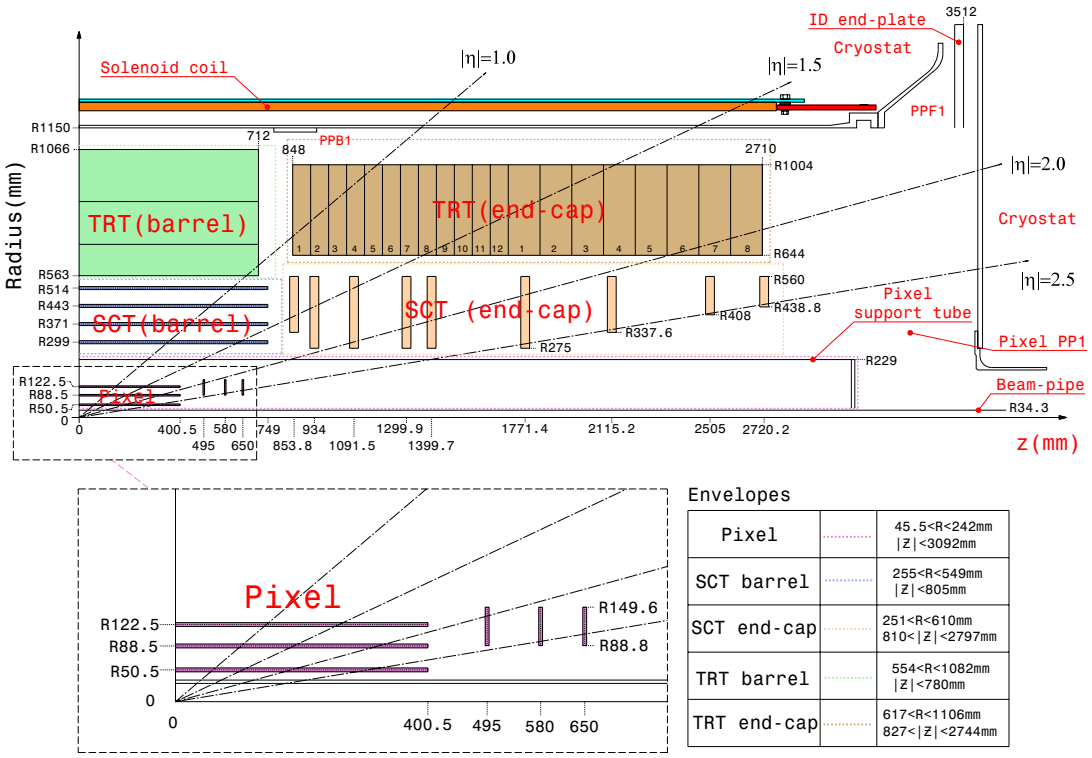


Figure 3.8: Schematic view of the Inner Detector section in the $y-z$-plane [99].

$\sim 1.3m$ corresponding to the pseudo-rapidity coverage of $|\eta| < 2.5$. It is composed of three cylindrical barrel layers aligned with the beam axis, and six endcap discs (three in each endcap) perpendicular to it. It consists of 1744 modules of a surface of 63.4×24.4 mm^2, each carrying 47232 pixels read out by 46080 electronic channels, which makes a total of 80.4 M read-out channels. Individual pixel sensors have a minimum surface of 50×400 μm and are 180 μm thick. The spacial resolution of the pixel detector is therefore 10 μm in $R-\phi$ and 115 μm in z (R) direction for the barrel (endcaps). The Pixel detector is exposed to an enormous radiation. The b-layer withstands an ionisation dose of 160 kGy/y and an 1 MeV neutron equivalent fluence (F_{neq}) of $\sim 8 \times 10^{13}$ cm^{-2}. This causes temperature dependent annealing and increase in the sensor leakage current. To minimise these two problems, the detector operates in the temperature range from -5°C to -10°C.

3.2.2.2 Semi-conductor Tracker

The Semi-conductor Tracker (SCT) provides further position measurements along the charged particles traces (in addition to those provided by the Pixel detector), which are crucial for track reconstruction. The SCT is composed of four cylindrical barrel layers with the beam line as a common axis, and eighteen endcap discs (nine per endcap) perpendicular to it, composed of 4088 modules with a total surface of 63 m^2 and providing a total of ~ 6.3 M readout channels. An SCT module consists of two pairs of silicon single-sided micro-strip sensors. The pairs of sensors are glued back-to-back with a separation of ~ 300 μm and a stereo alignment angle of 40 mrad between the two planes. Two pairs are bonded along one edge to form a plane sensor with an overall strip length of 126 mm. An SCT module carries 1536 readout channels (768 on

each side). The SCT strips are 285 ± 15 μm thick and 126 mm long. The barrel strip sensors are rectangular with a pitch of 80 μm, whereas those for the endcaps are trapezoidal with a mean pitch of $\sim$ 80 μm and a constant azimuth. The energy deposited by charge particles is collected in the strips. Two such energy deposits (so-called hits) from both sides of a module are combined into a geometrical space-point. Such space-points are used in for track reconstruction. The SCT provides up to four measured space-points along the traversal of one charged particle. The intrinsic accuracy of the SCT measurements is 17 μm in the $R-\phi$ plane, and 580 μm in z (R) direction for barrel (endcaps). The SCT layers and discs are exposed to an ionisation dose of 3-7.6 kGy/y and radiation fluence (F_{neq}) of $9-16\times10^{12}$ cm^{-2} and thus operate at a temperature of -5°C.

The signals from the SCT are read out by application-specific integrated circuits (ASICs) known as ABCD3TA. Each of these chips reads out 128 channels, i.e. twelve chips read out an SCT module. The SCT data-taking is governed by the SCT data acquisition (DAQ) system, which is responsible for configuring the front-end ASICs, communicating first-level trigger information, and transfer data from the front-end chips to the ATLAS high-level trigger system. The information communicated by the SCT DAQ has a byte-stream format. The errors occurring in the byte-stream readout can result in the absence of information vital for data quality estimation. The impact of the byte-stream error on the quality of data collected in 2010 was studied by the author of this thesis. The study is presented in Appendix A.

3.2.2.3 Transition Radiation Tracker

The Transition Radiation Tracker (TRT) is the outermost component of the ID. Unlike to the other two silicon based trackers, the TRT is composed of a system of drift chambers. The basic cell of the TRT is a coaxial capacitor in a form of a straw tube (cathode) with a diameter of 4 mm, with a gold plated wire in the middle (anode), interleaved with foils with different refractive indices (transition radiation material). The tubes are filled with a gas mixture composed of Xe (70%), CO_2 (27%) and O_2 (3%) with $5-10$ mbar over-pressure. In the barrel, 144 cm long tubes are arranged in 73 layers parallel to the beam axis between the radii or 56.3 and 106.6 cm, while in the endcaps the 37 cm long straws are organised in 160 planes, covering the radial space between the radii of 64.4 cm and $\sim$ 1 m, and extending up to 2.7 m away from the IP in the z direction, which corresponds to a pseudo-rapidity coverage of $|\eta| < 2$. When traversing a TRT cell, a charged particle ionises the gas in the tube, which inside the electric field triggers the development of an ionisation cascade. The free electrons from the cascade are collected on the anode forming a signal proportional to the energy of the particle that triggered the ionisation. The position of the incoming particle inside the tube can be determined from the drift time. Under normal operating conditions, the electron collection normally takes up to 48 ns with a drift-time accuracy of 130 μm. In addition, when traversing the transition radiation foils, a charged particle induces the production of transition radiation photons, which also trigger an ionisation cascade with an energy proportional to the energy-to-mass ratio (E/m) of the incoming particle. Measurement of the transition radiation energy combined with the primary cascade measurements provides an efficient discrimination between electrons and hadrons (mostly pions). By adjusting the front-end electronics threshold the transition radiation can efficiently be separated from the minimal ionising signal. Contrary to the silicon trackers, the TRT operates at the room temperature maintained by the heater pads placed on the thermal enclosures surrounding the SCT detector.

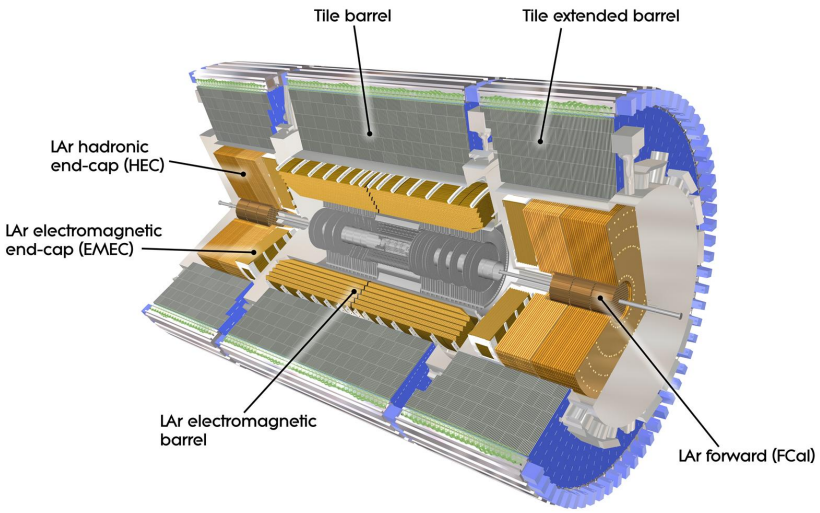


Figure 3.9: Schematic view of the ATLAS Calorimeter [99].

3.2.3 Calorimeter System

The energy which a particle emerges from the hard process is fully or partially deposited in the detector material through which it passes. The energy deposition mechanism depends on the initial energy and the mass of the particle, as well as the interactions it exhibits. Electrons, photons and to some extent muons undergo electromagnetic interaction with the electrons in the atoms of the material. For hadrons, the dominant interaction with material is the strong interaction with the nuclei of the atoms.

A high energy electron entering material induces an electromagnetic shower resulting from a succession of the bremsstrahlung and the electron-positron pair creation processes. The shower development continues as long as these processes are energetically allowed, after which the particle energy is deposited in the form of the measurable ionisation charge. Energy measurement based on this principle is performed in the electromagnetic calorimeter (EMCal). The spatial development of the shower is characterised by the so called radiation length X_0.

High-energy hadrons passing through material render their energy through production of particles, excitation of nuclei, spallation or ionisation, which results in the formation of hadronic showers. In ATLAS, such processes occur in the hadronic calorimeter (HCal). A part of the energy originating from spallation or excitations of nuclei, or carried by neutrinos, escapes the detector. In the so-called *compensating* calorimeters, such as the ATLAS HCAL, these energy losses are compensated for. The spatial development of the shower is characterised by the so called interaction length λ_I.

The ATLAS calorimeters are constructed with sampling geometry, meaning that particle showers develop in a passive material made of heavy atoms (absorber), while the signal produced in the shower in form of ionisation charge or scintillation light is collected in the light active material. The energy resolution of a calorimeter is given by the following expression:

$$\frac{\sigma_E}{E} = \frac{a}{E} \oplus \frac{b}{\sqrt{E}} \oplus c. \tag{3.5}$$

The first term of the equation 3.5, called noise term, parametrises instrumentation effects, such as noise or pedestal. The second term is the so called stochastic term, which takes into account statistical fluctuations related to the shower development, the material or the geometry. For majority of detectors, this is the dominant term. The third term is the so-call constant term, which describes the imperfections of the detector.

The ATLAS calorimeters cover the range $|\eta| < 4.9$. The detector construction over this range is adapted to the physics processes of interest as well as the radiation environment. In the central region matched to ATLAS ID is, fine granularity of the ATLAS EMCal is dedicated to precision measurements of electrons and photons. The rest of the calorimeter is built with a coarser granularity, sufficient for jet and $E_{\mathrm{T}}^{\mathrm{miss}}$ measurements.

The signal pulses created in the sensitive cells of the calorimeter components are read out by the on-detector Front-End Boards (FEBs), then transferred to the off-detector Read-Out Drivers (RODs), and finally sent to the Data Acquisition system (DAQ) via an optical link.

3.2.3.1 Electromagnetic Calorimeter

The ATLAS EMCal is installed around the solenoid magnet with a 6.4 m long and 53 cm thick barrel (EMB), covering the region $|\eta| < 1.475$, and two 63.2 cm thick endcaps (EMEC) with a radius of ~ 2 m, extending the pseudo-rapidity coverage to $|\eta| < 3.2$. The EMCal is a sampling calorimeter with liquid argon (LAr) as active material and lead as absorber. This is a robust and efficient combination of materials which takes advantage of the radiation hardness of LAr and high atomic number of lead. The absorber layers are shaped in an accordion geometry. This arrangement ensures full azimuthal coverage. The read-out electrodes made of copper and kapton clad are placed between the absorber plates. The space between plates is filled with LAr. The basic units of the EMCal are modules (Figure 3.10 left) built in such sandwich geometry with the dimensions varying among different parts of the detector. The LAr is kept at a temperature of 88 K by a cryostat system, while the calorimeter electronics operates at the room temperature. The nominal energy resolution of the electromagnetic calorimeter is $\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$.

The EMB is split into two half-barrels in the $\eta = 0$ plane and separated by a 4 mm gap. The thickness of the EMB corresponds to ~ 22 radiation lengths (X). The EMEC are composed of two coaxial wheels: an inner ($|\eta| < 2.5$) and an outer ($2.5 < |\eta| < 3.2$) wheel. Their thickness corresponds to ~ 24 radiation lengths. The full barrel and the endcaps in the region up to $|\eta| < 1.8$ are complemented by a pre-sampler, an active LAr layer with a moderate granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$ and a thickness of 11 mm in barrel and 5 mm in endcaps, which corrects for electron and photon energy losses upstream of the calorimeter.

The EMB is divided into three coaxial layers with different granularity. The first layer has a fine η-granularity of $\Delta\eta \times \Delta\phi = 0.025/8 \times 0.1$ in the region $|\eta| < 1.40$ and somewhat coarser η-granularity but finer ϕ granularity of 0.025×0.025 in $1.40 < |\eta| < 1.475$. The thickness of the first EMB layer corresponds to ~ 4.3 radiation lengths. This layer is dedicated to discriminate between close-by hard-process photons and those originating from π^0-decays. The middle layer, in which the bulk of the shower is developed and absorbed, has a $\Delta\eta \times \Delta\phi$ granularity 0.025×0.025 in the region $|\eta| < 1.40$ and a coarser η-granularity 0.075×0.025 in the region $1.40|\eta| < 1.475$. The thickness of the middle layer is equivalent to 16 radiation lengths. The third layer, responsible for absorbing the last and least energetic cascades of the

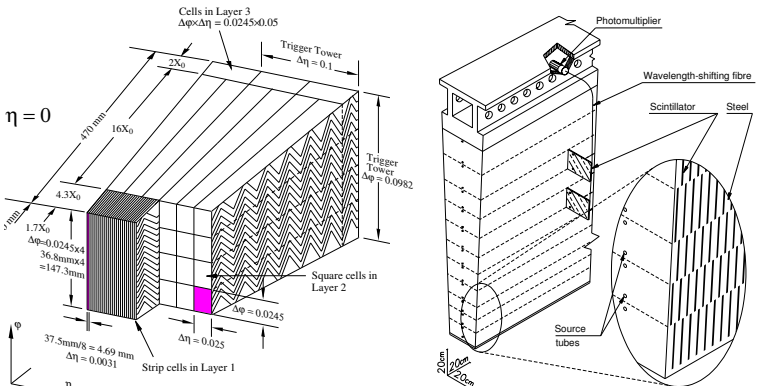


Figure 3.10: Left: sketch of a LAr EMCal barrel module showing layers with different granularity; right: a tile HCal element showing the mechanical assembly and the optical readout [99].

shower, covers the region $|\eta| < 1.35$, and has a depth of 2 radiation lengths with a granularity of 0.050×0.025. The EMB anode operates at a high voltage potential of +2 kV.

Two EMECs have fully symmetric construction. Each EMEC consists of an outer wheel which has three layers in the region $1.5 < |\eta| < 2.5$ and two layers in the region overlapping with the EMB $1.375 < |\eta| < 1.5$, and an inner wheel covering up to $\eta = 3.2$ with the finer granularity at lower $|\eta|$. Depending on the $|\eta|$ range, the EMEC anodes operate on high voltages in the range from 1-2.5 kV.

The signal pulses produced in ECal have rather long tails and at the same time a rather long integration time. This is why the ECal signal of an event often suffers a contribution from a subsequent bunch-crossing. In other words, the ECal is rather sensitive to the out-of-time pile-up.

3.2.3.2 Hadronic Calorimeter

The ATLAS HCal consists of a barrel and two endcaps built in different architectures.

The HCal barrel (HCB), covering the central region within $|\eta| < 1.0$, and two barrel extensions, covering $0.8 < |\eta| < 1/7$, are built as a sampling calorimeter with steel as absorber and scintillating tiles made of polystyrene as the active material. The tiles have a radial orientation (Figure 3.10 right), normal to the beam line. It surrounds the ECal, in the radial range of 2.28 to 4.25 m. With 64 azimuthal modules, the barrel and its extensions cover almost the full azimuthal angle. The barrel and its extensions are radially segmented in three layers of 1.5, 4.1 and 1.8 interaction lengths in thickness for the barrel and 1.5, 2.6 and 3.3 for the extensions. This makes an overall calorimeter thickness of 7.4 interaction lengths. In the first and second layer, tiles have a fine segmentation of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ while in the third the granularity is slightly coarser with 0.2×0.1. An ionising particle traversing a tile produces ultraviolet light which is collected with wave-length shifting optical fibres, grouped into the photomultipliers. The tile calorimeter components operate at the room temperature.

The HCal has two endcaps (HEC) made in sampling geometry with copper plates as absorber and LAr as active material. This choice had been made because of the good radiation hardness of LAr compared to scintillation tiles. Each HEC contains two wheels (HEC1 and HEC2) composed of 32 identical wedge-shaped modules, covering the range $1.5 < |\eta| < 3.2$. Each wheel consists of two depth segments. The wheels are composed of parallel copper plates, 25 mm thick close to the IP and 50 mm thick further away, serving as absorber. The plates have

an outer and an inner radii of 203 cm and 47.5 cm respectively, except in the overlap region with the forward calorimeter, where the inner radius is 37.2 cm. The plates are sandwiched with the 8.5 mm gaps filled with LAr, acting as the active material. Three electrodes dividing the gap in four equal gaps are immersed into LAr to collect the ionisation charges produced by the particles of the hadronic shower traversing the HEC. The overall thickness of a HEC is equivalent to 10 interaction lengths. The readout cells in form of pads placed onto the central foil of each gap. They have a resolution $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the region $1.5 < |\eta| < 2.5$ and 0.2×0.2 for higher $|\eta|$. To minimise the lost space and dead material, the HECs are contained in the common cryostat vessel with the EMECs and forward calorimeters. The ATLAS HECs operate at a high-voltage of 1.8 kV. As the other LAr calorimeters, HEC LAr components operate at a temperature of ~ 88 k.

The nominal energy resolution of the hadronic calorimeter is $\sigma_E/E = 50\% \sqrt{E} \oplus 3\%$.

3.2.3.3 LAr Forward Calorimeter

Two forward calorimeters (FCals) of a total thickness of approximately 10 interaction lengths are inserted in the narrow space between the beam line and calorimeter endcaps, EMEC and HEC, at both sides of the detector, to extend the η-coverage of the calorimeter system to the region $3.1 < |\eta| < 4.9$. Each FCal is composed of three modules along the z-direction. All FCals modules use LAr as the active material. The inner module of a FCal has copper as absorber and serves as electromagnetic calorimeter. The other two modules have absorber plates made of tungsten and are optimised to measure the energy of hadronic showers. The FCals share the cryostat vessels with EMECs and HECs. The nominal energy resolution of the forward calorimeter is $\sigma_{p_T}/p_T = 100\% \sqrt{E} \oplus 10\%$.

3.2.4 Muon Spectrometer

The muon spectrometer (MS) is constructed as a tracker detector situated outside of the calorimeters aimed to measure momentum and traces of muons, the only charged particles not absorbed in the calorimeters. In the MS barrel (MSB), the sensitive elements are arranged in three cylindrical layers parallel to the beam line, placed at radii of approximately 5 m, 7.5 m and 10 m, covering together the range $|\eta| < 1$. Two MS endcaps (MCEC) consist of four wheels perpendicular to the beam line each, located at approximately 7.4 m, 10.8 m, 14 m, and 21.5 m away from the IP in $|z|$ and extending over the pseudo-rapidity range $1 < |\eta| < 2.7$. The MS is installed inside of the ATLAS toroid magnet system which bends the muon traces. The nominal transverse momentum resolution of the MS is $\sigma_{p_T}/p_T = 10\%$ at $p_T = 1$ TeV. The resolution measured on the Z-boson pole mass reconstructed as a di-muon in the 2011 ATLAS data and that expected from the simulation are shown in figure 3.16.

The MS is composed of muon chambers of different technologies. Precision measurement of muon track coordinates in the principal bending plane is performed across most of the η-coverage by Monitored Drift Tubes (MDTs), and by Cathode Strip Chambers (CSCs) at higher pseudo-rapidity ($2 < |\eta| < 2.7$). Position measurement in the non-bending plane and triggering of the events containing muons is performed by the Resistive Plate Chambers (RPCs) in the barrel region and Thin Gap Chambers (TGCs) in endcaps, covering together the range $|\eta| < 2.4$.

An MDT is a pressurised drift tubes made of 0.4 mm thick Al with a diameter of 29.970 mm,

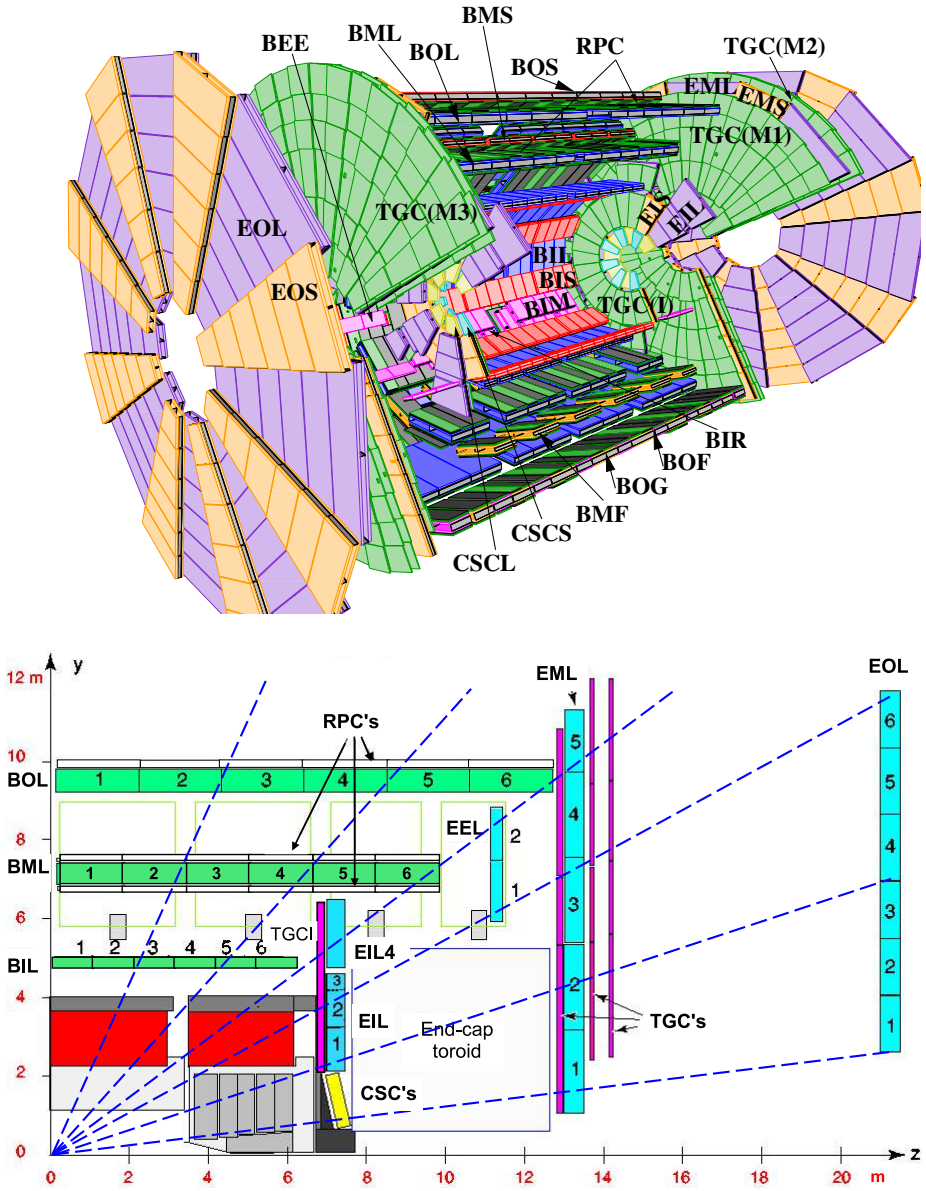


Figure 3.11: Principal MS configuration (top) and the MS cross section along the beam axis in the bending place showing its four chamber sub-systems: the tracking chambers MDT's and CSC's as well as the trigger chambers RPC's and TGC's. The MDT components are labeled with the first letter B (E) corresponding to the barrel (endcap). The CSC components are CSCL and CSCS. TGC wheels M1, M2 and M3 are shown in light green colour. [99].

operating with a gas mixture composed of Ar (93%) and CO_2 gas (7%) at 3 bar. MDTs are used for muon tracking in the region $|\eta| < 2.7$. Ionisation electrons are collected at the gold-plated W/Re (97%/3%) anode with a diameter of 50 μm, placed in the centre of the tube, and operating at 3080 V. The MDTs are arranged in the MDT chambers along ϕ in both barrel and endcaps. The spatial resolution of a single tube is about 80 μm, which allows for a resolution of maximally 35 μm for two multi-layer chambers. MDTs operate safely with counting rates of up to 150 Hz/cm^2, which is exceeded for $|\eta| > 2$.

In the higher η-range and more challenging particle fluxes, muon tracking relies on CSCs, which can operate with the counting rates up to about 1 kHz/cm^2. CSCs are multi-wire proportional chambers, composed of four parallel planes made of parallel anode wires, installed such that the central wire points in the radial direction. Each CSC contains two cathodes, one parallel to the anodes and the other one perpendicular to them in order to allow for a precise coordinate measurement. The distance between the wires is 2.5 mm. Both cathodes are segmented in strips. The electrodes operate at the nominal voltage of 1.9 kV. The CSCs provide a spatial resolution of 40 μm in the bending plane and 5 mm in the non-bending direction. The time resolution achieved by the CSCs is typically 7 ns and is dictated by the electron drift time, which is shorter than 40 ns. The CSC panels are installed 7 m away from the IP along the z-direction and about 90 cm away from the beam axis in the radial direction, covering the pseudo-rapidity range $2 < |\eta| < 2.7$ in the innermost layer.

An RPC consists of a pair of resistive plates, made of phenolic-melaminic plastic laminate, with a 2 mm space between them kept with insulating spacer, and filled with a gas mixture composed of $C_2H_2F_4$ (94.7 %), Iso $- C_4H_{10}$ (5 %), and SF_4 (0.3 %). The plates are kept under a voltage of 9.8 kV. Metal strips mounted on the resistive plates make a capacitive coupling, which enables signal readout and position measurement. RPCs are covering the barrel region within $|\eta| < 1.05$. Their spatial resolution is 10 mm in each direction.

TGCs are multi-wire proportional chambers with the wires arranged with 1.8 mm among them and with a distance of 1.4 mm to the cathodes, and operating at 2.9 kV. They are used in the endcaps, covering the region $1.05 < |\eta| < 2.4$. The chambers are filled with a highly quenching gas mixture of CO_2 and n$-C_5H_{12}$. The chambers operate in a quasi-saturated mode with a gas gain of $\sim 3\times10^5$. The TGCs provide a measurement in the bending plane by collected ionisation charges on the wires, and in the non-bending plane by the strips mounted on the cathode side perpendicular to the wires. The spatial resolution of the TGCs in the radial direction is 2-6 mm, and in the azimuthal direction 3-7 mm. The time resolution of TGCs is 4 ns.

3.2.5 Forward Detectors

In addition to the detectors described in the previous sections, dedicated to measure the physical quantities relevant for the hard process, the forward region of the ATLAS detector is instrumented with a number of detectors specialised for luminosity measurement, monitoring of the LHC beam conditions and low-p_T trigger, shown in Figure 3.12.

The Beam Condition Monitor (BCM) consists of four modules, each consisting of a pair of 500 μm thick diamond sensors of surface 1 cm $\times$ 1 cm, specialised to measure luminosity bunch-by-bunch and monitor beam anomalies. The BCM is situated only 5.5 cm away from the beam axis, and is thus exposed to enormous amount of radiation. High energy charged particles traversing BCM detector produce ionisation charge, which are collected at electrodes operating

at 1 kV. The BCM readout is very fast, with a latency at the order of ns.

The Minimum Bias Trigger Scintillators (MBTS) are placed 3.6 m away from the IP along the z axis, covering the η-range $2.09 < |\eta| < 3.84$, with a task to trigger on the low-p_{T} (minimum bias) scattering events. The MBTS contain 32 scintillator counters installed at inner walls of the EMEC cryostat. The scintillation light is collected by the wave-length shifting fibers, which transport it to photomultipliers, which transform it into electronic signal.

Two Luminosity Measuring Čerenkov Integration Detectors (LUCIDs) serve for measuring relative luminosity in ATLAS. They are placed 17 m away from the nominal IP along the z-direction and cover the region $5.6 < |\eta| < 6$. A LUCID is composed of 20 aluminium tubes filled with C_4F_{10} gas. The tubes are 1.5 m long and have a diameter of 15 mm. Charged particles traversing the gas produce Čerenkov light, which is detected by photomultipliers. Thank to its good time resolution, of the order of few ns, LUCID is capable of observing individual LHC bunches.

Two Zero Degree Calorimeters (ZDCs) are installed 140 m away form the nominal IP along the z-axis. They are dedicated to energy measurement of neutral particles (neutrons, photons) at a very forward region, $|\eta| > 8.3$. ZDCs also serve as additional low-p_{T} trigger aimed to suppress the contribution of the beam-gas events. ZDCs are built in sampling geometry, with one electromagnetic and several hadronic modules consisting of tungsten and steal absorber plates and quartz strips as active material. The signal produced in quartz by ionising particles is Čerenkov light, which is then collected by photomultipliers.

The beam pick-up detectors (BPTX) are trigger elements laying 175 m away from the IP at the beam direction and providing highly accurate timing signals for bunches passing through the ATLAS detector. These signals are used in the trigger chain as well as for keeping track of the collision time relative to the LHC clock and beam conditions monitoring.

The most distant ATLAS components located at even 237 away from the IP are the ALFA (Absolute Luminosity For ATLAS) detectors, the forward detectors measuring absolute luminosity of the ATLAS events by detecting the elastic pp-collisions in the range $10.6 < |\eta| < 13.5$. ALFA detectors are based on the roman pot stations, consisting of a pair of modules each made of

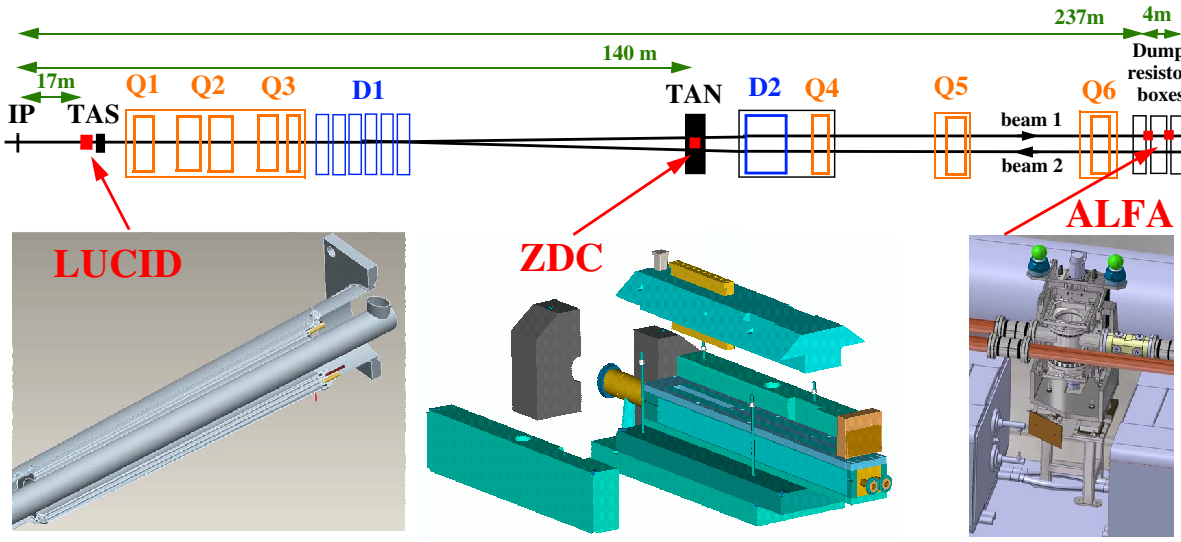


Figure 3.12: Positions of the forward detectors with respect to the IP [99].

scintillating fibres, placed one below and the other one above the beam.

3.2.6 ATLAS Detector Performance

The ATLAS detector was designed to achieve a specific level accuracy defined for each detector component and summarised here in Table 3.1.

Sub-detector	Design resolution	η measurement	η trigger
ID	$\sigma_{p_T}/p_T = 0.05\% p_T \oplus 1\%$	± 2.5	
EMCal	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	± 3.2	± 2.5
HCal	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	± 3.2	± 3.2
FCal	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$	$3.1 < \lvert\eta\rvert < 4.9$
MS	$\sigma_{p_T}/p_T = 10\%$ at $p_T = 1$ TeV	± 2.7	± 2.4

Table 3.1: Design resolution and pseudo-rapidity coverage for measurements and for triggering of the ATLAS sub-detectors. Taken from [99].

The performance of the individual sub-detectors observed in the 2011 collision data is illustrated in the following figures.

Figure 3.13 shows that the vertex resolution in the transverse (longitudinal) plane ranges between $\sim 20\mu$m ($\sim 40\mu$m) for vertices to which tracks with a high total p_T are associated, and $\sim 400\mu$m ($\sim 700\mu$m) for those with a low total track p_T.

Figure 3.14 shows the invariant mass of the Z-boson reconstructed as the invariant di-muon mass using only the ID tracks corresponding to the combined muons. A mass resolution of ~ 6 GeV (~ 10 GeV) is achieved if both muons used are reconstruction in the barrel (positive endcap) region.

Figure 3.15 illustrates the energy resolution of the LAr calorimeters. The invariant mass of the Z-boson can be reconstructed as the invariant di-electron mass with a resolution of ~ 10 GeV in the EMEC and FCal.

Figure 3.16 shows that the di-muon mass resolution varies in the range $2-3$ GeV depending on the pseudo-rapidity region when reconstructed using combined muons.

The understanding of the detector as well as the SM physics processes was refined during the course of the data taking. Collision data at $\sqrt{s} = 8$ TeV recorded in 2012 show a higher level of the agreement between the data and the simulation, as well as an improved detector resolution. The improvements of the ID, EMCal and MS resolutions are demonstrated in Figure 3.17.

3.3 Data Acquisition and Data Processing

The collision events of interest to physics analyses are selected and processed "online" by the ATLAS trigger and data acquisition system [99,114] described in section 3.3.1. The data is then further processed "offline" from raw detector read-out signals to the reconstructed data format used for physics analyses. An overview of the offline processing chain is given in section 3.3.2.

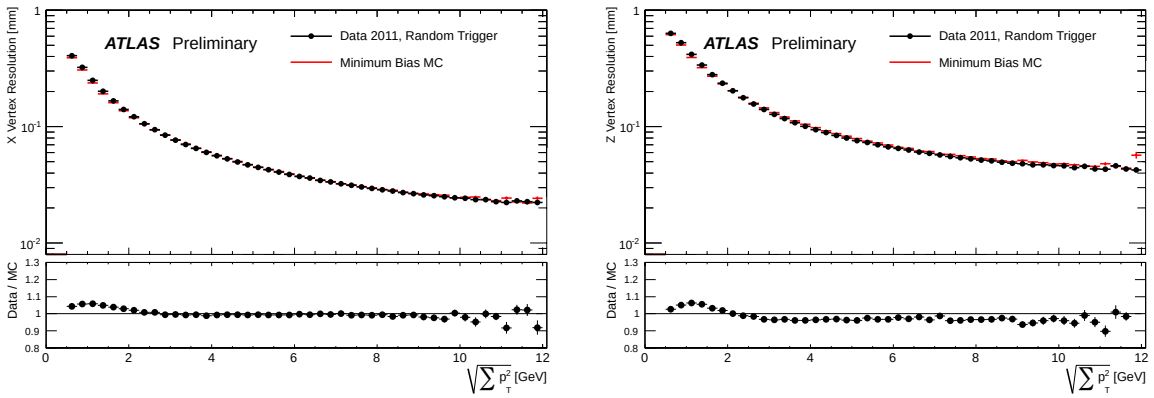


Figure 3.13: Vertex resolution in the x- (left) and z-direction as a function of the sum of the p_T of the tracks associated to the vertex derived from the minimum bias simulated samples and measured in the G4 data taking period of the 2011 data taking campaign with the random trigger. Taken from [107].

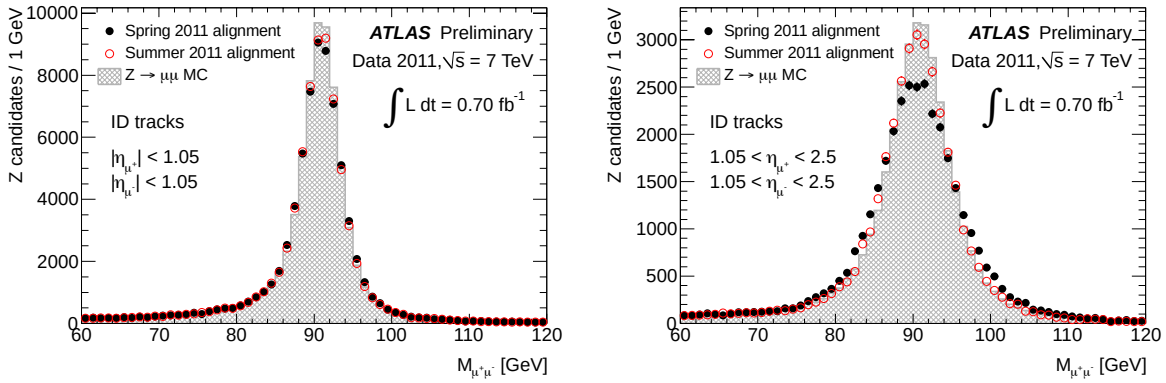


Figure 3.14: Invariant di-muon mass reconstructed using the track parameters of the Inner Detector tracks corresponding to the combined muons measured in barrel region (left) and the endcap A (right), and compared to the simulated $Z \to \mu^+\mu^-$ events. Taken from [108].

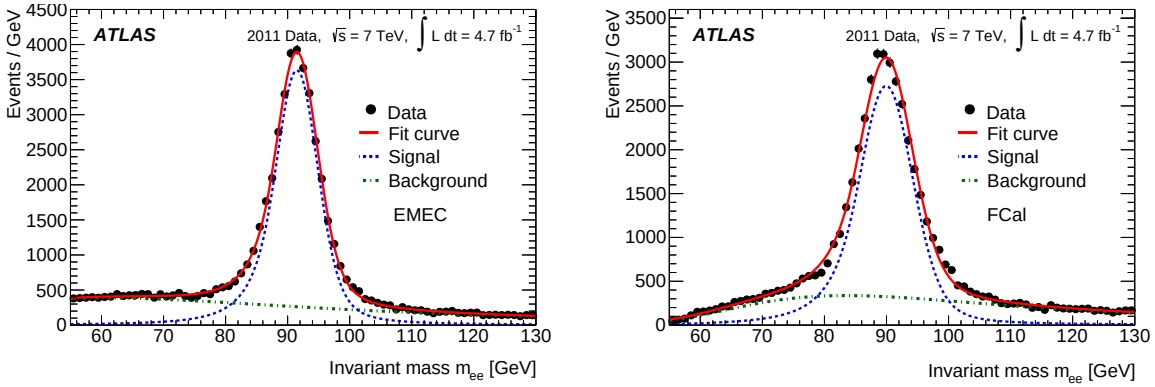


Figure 3.15: Left: The di-electron mass distributions reconstructed using the energy measurements from EMEC [109].

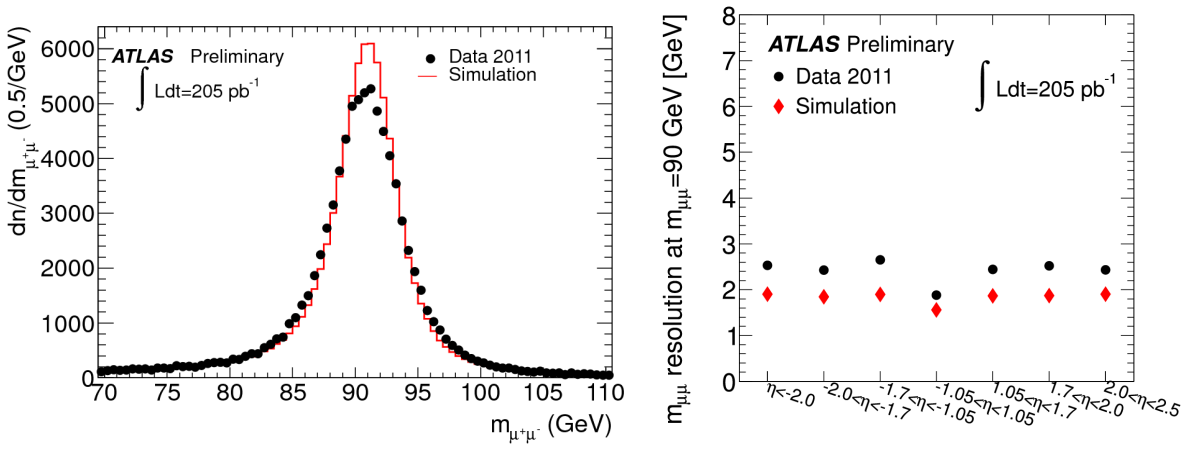


Figure 3.16: Left: Di-muon invariant mass distribution for oppositely charged muon pairs with transverse momentum above 20 GeV and calorimeter isolation. Right: Di-muon mass resolution of combined muons in different pseudo-rapidity regions. Taken from [110]

3.3.1 Trigger and Data Acquisition

The LHC is designed to provide a bunch-crossing frequency of 40 MHz with about 25 proton-proton interactions per bunch crossing at the nominal instantaneous design luminosity of 10^{34} cm^{-2}s^{-1}. Assuming a per-event size of 1.5 MB, 60 TB/s of data would have to be processed and stored to record every collision. The task of the ATLAS trigger and data acquisition system is to reduce this enormous data output by a factor $2 \cdot 10^5$ to about 300 MB/s by selecting and processing the rare events of interest to physics analyses and rejecting the large backgrounds from of low-p_{T} collisions. The ATLAS trigger consists of three levels, called Level-1 (L1), Level-2 (L2), and Event Filter (EF), the latter two being collectively referred to as High Level Trigger (HLT).

The L1 trigger is implemented in custom hardware and is designed to reduce the initial bunch-crossing rate of 40 MHz to 75 kHz. It searches for signatures from high-p_{T} electrons, photons, muons, jets, and τ-leptons, as well as missing transverse energy in the calorimeter and muon system and defines so-called regions of interest (RoIs) in terms of (η,ϕ)-coordinates. The L1 Calorimeter Trigger (L1Calo) relies on the energy deposits summed over sets of calorimeter cells resulting in super-cells of an effective granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. A sliding window algorithm combines a number of neighbouring super-cells to form so-called trigger towers. The information used in the L1 decision is the number of trigger towers in the event with an E_{T} exceeding a given energy thresholds. The L1 muon trigger uses signals in the RPCs in barrel and TGCs in the end caps to identify patterns compatible with high-p_{T} muons. The information used for the trigger decision is the number of muons passing given p_{T} thresholds. In addition there are Minimum Bias Trigger Scintillators (MBTS) and a number of other specialised triggers which can be used e.g. to distinguish the collision events from so-called beam-gas or beam-halo events. The overall L1 decision is made by the Central Trigger Processor (CTP) within a maximum latency of 2.5 μs. While the L1 decision is being made, the complete readout information of the event is temporarily stored in pipeline buffers of the Front-end Electronics (FE), usually installed directly on the detector or close-by. If an event is accepted by L1, the RoIs are passed to the HLT for further processing, while the full event data is transferred from the FE via Read Out Drivers (RODs) to the Read Out Buffers (ROBs).

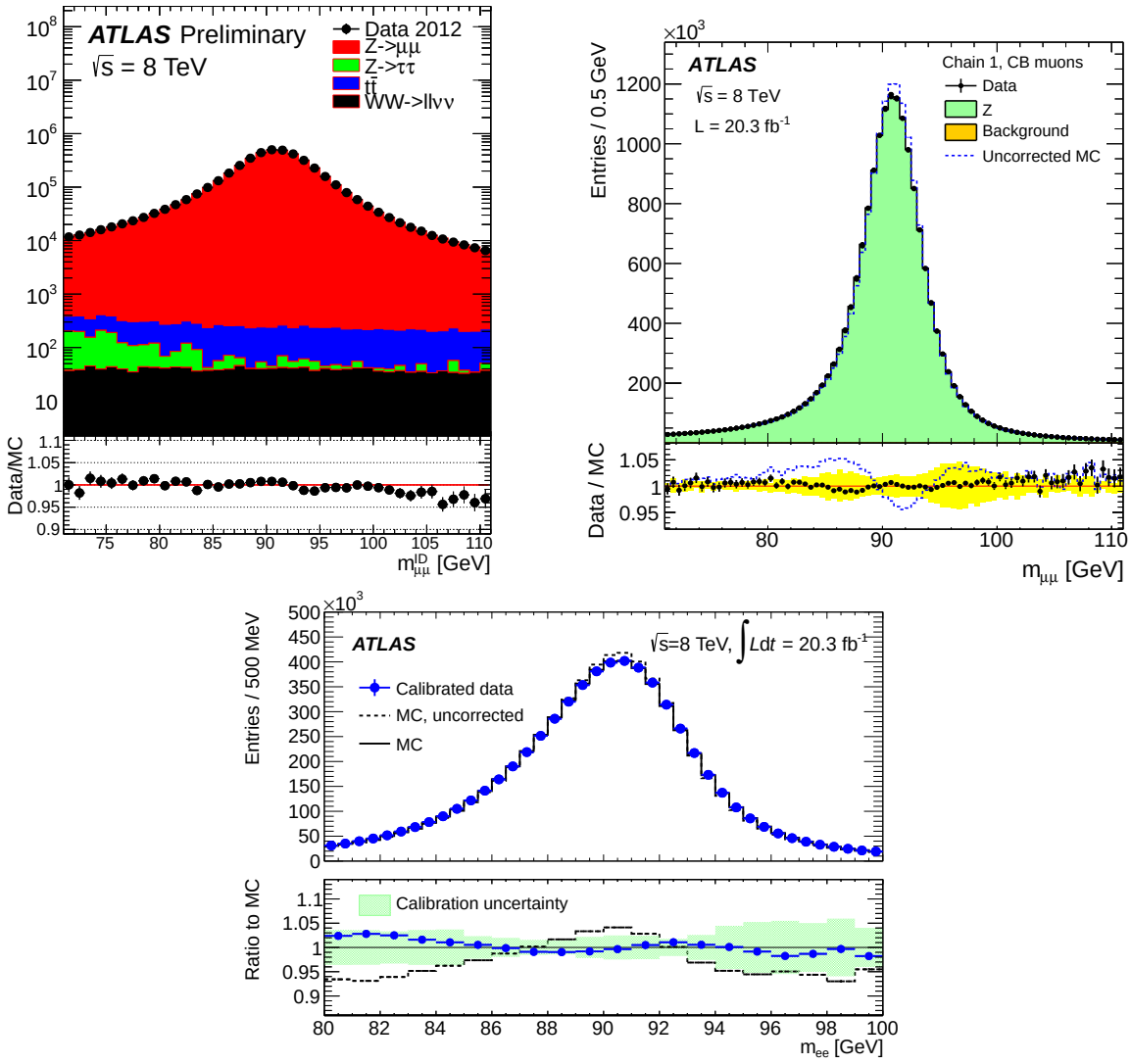


Figure 3.17: Top left: Di-lepton invariant mass distribution at the Z boson pole reconstructed using ID tracks only. The black points show data, while the stacked distribution shows simulated $Z \to \mu\mu$ (red), $Z\tau\tau$ (green), $t\bar{t}$ (blue), and di-boson (black) production. Taken from [111]. Top right: Di-muon invariant mass distribution of $Z \to \mu\mu$ candidate events reconstructed with combined muons. The points show the data, the filled histograms show the simulation with the MC momentum corrections applied and the dashed histogram. Background estimates are added to the signal simulation according to their expected cross sections. The sum of background and signal MC is normalised to the data. Taken from [112]. Bottom: Di-electron invariant mass distribution for $Z \to ee$ decays in data and simulation. Energy scale corrections are applied to the data. The simulation isnormalised to the number of events in data. Taken from [113].

The RoIs, covering typically $\sim 2\%$ of the total event data, are passed to the software-based HLT running on CPU farms. The first step of the HLT, the L2 trigger, executes fast reconstruction algorithms inside these RoIs using high granularity information from all detector components including the ID. The L2 trigger is designed to further reduce the event rate from 75 kHz to 3.5 kHz, with an average processing time of about 40 ms.

In the final processing step of the HLT, the EF, sophisticated reconstruction algorithms using full precision information over the entire detector acceptance are used to refine the trigger decision. With an average event processing time of about 4 s, the EF reduces the event rate to roughly 200 Hz, at which events are written to the CASTOR (Cern Advanced STORage manager) storage system. The events accepted by the EF are arranged into so-called data streams. Based on the trigger selection the events are written to the `Egamma`, `Muon`, `JetTauEtmiss` or `MinBias` physics streams. These streams are not exclusive, however they are defined in the way to minimise the overlap between them. The so-called `Express` stream, composed of a small portion of events from different physics streams, is dedicated to the offline data quality monitoring, as well as the calculation of detector calibration constants. In addition there are `Calibration` and `Debug` streams which are used for performance studies and the analysis of processing errors.

Triggers are organised in so-called trigger chains, a combination of a L1 item and an HLT trigger chain. The L1 item defines the requirements on the type, energy threshold, and multiplicity of the trigger objects at L1. The HLT trigger chain consists of a sequence of L2 and EF algorithms that reconstruct the physics objects of interest ("feature extraction algorithms") and apply selection criteria ("hypothesis algorithms") on them. The so-called trigger menu defines the set of trigger chains to be used for a given data taking period. The trigger chains used for the analyses in this thesis are further described in Section 7.4.1.

3.3.2 Offline Data Processing

The collision events recorded by the ATLAS trigger and data acquisition system are further processed "offline" at the CERN Tier-0 computing centre in order to prepare them for physics analysis. After evaluation of the data quality and detector calibrations based on a subset of the data from the `express` stream, the physics objects and quantities are reconstructed from the RAW data and stored in a format called Event Summary Data (ESD) along with detector-level hit information. The per-event ESD size is approximately 1.5 MB. The data is then further processed into the Analysis Object Data format (AOD) which contains a subset of the ESD information and brings a reduction of the per-event size of approximately a factor of 10. For analysis purposes further derivations of ESDs and AODs in the so-called Derived Physics Data (DPD) format are used. DPDs are often tailored to specific analyses and have unnecessary objects, object properties, or events removed in order to optimise file size and processing time. The most widely used DPD format is the so-called D3PD, in which events are stored in the form of ntuples defined by CERN's object oriented analysis framework ROOT [115]. The reconstructed data in the above described data format is distributed through the global computing grid to make it available to ATLAS institutions world-wide.

Chapter 4

Event Simulation

Proton-proton collision data recorded by the ATLAS detector are analysed by examining the compatibility of the kinematic distributions derived from data to those simulated according to the state-of-the-art understanding of the pp collision physics and the experimental setup. In this chapter, basic concepts and mechanisms used in the simulation of pp collision events are explained. The concrete prescriptions used to simulate event samples used in the searches presented in this thesis are discussed in Section 7.4. Event simulation is a complex procedure divided into three major steps. Simulation of the physics process including the modelling of the incoming protons, their collision and the subsequent event development up to the decays into stable particles, is performed by the so-called event generation explained in Section 4.1. As the physics processes occurring in pp collisions are probabilistic in nature, event generation involves random numbers, which are generated using Monte Carlo (MC) techniques. Simulation of the detector and its response to the particles traversing it is performed by the so-called detector simulation described in Section 4.2. Finally, from the simulated energy deposits left by the particles in the sensitive layers of the detector, signals are produced in the digitisation procedure, briefly described in Section 4.3.

4.1 Event Generation

The modelling of a pp collision requires a detailed understanding of the structure of a relativistic proton, as well as the dynamics of strong interaction at high (perturbative QCD) and low energies (non-perturbative QCD). The main elements of the event generation are summarised in this section.

4.1.1 Hard Process

In our current understanding of the SM, a proton is a bound state composed of point-like quarks (q) and gluons (g) interacting among themselves via the constant exchange of soft virtual gluons. According to the uncertainty principle, the time scale of a virtual gluon interaction is inversely proportional to its virtuality q, i.e. $t \sim 1/q$, so that gluons with higher virtuality usually are absorbed by the same quark by which they were radiated [116]. In a high energy pp collision, quarks and gluons from one incoming proton collide with those from another

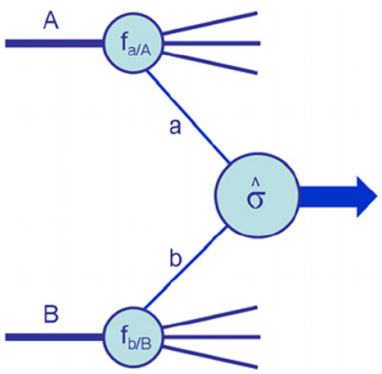


Figure 4.1: An illustration of a hard scattering. A parton of a 4-momentum a originating from a proton of a 4-momentum A undergoes a hard collision with a parton carrying a 4-momentum b coming from a proton which has a 4-momentum B. The spectator partons do not influence the hard interaction and continue their trajectory with slightly distorted directions.

(Figure 4.2), resulting in one or multiple qq, gg or qg collisions. Since the interaction cross section of such collisions decreases as the momentum exchange Q between the colliding particles raises, a pp collision typically consists of one so-called *hard scattering/process* with a high momentum exchange Q, which is sometimes accompanied by further soft collisions, the so-called *multi-parton interactions*, explained in Section 4.1.5. The proton constituents involved in the hard scattering are deflected at a time scale of the order of $1/Q$, which is too short for these constituents to undergo any soft interaction with the rest of the parent proton, so that the hard process can be considered independent from the interactions within the proton. In this limit, the proton is described as a cloud of quasi-free point-like particles, called *partons* by the so-called *Parton Model* [33].

To estimate the production probability of a final state X in a pp collision, all possible parton-parton collisions able to produce X need to be taken into account. In other words, the production cross section $\sigma_{pp\to X}(\sqrt{s})$ of X in a pp collision at a given $\sqrt{s}$ is composed of a sum of partonic cross sections $\hat{\sigma}_{ab\to X}(\sqrt{\hat{s}})$ at a parton-parton centre-of-mass energy $\sqrt{\hat{s}}$, taking all possible combinations of the partons a and b able to produce X, into account (Figure 4.1). To account for the probability of achieving a sufficient $\sqrt{\hat{s}}$ for producing X in an ab collision, the partonic cross section needs to be convoluted with the so-called *parton distribution functions* (PDFs), that describe the distribution of the 4-momentum of the proton among its constituents (see Section 4.1.2). This decomposition of a pp collision into a hard QCD scattering and a soft QCD PDFs is described by the factorisation theorem [117]. This theorem introduces the so-called *factorisation scale* μ_F, which represents the estimated limit between the perturbative and non-perturbative regime of QCD, i.e. the energy at which the running α_s becomes too large for achieving a desired convergence of the development in powers of α_s. Thus, the factorisation scale is a parameter on which the soft QCD observables depend, in particular the PDFs $f_a(x_a, \mu_F)$, which provide the probability that the parton a carries the fraction $x_a = p_a/P_{proton}$ of the total proton momentum. In addition, the cross section calculation depends also on the choice of the renormalisation scale μ_R of QCD, at which α_s is evaluated (see Section 2.2.6). In this picture, the $\sigma_{pp\to X}(\sqrt{s})$ expression reads as:

$$\sigma_{pp\to X}\left(\sqrt{s}\right) = \sum_{a,b} \int_0^1 dx_a dx_b f_a\left(x_a, \mu_F\right) f_b\left(x_b, \mu_F\right) \hat{\sigma}_{ab\to X}\left(x_a, x_b, \sqrt{\hat{s}}, \alpha_s(\mu_R), \mu_F\right). \quad (4.1)$$

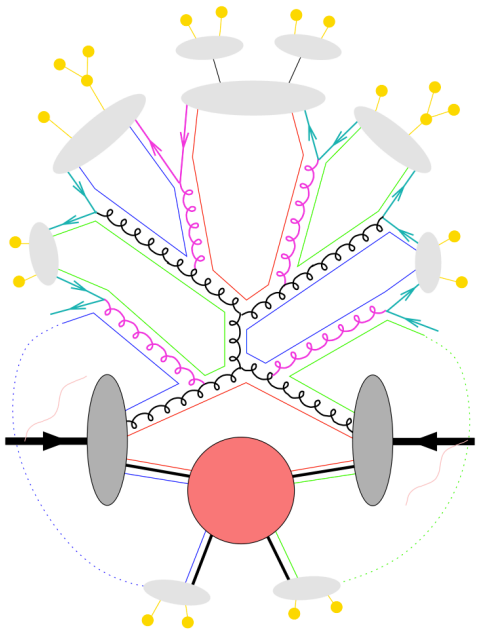


Figure 4.2: A typical *pp*-collision event showing the gluon-gluon hard scattering (black curly lines) with initial and final state gluon radiation (magenta curly lines) and its evolution through the stages of parton shower development (magenta and green strait quark lines), hadronisation (grey areas) and hadronic decays (yellow dots). The underlying event and multiple interactions are collectively shown by the red circle. The color fields are shown as red, blue and green straight lines painted along the gluons and quarks. The blue and green bows connecting the the parton shower products to the underlying event products represent the colour connection between the both parts of the event.

The partonic cross can be expressed as:

$$\hat{\sigma}_{ab\to X}\left(x_a, x_b, \sqrt{\hat{s}}, \mu_F, \mu_R\right) = \int \frac{1}{2x_a x_b \hat{s}} |\mathcal{M}_{ab\to X}\left(x_a, x_b, \sqrt{\hat{s}}, \alpha_s(\mu_R), \mu_F\right)|^2 d\Omega, \tag{4.2}$$

where the factor $1/2x_a x_b \hat{s}$ corresponds to the particle flux, while $|\mathcal{M}_{ab\to X}|^2$ stands for the square of the matrix element of the scattering in question, averaged over colour and spin states of a and b. Differential $d\Omega$ represents the phase space in which the process $ab \to X$ is physically possible. For a scattering with a sufficiently high momentum transfer, the matrix element can be calculated perturbatively, as an interference of diagrams of different orders. As the factorisation and normalisation scales describe the not precisely known borders between two physics domains, their values are usually related to some quantities characteristic for the modelled process (e.g. the mass of the particle produced). The uncertainties arising from such a convention are typically estimated by comparing the cross section values $\sigma_{pp\to X}$ evaluated using $\mu_n/2$ (μ_n) and μ_n ($2\mu_n$) for the down (up) variation estimation, where μ_n designates the chosen nominal value of μ_F or μ_R.

As the final state particles of the hard scatter move away from the interaction point, their kinetic energy transforms into the potential energy of the strong interaction, which is then converted into quark pair production, i.e. into the development of a *parton shower*, explained in Section 4.1.3. These quarks then recombine into observable particles in the process called *hadronisation*, summarised in Section 4.1.4. The hard scattering is also accompanied by the so-called *underlying event*, described in Section 4.1.5, which is induced by the partons which do not participate in the hard process. The parton shower development, hadronisation and underlying event, illustrated in Figure 4.2, are in general soft processes and as such described by non-perturbative QCD.

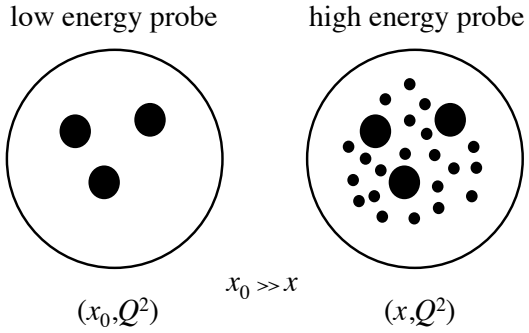


Figure 4.3: Illustration of the Bjorken scaling violation. When probed at low energies, proton appears to consist only of three valence quarks, each carrying a substantial fraction of the proton momentum (left). A high energy probe reveals the sea of quarks and gluons, which take over a fraction of the proton momentum, leaving less momentum for the valence quarks (right).

4.1.2 Parton Distribution Functions

According to the Parton Model, the longitudinal momentum of a relativistically moving proton p is distributed among its constituents - partons. The probability of finding a parton of a type a with a proton momentum fraction in the interval $[x_a,\ x_a + dx_a]$ within a proton p is given by $f_a^p(x_a)\,dx_a$, where $f_a^p(x_a)$ is the *Parton Distribution Function* (PDF). The PDFs reflect the soft interactions inside the proton, and are thus governed by the non-perturbative QCD.

At leading order, PDFs are independent of the momentum transfer Q of the hard process. This suggests that the proton consists of point-like constituents independently of the energy at which it is probed. This picture is referred to as *Bjorken scaling* [118]. At higher probing scales, however, these seemingly point-like partons start to appear as groups of softly interacting particles, resulting in an increased number of partons with a low x and correspondingly a decreased number of partons with a high x in the proton (in other words, an increase of the corresponding PDF at low x and a decrease at high x). This behaviour is known as the *breaking of the Bjorken scaling* and manifests itself through higher order corrections.

The higher order diagrams, however, are sensitive to the existence of soft and collinear divergences, i.e. to the radiation of particle with a negligible transverse momentum compared to the hard process. These divergences can be absorbed in the Q-dependent formulation of the PDFs. This feature has an important consequence. If a PDF is known for a given Q_0, it can be evaluated for a different Q using an evolution given by the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [119–121]. This scale dependence of the PDFs motivates the introduction of the so-called *factorisation scale* μ_F, which becomes a parameter of PDFs and thereby also of the production cross section given in Equation 4.1.

However, PDFs still need to be determined for some Q_0. The usual procedure of determining the PDFs is to obtain them by fitting the experimental data from deep-inelastic scattering experiments, in which the hadron is probed by a point-like probe (usually an electron or another proton) with known energy. There are several different schemes for fitting PDF sets. The schemes provided by the CTEQ [122] and MSTW [123] collaborations are in most common use among the colliding experiments. For an illustration, proton quark and gluon PDFs from the CTEQ6 PDF set are shown in Figure 4.4 for two different Q^2 values.

The proton consists of two u and one d valence quarks, and a sea of gluons and quark-antiquark pairs that are constantly produced and annihilated. A relativistic proton is modelled by a set of PDFs, which contains one function for each kind of partons producible at the momentum transfer Q [124].

4.1.3 Parton Shower

Partons involved in a hard scatter process normally have a very high energy (> 1 GeV) for which α_s is small ($\ll 1$). At such energies, quarks and gluons are likely to radiate off a gluon, carrying a portion of the energy of its mother-particle and a colour connection to it. These gluons can then decay into further gluons or quark-antiquark pairs. This process continues as long as the particles produced have sufficient energy to reach a distance from the initial particle at which the colour field breaks up into a quark-antiquark pair. In this way a so-called *parton shower* (PS) is formed.

The formation of a PS is a non-perturbative process. The PS contribution to the hard process cross section is thus estimated by including only the dominant contribution to each order [126]. These dominant contribution usually come from the nearly collinear parton splitting, or the low-energy gluon radiation. The correction due to collinear splitting is estimated as the contribution which an addition of a parton collinear to one of n partons in an assembly makes to the differential cross section $d\sigma_n$ of a given process involving only these n partons. Starting from $d\sigma_n$, a differential cross section for $n+1$ particles is calculated by integrating over all possible opening angles θ and azimuth angles ϕ of the additional parton, and parametrising the probability that the new particle j carries a fraction z of the energy of its mother-particle i by a splitting function $P_{ji}(z, \phi)$:

$$d\sigma_{n+1} \approx d\sigma_n \frac{\alpha_s}{2\pi} \frac{d\theta^2}{\theta 2} dz\, d\phi\, P_{ij}(z, \phi). \tag{4.3}$$

The values of z, θ and ϕ are quasi-random numbers obtained from a MC generation according to the $P_{ji}(z, \phi)$ distribution. The shower formation continues as long as the virtual mass squared

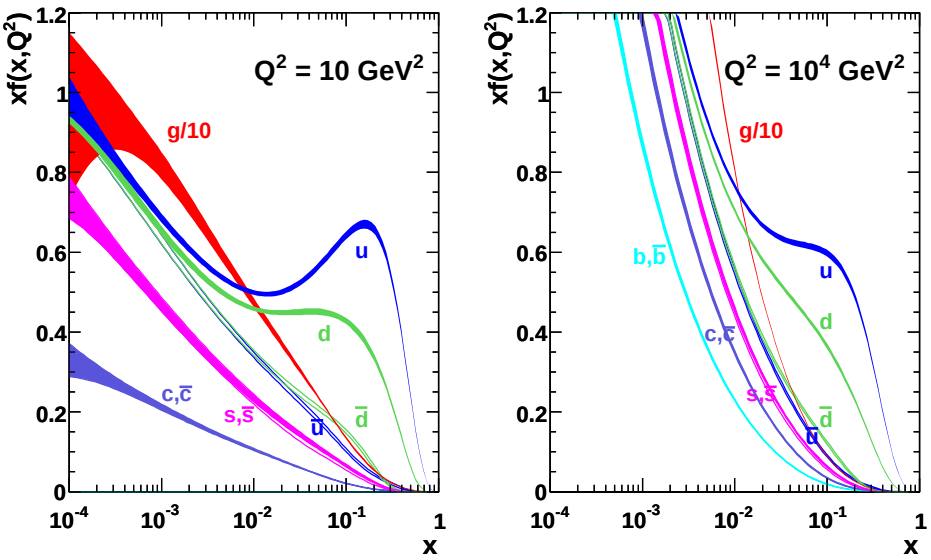


Figure 4.4: Proton PDFs for u, d, c and s quarks as well as the gluon from the MSTW 2008 PDF set fitted at NNLO for a Q^2 of 10 GeV (left) and 10 TeV (right). The u- and d-quark PDFs are showing pronounced enhancement for the high x-values due to valence quarks over the monotonically falling sea-quark distribution. Sea-quark and gluon PDFs grow significantly with increasing Q^2 while the valence quarks PDFs accordingly decrease. Taken from [125].

(or virtuality) q^2 of the particles in the shower is higher than the so-called hadronisation scale, at which particles are no longer able to overcome confinement.

Virtual loop corrections to the hard process are taken into account by calculating the probability that a particle loosing it's virtuality from the initial q_i the final q_f does not split. These probabilities are given by the so-called Sudakov form factor:

$$\Delta_i(q_i^2, q_f^2) = \exp\left(- \int_{q_f^2}^{q_i^2} \frac{dq^2}{q^2} \frac{\alpha_s}{2\pi} \int_{Q_0^2/q^2}^{1-Q_0^2/q^2} dz \int_0^{2\pi} d\phi \, P_{ij}(z, \phi) \right), \tag{4.4}$$

assuming that $d\theta^2/\theta^2 = dq^2/q^2$. Along with the virtual loops, this expression also accounts for unresolvable splittings, i.e. such with a too low energy or a too small splitting angle to be detected.

Final state radiation (FSR), i.e. a gluon radiated off a final state parton is generated through this parton shower procedure. For the initial state radiation (ISR), however, this procedure is not suited, as the momenta of the partons initiating the hard process need to be precisely adjusted to MC simulation of the hard process (for example a gluon decaying into a $t\bar{t}$ pair). Thus, the longitudinal momentum fractions x_1 and x_2 of the incoming partons need to be simulated first, and the momentum and angle of the ISR is found by the backwards evolution, which involves a PDF dependent correction to the Sudakov form factors. Due to the colour-connection, the ISR tends to be aligned with its mother-parton.

4.1.4 Hadronisation

Partons produced in the parton shower reach the distances from their production vertex at which the strong coupling constant $\alpha_s(Q^2)$ becomes large and pulls them into confinement, i.e. forces them to recombine in colourless states. This process is known as *hadronisation*. It occurs in the non-perturbative regime of QCD and thus relies on phenomenological models. Two mostly used such models are the *Lund String Model* [127] and the *Cluster Model* [128].

The Lund String Model is based on the observation from lattice QCD[1] [129], that the potential energy of the colour field between a quark and an antiquark carrying opposite colour charges increases linearly with the distance between the quark and the antiquark, and can be represented as narrow flux tubes. Thus, this field is described as a string stretching between the quark and the antiquark. The presence of gluons in the system modifies the geometry of the field causing so-called kinks in the string. When the string energy overcomes the mass threshold of a given quark-antiquark pair, it can break forming an antiquark to match the original quark, and a quark to match the original antiquark, and leaving three shorter strings with lower potential energy. This procedure continues until the energy of the particles drops below the point at which they can no longer escape the confinement. At that point they combine forming the final state hadrons.

The Cluster Model relies on the pre-confinement property of QCD, which, at an energy scale much lower than that of the hard scattering, pushes the partons in a shower to cluster in colourless groups, proto-hadrons, in a way which does not depend on the underlying hard

[1] Lattice QCD is a method of approximate computation of the QCD phenomena in the non-perturbative regime on a grid of phase space points.

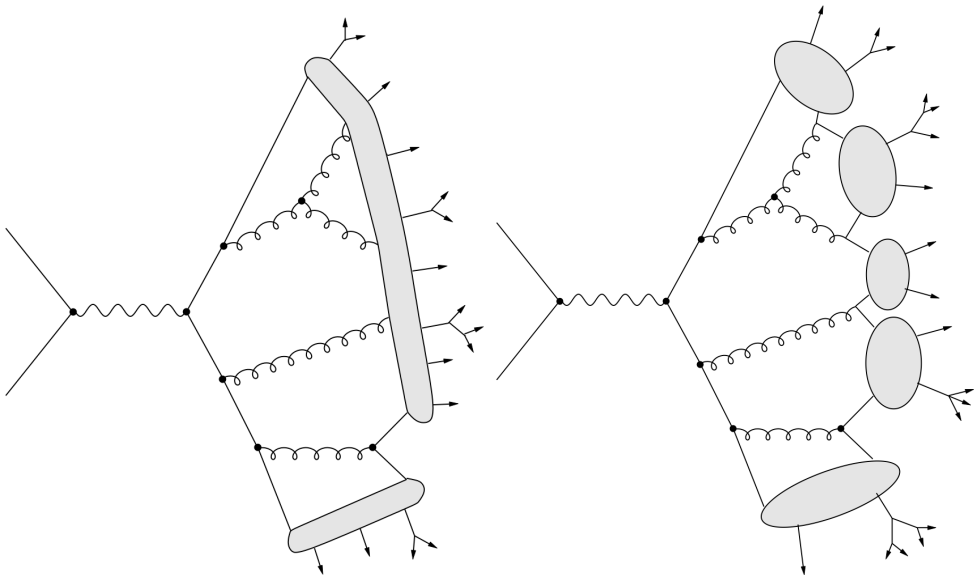


Figure 4.5: Illustration of two hadronisation models: Lund string model (left) and cluster model (right). Taken from [128].

scattering. Thereby, all remaining gluons in the shower are forced to split into quark-antiquark pairs, which participate in the formation of clusters. More massive proto-hadrons subsequently decay into lighter ones. Contrary to the Lund String Model, in which a gluon makes a kink in the string, in the cluster model a gluon imperatively breaks the string.

4.1.5 Underlying Event

The processes discussed so far, i.e. the hard process, its higher order corrections (parton shower) and development (hadronisation), are initiated by one parton from each incoming proton. However, the event development obtains a contribution from additional soft or moderately hard processes, jointly referred to as *underlying event* (UE). The description of the UE employs phenomenological models, because of the non-perturbative nature of the processes involved. The UE is considered to be composed of thee dominant components: *multi-parton interaction*, *beam remnants* and *pile-up*.

Multi-parton interactions correspond to moderately hard interactions of the spectator partons of the incoming protons, that is those partons not participating in the hard process. This interaction typically results in a pair of low-p_{T} back-to-back jets that are colour connected with the rest of the event.

Beam remnants represent the soft QCD activity due to the spectator partons which have not undergone any inelastic scattering. Beam remnants are modelled by conserving the colour connection and momentum within the event.

Pile-up represents the presence of moderately hard interactions in the event originating from multiple inelastic pp scatterings per bunch crossing[2]. This contribution represents a serious

[2]This process is also referred to as in-time pile-up, to be differed form the out-of time pile-up which involves collisions from the subsequent bunch crossing taken into account due to the long integration time of the EM

challenge to the reconstruction of the hard process as the products of pile-up events largely overlap with those of the hard process (although typically originating from a vertex displaced with respect to that of the hard scattering). The mean value of multiple pp interactions per bunch crossing is proportional to the instantaneous luminosity of data taking. Thus, for the collisions at high instantaneous luminosities, as at the LHC, the pile-up contribution becomes significant and requires a precise modelling. Within ATLAS, pile-up activity is modelled by enriching the original hard scattering by adding so-called *minimum-bias events*. Minimum-bias events represent soft collisions involving a colour-singlet (diffractive events) or a colour (non-diffractive events) exchange between the initial protons, resulting in a dissociation of one (single-diffractive events) or both (double-diffractive and non-diffractive events) protons, in which plenty or soft coloured objects are produced. Experimentally, such events are selected by imposing the loosest possible event selection, normally involving only a trigger requirement (e.q. a signal in the MBTS elements), which filters the beam gas and beam halo events out and keeps as many pp collision events as possible (Section 7.4.2).

4.1.6 Event Generators

Event generators are tools which generate collision events and their development as described in the previous section, using the MC method. Generators can be either *general-purpose*, performing all steps in the event development, or *specialsed* for a particular sub-process. Commonly used generators differ in the approach to various elements of event generation. While some use only the LO matrix element calculation, others include higher order correction, as well as various parton shower or hadronisation models. In this section, the general features of the generators used in the studies documented in this thesis are briefly summarised. Details of the event simulation for each individual process is described in Section 7.4.2.

Pythia [130] is a general-purpose generator developed to simulate the full event evolution on e^+e^-, ep and $pp/p\bar{p}$ collision events. Pythia is equipped to simulate around 200 hard processes of the type $2 \to n$ with $n \leqslant 3$ using a LO matrix element. The hard scatters comprise elastic, non-diffractive, single and double-diffractive processes. Pythia uses p_{T}-ordered parton showers in the leading-logarithmic (LL) approximation, and the Lund string hadronisation models. This generator is often used for the in-time pile-up simulation.

Herwig (Hadron Emission Reactions With Interfering Gluons) [131] is a LO matrix element generator for $2 \to 2$ hard processes using an angular-ordered parton shower in the LL approximation, and the cluster hadronisation model. Multi-parton interactions is modelled using the Jimmy package [132]. This package assumes the matter distribution independent from the momentum fraction x in impact parameter space. With this assumption the the multi-parton cross section can be calculated as a function of the hard process cross section, PDFs and area overlap between the colliding protons.

MC@NLO [133] generates events with NLO matrix element which is smoothly matched to LL or nexto-to-leading-log (NLL) parton shower MC simulation. Matching a parton shower approximation to an NLO perturbative calculation requires the resummation of the enhanced higher order terms. This procedure, however, introduces double counting at the same order, which needs to be subtracted. The subtraction scheme used in the MC@NLO generator results in a small and well controlled portion of events with negative event weights.

calorimeter.

AcerMC [134] is a matrix element generator specialised for SM background processes in pp-collisions. It is usually combined with Pythia or Herwig modelling of the event development.

Alpgen [135] is dedicated to generate multi-parton hard processes in hadronic collisions with matrix element calculations at LO in QCD and EW. This generator provides complete parton-level events ready for the further event development perform by some other generator (usually Herwig and Jimmy).

Powheg [136] (the Positive Weight Hardest Emission Generator) is an NLO parton-level generator that produces positive-weighted events. It generates the hardest radiation in the event using the exact NLO matrix element and is normally interfaced with another parton-shower MC generators (usually Pythia or Herwig) for showering and hadronisation.

4.2 Detector Simulation

Fully generated events are processed through the ATLAS detector simulation to model their interaction with the sensitive and dead material of the detector. The full detector simulation makes use of the Geant4 framework [137, 138], which provides the most accurate possible modelling of the detector geometry. However, as this accurate modelling requires enormous computing resources, various fast simulation algorithms have been developed for specialised purposes, based on the parametrisation of the detector response.

The full detector simulation relies on two databases: the *geometry database*, containing the detailed description of the physical geometry of the detector used within the Geant4 framework, and the *conditions database*, containing the variable detector conditions (e.g. information about temperatures or disabled modules), which enables a realistic emulation of run-by-run detector conditions during data-taking. In the full simulation, each particle of a generated event is propagated through the full ATLAS detector using Geant4. During the passage through the sensitive elements of the detector, the particle leaves energy deposits, which are recorded as hits, containing the information about total energy deposition, position, and time of the hit. These hits are passed to the digitisation process (Section 4.3).

Some of the MC samples used in the analyses presented in this thesis make use of the ATLFAST-II fast simulation algorithm [139]. This algorithm was developed to increase the simulation speed while keeping the compatibility of its output with the accuracy requirements for the standard ATLAS reconstruction. ATLFAST-II consists of two components: the Fast ATLAS Tracking Simulation (FATRAS) [140] for the ID and MS simulation, and the Fast calorimeter Simulation (FastCaloSim) [141] for the calorimeter simulation. FATRAS keeps the full geometry description of the sensitive detector elements while the other materials have a simplified description. FastCaloSim does not propagate every particle of a shower through the calorimeter material, but propagates a single particle while parametrising the longitudinal and lateral shower development.

For the analyses presented on this thesis, the fast simulation is only used for studies of few systematic uncertainties, while all the other samples are reconstructed using the full detector simulation.

4.3 Digitisation

Hits in the sensitive detector elements produced in the detector simulation are input to the digitisation process, where they are transformed into so-called *digits*, i.e. voltage or current signals, which serve as inputs to RODs in the detector readout electronics. At this stage, the hits are overlaid with the pile-up activity of minimum bias events and detector noise. Also the L1 trigger hypotheses are evaluated, although no events are discarded. The output of the ROD emulation is provided in the Raw Object Data (ROD) format, in which real data are also provided. Events in the ROD format are passed through the same HLT trigger and reconstruction chain as the real data events.

Chapter 5

Event Reconstruction

As particles traverse a detector they interact with the material leaving their visible trace in a form of energy deposit, ionisation charge or scintillation light. Such raw signals are used to reconstruct objects which approximately correspond to the particles which produced them, their traces or production vertices. Object reconstruction employs complex and sophisticated reconstruction algorithms.

With the dominant decay $T \to Wb$ and the lepton+jets final state assumed in the analyses presented in this thesis, the reconstructed objects present in the experimental signature of $T\bar{T}$ events are electrons, muons, jets (including those recognised as originating from b-quarks), and missing transverse energy due to the presence of a neutrino. The reconstruction of these objects crucially depends on the reconstruction of interaction vertices, ID and MS tracks, as well as the association of energy deposits in the calorimeter. The reconstruction procedure of all these objects is briefly explained in this chapter.

As in the assumed decay mode the $T \to Wb$ events result in exactly the same experimental signature as $t\bar{t}$ events in the same channel, the analyses documented in this thesis make use of the reconstructed objects optimised for the top quark analyses with the 2011 ATLAS data. The t' search described in Chapter 8 and the VLQ search documented in Chapter 9 use slightly different object definitions. As these differences are rather subtle, and as they do not have a great impact on the performance study presented in Chapter 8, only the prescriptions used for the VLQ search are presented here. These prescriptions are described in detail in Section 1 of [142].

5.1 Charged Track Reconstruction

Trajectories of charged particles reconstructed from their signal in the Inner Detector are referred to as *reconstructed tracks*. As the Inner Detector is immersed in the solenoid field (see Section 3.2.1), tracks bend in the transverse (x, y)-plane and their trajectories have an approximately helical shape. When charged particles traverse sensitive detector elements of the Inner Detector, they deposit their energy through ionisation. These energy deposits are read out as *hits*, which are used to form *space points*. Each Pixel hit corresponds to one space point, while SCT space points are formed from pairs of SCT hits from each side of a module.

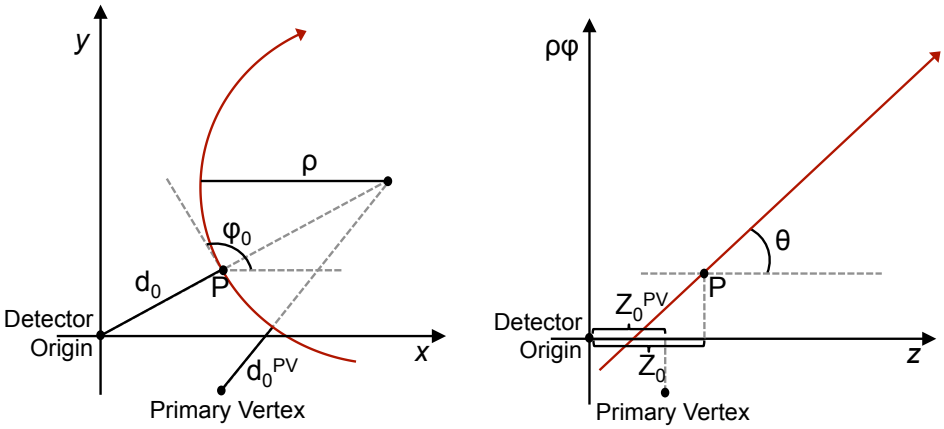


Figure 5.1: Illustration of the geometric definition of the track parameters $\tau = (d_0, z_0, \phi_0 \cot\theta, q/p_\mathrm{T})$ in the (x, y)-plane (let) and the $(\rho\phi, z)$-plane (right). The particle trace is shown by the red line and its perihelion is marked by the point P.

Tracks are reconstructed in two stages. First, space points are grouped into sets that could originate from a single charged particle in the process called *pattern recognition*. Such track candidates then undergo *track fitting* to obtain the best estimate of the parameters by which a track is described. These *track parameters* are usually given by a set $\tau = (d_0, z_0, \phi_0 \cot\theta, q/p_\mathrm{T})$. The track parameters are illustrated in the graphics in Figure 5.1. The point P in the graphics represents the closest point of the track to the detector origin given by the coordinates $(x, y, z) = (0, 0, 0)$, i.e. the perihelion. Accordingly, d_0 and z_0 stand for the track impact parameters in the transverse and longitudinal plane respectively, i.e. the distance from perihelion to the origin, and the z coordinate of the perihelion. These two impact parameters are often expressed with respect to the production vertex of the hard process, i.e. the primary vertex (d_0^{PV}, z_0^{PV}), or the beam spot[1] (d_0^{BS}, z_0^{BS}). The angle ϕ_0 is defined by the x-axis and the track tangent in the point P, and the angle θ by the z-axis and the track plane. The radius of the track curvature is given by $\rho = 1/p_\mathrm{T}$. It is assigned the charge of the particle, determined by the sense in which the track bends.

Track reconstruction employs two complementary pattern recognition algorithms: the main *InsideOut* and consecutive *OutsideIn* reconstruction [144]. The *InsideOut* reconstruction (see Figure 5.2 for illustration) starts with establishing track seeds made of three silicon space points. The track seeds define roads along which the space points are added to a track candidate outwards. Space points shared between two or more track candidates are assigned to exactly one of them using a sophisticated algorithm for resolving ambiguities. In such a way, silicon track candidates are formed and extrapolated into the TRT. Final ID tracks are formed using a track fitter with a χ^2 function as a metric.

The *OutsideIn* reconstruction is performed on the remaining space points after the *InsideOut* reconstruction, seeded in TRT and extrapolated into silicon detectors. This algorithm is targeted to the reconstruction of secondary tracks not originating from a primary vertex. For a charged particle multiplicity measurement, in which it is crucial to reconstruct as many tracks coming

[1]The beam spot represents a region close to the centre of the detector in which pp collisions take place. It is described by a Gaussian distribution with a longitudinal width of $\sim$ 55 mm and radial radial width of $\sim$ 15 μm [143]

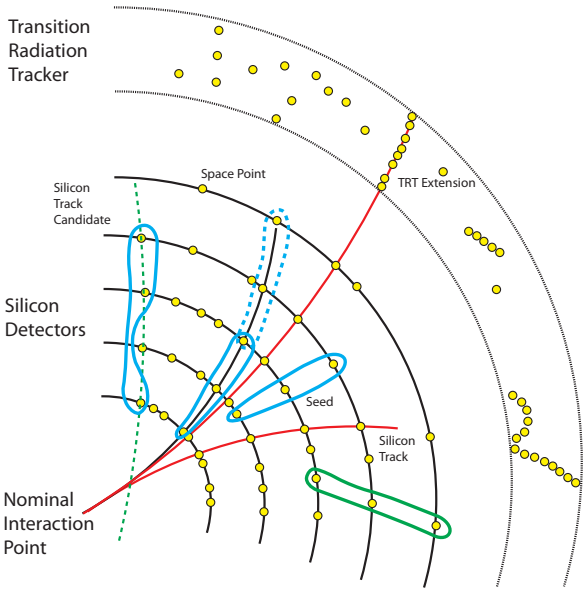


Figure 5.2: Illustration of the *InsideOut* pattern recognition with different reconstruction scenarios in a simplified model of the Inner Detector [145]. The space points are shown by yellow dots. The silicon track seeds are marked by blue contours. The dashed blue contour represents a seed which belongs to the same particle trajectory as another seed. The green contour corresponds to a seed incompatible with a trace coming from the nominal IP. The dashed green line shows a track candidate with one silicon hole incompatible with the nominal IP. The full lines show the fully reconstructed ID track candidates. The longest one among them uses the information from all the sub-detectors while the other two correspond to the silicon track candidates with no TRT extension.

from primary vertices as possible, another *InsideOut*-like sequence with relaxed requirements on the number of silicon hits and impact parameters, called the *LowPt* algorithm, is performed with the space points remaining after the *OutsideIn* reconstruction.

The tracking reconstruction efficiency ϵ_{trk} is estimated from the simulation in bins of p_{T} and η. It is defined as a ratio of the number of reconstructed tracks matched to a generated particle $N_{\text{reco}}^{\text{matched}}$, and the number of generated particles N_{gen}:

$$\epsilon_{\text{trk}}(p_{\text{T}}, \eta) = \frac{N_{\text{reco}}^{\text{matched}}(p_{\text{T}}, \eta)}{N_{\text{gen}}(p_{\text{T}}, \eta)} \tag{5.1}$$

For events with at least 2 charged particles with $p_{\text{T}} \geqslant 100$ MeV and $|\eta| < 2.5$, like those used for the charged particle multiplicity measurement with 2010 data, the tracking efficiency is presented in Figure 5.3. It reaches $\sim 80\%$ in the central region and drops below 50% in the forward region. This efficiency raises steeply from $\sim 10\%$ for the tracks with $p_{\text{T}} \sim 100$ MeV to $\sim 80\%$ for those with $p_{\text{T}} \sim 500$ MeV, where it flattens out, reaching as high as $\sim 90\%$ for tracks with $p_{\text{T}} > 5$ GeV. In analyses with 2011 data the track p_{T} requirement is raised to 400 MeV.

5.2 Primary Vertex Reconstruction

The primary vertex (PV) of an event designates the point of the proton-proton collision vertex. Various objects rely on an accurate reconstruction of the PV. In the high luminosity collision runs multiple pp-interactions occur (referred to as pile-up), resulting in multiple PV candidates per event.

PVs are reconstructed as points from which fitted tracks originate. The reconstruction proceeds in two steps: the *primary vertex finding*, in which reconstructed tracks are associated to the vertex candidates, and the *vertex fitting*, in which the vertex position and the corresponding uncertainties are estimated [147].

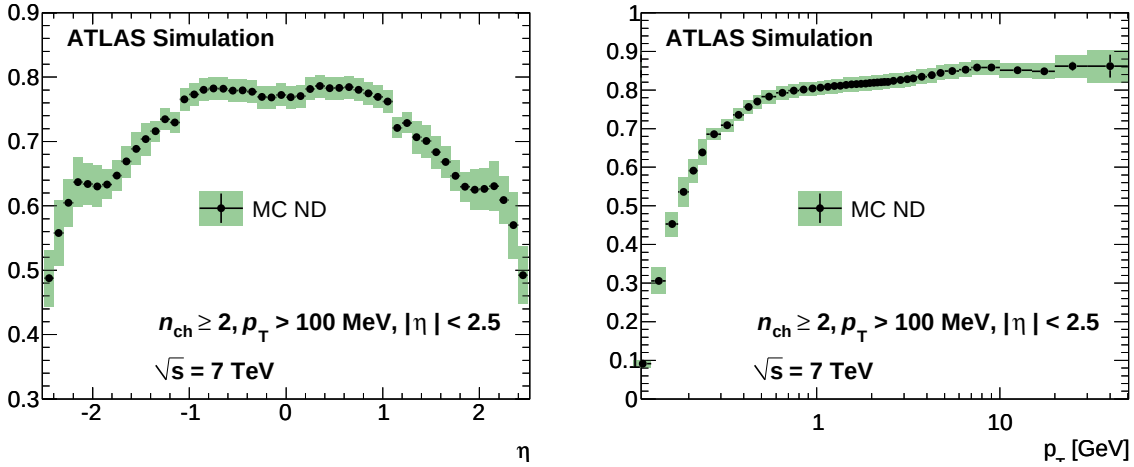


Figure 5.3: The track reconstruction efficiency as a function of pseudo-rapidity (left) and transverse momentum (right) derived from simulated non-diffractive events and measured in pp collision data at $\sqrt{s} = 7$ GeV [146]. Black points stand for data points, the black lines for statistical uncertainties on the measurement and the green shaded areas for the expected systematic uncertainties.

The vertex finding is an iterative procedure that runs over the set of tracks which pass certain quality criteria in terms of a p_{T} threshold, number of hits in the silicon detectors as well as the longitudinal and transverse impact parameters and their resolutions. A vertex seed is established by finding the global maximum of the z_0^{BS} (see the definition of the track parameters in the previous section) distribution of the available tracks. The vertex seed and the tracks surrounding it are used as an input to the *adaptive vertex fitting* procedure [148] in which the vertex position is fitted. Tracks incompatible with the fitted vertex by more than 7 standard deviations (s.d.) are deemed not to match to the vertex, and are returned to the vertex finding for the next iteration. Distributions of PVs with at least 3 associated tracks (with $p_{\mathrm{T}} > 400$ MeV) in the $x - y$ and the $z - x$ planes are shown in Figure 5.4 for ~ 24 pb^{-1} of data events produced at $\sqrt{s} = 7$ TeV.

Among the reconstructed PV candidates, the hard process PV is taken to be that with the highest sum of the squared transverse momenta of the tracks associated to it, $\Sigma_{tracks} p_{\mathrm{T}}^2$.

5.3 Electron Reconstruction

The electron candidates in ATLAS central region ($|\eta| < 2.47$) are reconstructed from energy deposits (clusters) in the electromagnetic (EM) calorimeter, that are associated to a reconstructed Inner Detector track [109, 149]. The reconstruction is refined by a set of requirements on calorimeter, track and combined variables suited to effectively separate electrons from jets. In this section, the electron reconstruction and identification are explained, followed by a description of the electron selection used in the searches presented in this thesis.

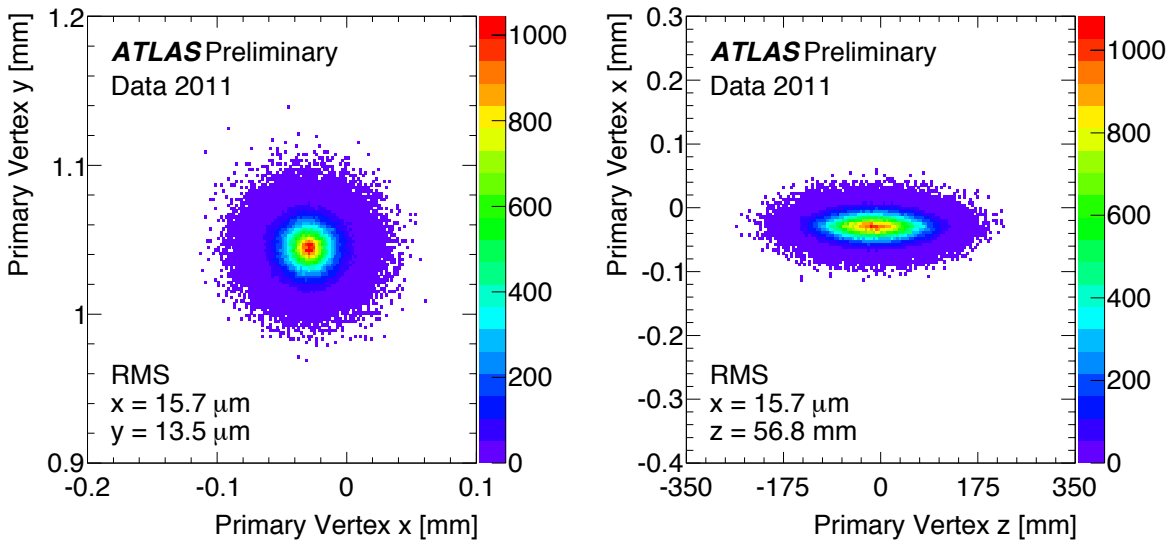


Figure 5.4: Distribution of reconstructed primary vertices in the $x-y$ (left) and $z-x$ (right) planes for ~ 24 pb^{-1} of data events produced at $\sqrt{s} = 7$ TeV. Only vertices are shown that have at least 3 reconstructed tracks are associated.

5.3.1 Electron Reconstruction

Electromagnetic (EM) clusters are reconstructed from *seed clusters* in the calorimeter middle layer. Using the so called *sliding-window* algorithm, seed clusters are established as longitudinal towers consisting of 3×5 calorimeter segments ($\Delta\eta \times \Delta\phi = 0.025 \times 0.025$) with energies above 2.5 GeV. In the central region, electron candidates are formed as seed clusters matched to at least one ID track. Reconstructed tracks are extrapolated from their outermost measurement with respect to the interaction point to determine their impact point into the middle layer of the calorimeter. A track is considered to be matched to a seed cluster if the distance of its impact point from the seed position satisfies $|\Delta\eta| < 0.05$, $\Delta\phi < 0.1$ in the direction of the track bending and $\Delta\phi < 0.05$ in the opposite direction. The $\Delta\phi$ condition is optimised to account for Bremsstrahlung losses. In the case that more than one track fulfils these requirements, that with the minimal space angle $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ between the track impact point and the seed cluster position is considered the best match. Thereby, primary tracks (i.e. those with silicon hits) are given a priority. EM clusters are then recomputed from the electron candidates using an extended sliding window, 3×7 in the barrel and 5×5 in the end caps, requiring the seed cluster to be a subset of the final cluster.

The cluster energy is determined as a sum of the measured energy deposit in the cluster and the estimated energy deposits in the material in front of the EM calorimeter, in the neighbouring EM segments (lateral leakage) and beyond the EM calorimeter (longitudinal leakage). The momentum of the matched track is also used in the calculation. The η and ϕ directions of an electron are included from the corresponding track, unless the track contains only TRT hits, in which case the η is taken from the cluster pointing (η_{cl}) due to the insufficient η resolution of the TRT.

5.3.2 Electron Identification

Electron identification is performed on established electron candidates by applying an offline selection, collectively called the isEM++ menu. The loosest electron definition consisting of a subset of the isEM++ menu is called *loose++* scheme. Applying additional requirements to this scheme, a somewhat tighter *medium++* and yet tighter *tight++* schemes are defined. The complete isEM++ menu is presented in Table 5.1. Some of the requirements shown in this table are optimised in bins of E_{T} and η.

Type	Description	Name
	loose++ selection	
Acceptance	$\lvert\eta\rvert < 2.47$	
Hadronic leakage	Ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster (used over the range $\lvert\eta\rvert < 0.8$ and $\lvert\eta\rvert > 1.37$)	R_{had1}
	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster (used over the range $\lvert\eta\rvert > 0.8$ and $\lvert\eta\rvert < 1.37$)	R_{had}
Middle layer of EM calorimeter	Ratio of the energy in 3×7 cells over the energy in 7×7 cells centred at the electron cluster position	R_η
	Lateral shower width, $\sqrt{(\Sigma E_i \eta_i^2)/(\Sigma E_i) - ((\Sigma E_i \eta_i)/(\Sigma E_i))^2}$ where E_i is the energy and η_i is the pseudo-rapidity of cell i and the sum is calculated within a window of 3×5 cells	$w_{\eta 2}$
Strip layer of EM calorimeter	Shower width, $\sqrt{(\Sigma E_i (i - i_{\mathrm{max}})^2)(\Sigma E_i)}$, where i runs over all strips in a window of $\Delta\eta \times \Delta\phi \approx 0.0625 \times 0.2$, corresponding typically to 20 strips in η, and i_{max} is the index of the highest-energy strip	w_{stot}
	Ratio of the energy difference between the largest and second largest energy deposits in the cluster over the sum of these energies	E_{ratio}
Track quality and track–cluster matching	Number of hits in the pixel detector ($\geqslant 1$)	n_{pix}
	Number of total hits in the pixel and SCT detectors ($\geqslant 7$)	n_{si}
	$\Delta\eta$ between the cluster position in the strip layer and the extrapolated track ($\lvert\Delta\eta\rvert < 0.015$)	$\Delta\eta$
	medium++ selection	
Track quality and track–cluster matching	Number of hits in the b-layer (> 0 for $\lvert\eta\rvert < 2.01$)	n_{BL}
	Number of hits in the pixel detector (> 1 for $\lvert\eta\rvert > 2.01$)	n_{pix}
	Transverse impact parameter ($\lvert d_0\rvert$ <5 mm)	d_0
	Tighter $\Delta\eta$ requirement (< 0.005)	$\Delta\eta$
TRT	Loose requirement on the ratio of the number of high-threshold hits to the total number of hits in the TRT	f_{TR}
	tight++ selection	
Track quality and track–cluster matching	Tighter transverse impact parameter requirement ($\lvert d_0\rvert$ <1 mm)	d_0^{tight}
	Asymmetric requirement on $\Delta\phi$ between the cluster position in the middle layer and the extrapolated track ($\lvert\Delta\phi\rvert < 0.02$)	$\lvert\Delta\phi\rvert$
	Ratio of the cluster energy to the track momentum	E/p
TRT	Total number of hits in the TRT	n_{TRT}
	Ratio of the number of high-threshold hits to the total number of hits in the TRT	f_{TR}
Conversions	Veto electron candidates matched to reconstructed photon conversions	

Table 5.1: Summary of variables used for electron identification [149]. For those for which requirements are E_{T}- and η-dependent no cut values are specified.

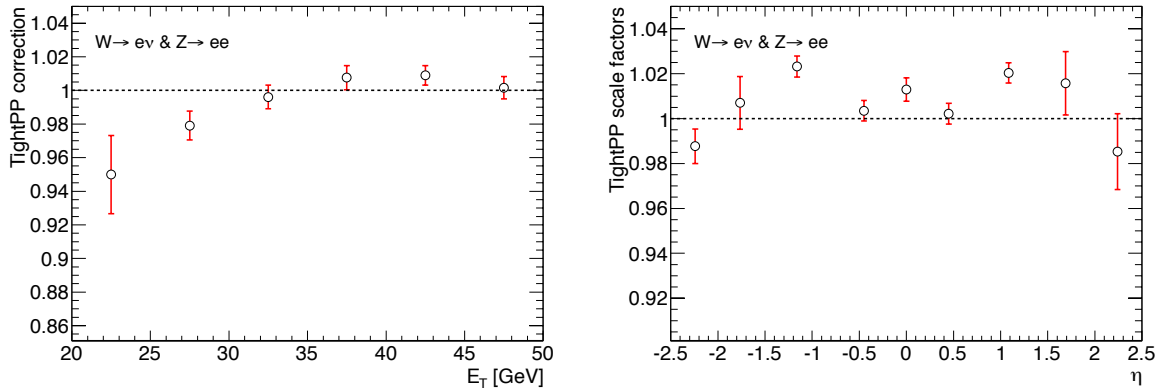


Figure 5.5: Electron *tight++* selection scale factors derived form the Zee and $W \to e\nu$ data samples using the T&P method shown as a function of pseudo-rapidity (left) and transverse energy (right) of the probed electron [142].

5.3.3 Electron Selection

The selected electron candidates are reconstructed in the central region of the detector with $|\eta_{cl}| < 2.47$ excluding the calorimeter transition region $1.37 < |\eta_{cl}| < 1.52$, and to have a transverse energy of $E_{\mathrm{T}} > 25$ GeV defined as $E_{\mathrm{T}} = E_{cl}/\cosh(\eta_{track})$ [150], where the index *cl* stands for the cluster parameters and the index *track* for the parameters of the associated track. Here η_{track} is included in the definition, as the angular coordinates of an electron are adopted from the associated track, as explained in Section 5.3.1. The electron candidates are considered to be identified as electrons if they pass the *tight++* selection [151].

To suppress the QCD multi-jet events with jets misidentified as electrons, tight isolation requirements are imposed on the total transverse energy and momentum within the corresponding cones around the electron direction. The total transverse energy E_{T} is computed using the energy deposits in the calorimeter cells within a cone of $\Delta R = 0.2$ around the direction of the electron candidate, subtracting the transverse energy of the electron candidate itself. It is corrected for the energy leakage into the isolation cone, and for the energy deposit coming from pile-up. The total transverse momentum p_{T} is calculated as a sum of the transverse momenta of all tracks other than that corresponding to the electron candidate, which have a direction within a cone of $\Delta R = 0.3$ around the electron candidate, a transverse momentum of $p_{\mathrm{T}} > 1$ GeV, and are compatible with the PV. In both cases the isolation cuts are chosen corresponding to approximately 90% efficiency according to the tag-and-probe (T&P) method with respect to the *tight++* selection [152]. In addition, selected electrons are required to be geometrically matched with the electron that triggered the event within a cone of $\Delta R = 0.15$. Finally, electrons are required not to overlap with jets. qAll jets within $\Delta R = 0.2$ around an electron are removed from the event. In case any further jet is present within $\Delta R = 0.4$ around the electron after this jet-electron overlap removal, the electron is discarded.

The electron trigger and reconstruction efficiencies are measured on the $Z \to ee$ events using the *Tag-and-probe* method (T&P) [149]. The differences in the trigger, reconstruction and identification efficiencies between data and MC are corrected by applying separate scale factors (SFs) dependent on η and E_{T} and defined as $\varepsilon_{data}/\varepsilon_{MC}$. All SFs are measured from the T&P method. While the trigger and reconstruction SFs are derived from the $Z \to ee$ samples, for the *tight++* selection SFs both $Z \to ee$ and $W \to e\nu$ samples are used. The latter SFs are shown

as a function of η and E_{T} in Figure 5.5. The dominant source of systematic uncertainties to these measurements comes from the background subtraction in data for these two data samples. The isolation cut efficiencies with respect to the *tight++* selection are measured on the $Z \to ee$ sample as well in bins of E_{T} and η.

5.4 Muon Reconstruction

ATLAS follows four different strategies of muon reconstruction, making use of information from different subsystems [153]. The muon candidates used for this analysis are *combined muons*, reconstructed using both MS and ID information. They are seeded by the *stand-alone muons*, reconstructed solely from the MS hits. The momentum of stand-alone muons is corrected for the energy losses in the calorimeter, and their tracks are extrapolated to the beam line to compute the momentum, direction of flight and the impact parameters with respect to the IP. Combined muons are reconstructed by the `Muid` algorithm combining stand-alone muons with matching ID tracks. The algorithm combines the space points assigned to both objects with a common refits for the momentum and the trajectory measurement. In addition, a correction accounting for the muon energy loss in the calorimeters is applied, assuming that muons interact with the material as minimum-ionising-particles.

The muon candidates used for this analysis need to pass tight quality criteria prescribed by the Muon Combined Performance (MCP) group [154]. Selected muons need to have a direction within the detector acceptance $|\eta| < 2.5$ and a transverse momentum $p_{\mathrm{T}} > 20$ GeV in order to be on the plateau of the single muon trigger efficiency. The ID track contained in the muon needs to satisfy all track quality cuts defined by the MCP group:

- Number of b-layer hits > 0, unless the track is passing through a dead b-layer sensor,
- Number of pixel hits + number of crossed dead pixel sensors > 1,
- Number of SCT hits + number of crossed dead SCT sensors $\geqslant 6$
- Number of pixel holes + number of SCT holes < 2,
- For $|\eta| > 1.9$: number of TRT hits (n_{hits}) + number of TRT outliers (n_{outliers}) > 5 and $n_{\mathrm{outliers}}/(n_{\mathrm{hits}} + n_{\mathrm{outliers}}) < 0.9$
- For $|\eta| < 1.9$: if number of TRT hits (n_{hits}) + number of TRT outliers (n_{outliers}) > 5, $n_{\mathrm{outliers}}/(n_{\mathrm{hits}} + n_{\mathrm{outliers}}) < 0.9$ is required

The muon candidates are require to fulfil two isolation requirements. The transverse energy within a cone with $\Delta R = 0.2$ around the muon candidate, from which the energy of the muon track is subtracted, is required to be $E_{\mathrm{T}}^{0.2} < 4$ GeV. Hereby, the energy of the muon track is computed assuming the minimum-ionising interaction of the muon with the detector material. Similarly, the transverse momentum within a cone with $\Delta R = 0.3$ to be $p_{\mathrm{T}}^{0.3} < 2.5$ GeV. The E_{T} within the cone is computed using the calorimeter cells, while the p_{T} is calculated using all tracks except that associated to the muon. These isolation cuts have been specially optimised to suppress the background coming from heavy flavour decays and to reduce the contamination coming from the pile-up present in the second half of the 2011 data taking period. For avoiding

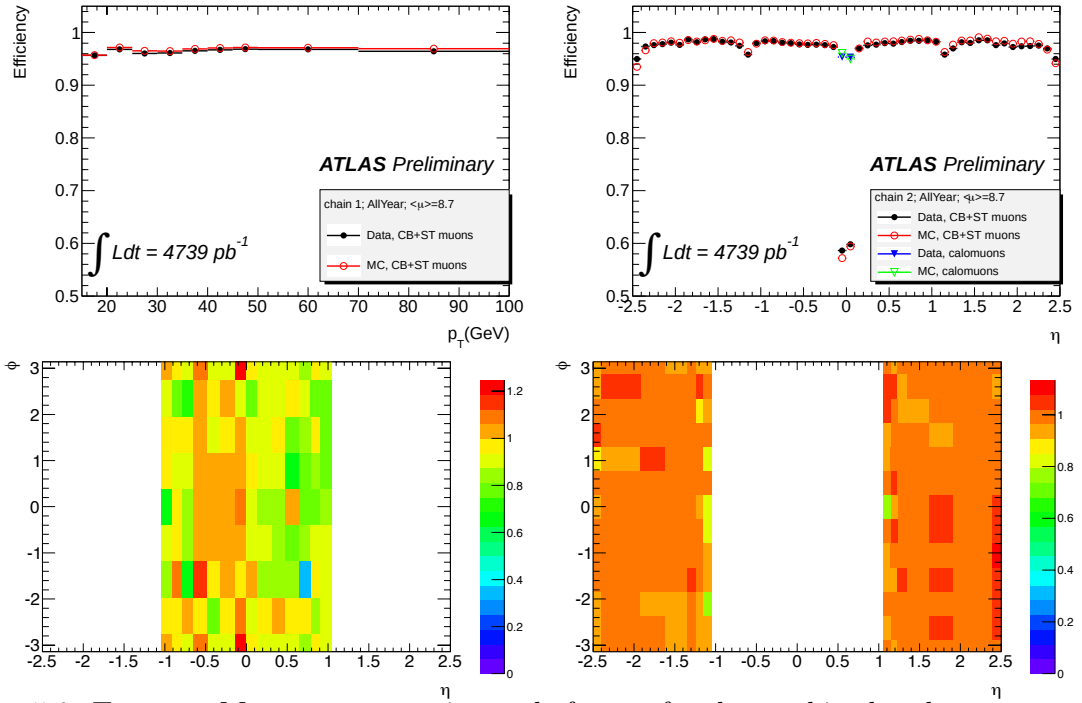


Figure 5.6: Top row: Muon reconstruction scale factors for the combined and segment tagged muons for the produced from the $Z \to \mu\mu$ events collected in the period B-M of the 2011 data taking campaign [155]. Bottom row: Muon trigger scale factors for the barrel (left) and the endcaps (right) regions corresponding to the mu18 trigger derived form the Z data samples using the T&P method. The SFs plotted in bins of η and ϕ are shown for L-M (bottom) [142].

muon-jet double counting, a muon-jet overlap removal is performed. Any muon fulfilling all mentioned criteria is rejected if it is closer than $\Delta R = 0.4$ to any jet with $p_{\mathrm{T}} > 25$ GeV and $|JVF| > 0.75$ (see next section for the explanation of the JVF variable). In addition, selected muons need to match the muon that triggered the event within a cone of $\Delta R = 0.15$.

The muon trigger, reconstruction and isolation efficiencies are corrected in the MC to match those in data by applying SFs measured by the T&P method in $Z \to \mu\mu$ data and MC for the periods B-I, J-K and L-M separately [154]. Summary plots of the muon reconstruction efficiency as a function of p_{T} and η for all data used (periods B-M) are shown in Figure 5.6, top row. Muon trigger SFs for the last two data taking perionds (L-M) are shown in the $\eta - \phi$ plane in Figure 5.6, bottom row.

5.5 Jet Reconstruction

The hadronisation of a quark or a gluon in a collision event produces collimated hadrons resulting from quark and gluon fragmentation, that are known as known as jets. In collisions involving hadrons, such as pp collisions at the LHC, jets are predominant elements of each event. Although jets originate from different partons, their development in the detector is so disperse that jets often overlap. It is thus a non-trivial question how to define jets on the level of reconstructed objects and there are various approaches to its solution.

5.5.1 Jet Definition

The jet algorithm used in the analyses presented in this thesis is the anti-$k_\perp$ clustering algorithm [156, 157]. In clustering algorithms, a distance measure between objects and a criterion for terminating the clustering procedure need to be defined. The distance between objects i and j is defined as:

$$d_{ij} = \min(k_{Ti}^{2p}, k_{Tj}^{2p}) \frac{(\Delta R_{ij})^2}{R^2}, \tag{5.2}$$

with k_{Ti} being the transverse momentum of the object i, $(\Delta R_{ij})^2$ the space angle in quadrature between objects i and j, and R the parameter of the algorithm. The distance between the object i and the origin of the event, which is approximated by the beam spot position called here B, is given by:

$$d_{iB} = k_{Ti}^{2p}. \tag{5.3}$$

A list of d_{ij} is made with all available objects. The objects i and j providing the minimal d_{ij} are merged and the list is recomputed with the new cluster. The procedure is then repeated until no single object is available.

The anti-$k_\perp$ algorithm corresponds to the case where $p = -1$ in Equations 5.2 and 5.3. While guarding infrared and collinear safety, this configuration ensures that clustering is initiated by high transverse momentum objects and that the soft objects preferentially recombine with a high p_T object rather than with each other, which provides rather conical jets. This is contrary to the $k_\perp$ algorithm, in which $p = 1$, resulting in very irregular jet shapes and recombination driven by adjacent soft objects.

Jet candidates used in this analysis are reconstructed using the anti-$k_\perp$ algorithm with $R = 0.4$ on the topological calorimeter clusters calibrated at the electromagnetic (EM) scale. Topological clusters are three-dimensional clusters of calorimeter cells with energies that are significant compared to the expected noise. The clusters are formed seeded by a single cell with the energy above a given threshold and subsequently including the neighbouring cells of sufficient energy. The EM scale is the basic calibration of the energy deposits left by electromagnetic showers in both EMCal and HCal derived from the test beam measurements for electron deposits. The energy scale of the energy deposits left by the electromagnetic showers is then corrected using the invariant di-electron mass around the Z pole in the EMCal, and validated using muons from the test beam and cosmic rays in the HCal.

5.5.2 Jet Energy Calibration

The measured energy of jets reconstructed as described in the previous section suffers losses in the hadronic shower energy due to the non-compensated invisible energy (originating from spallation or excitations of nuclei) or energy escaping the calorimeters (carried by neutrinos), as well as the interaction with the insensitive material in the detector. In order to best estimate the energy of the initial parton from which the jet has developed, these losses are corrected for by the jet calibration to the jet energy scale (JES). The calibration chain is illustrated in Figure 5.7. Jets are first corrected for both in-time and out-of-time pile-up in bins of jet pseudo-rapidity by applying a scheme parametrised according to: the number of PVs in an event, and the average number of interactions in a luminosity block. Then, the jet direction is corrected to point to the PV, and finally calibrated to the JES by applying the simulation-based p_T and η dependent correction factors following the co-called EM+JES scheme [158, 159].

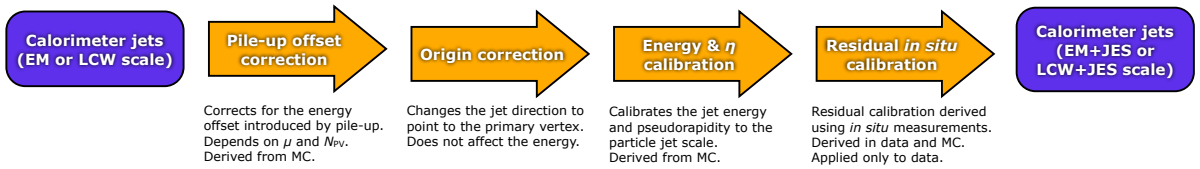


Figure 5.7: Overview of the ATLAS jet calibration scheme used for the 2011 dataset. The pile-up, absolute JES and the residual in-situ corrections calibrate the scale of the jet, while the origin and the eta corrections affect the direction of the jet. Taken from [158].

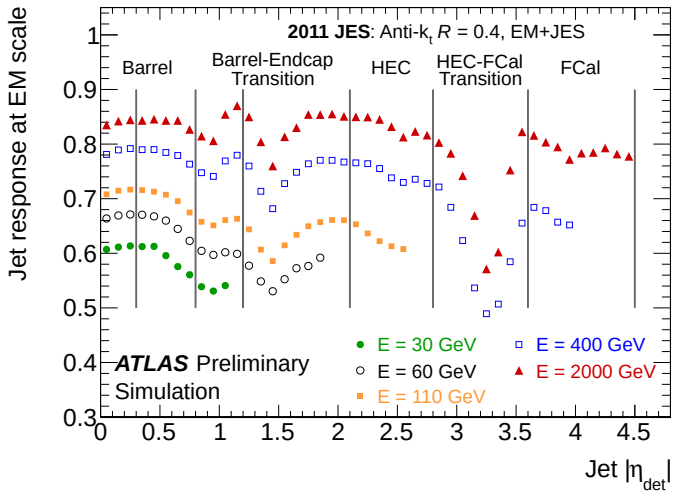


Figure 5.8: Average energy of jets calibrated at the EM+JES scale with respect to the truth jet energy ($E_{\text{jet}}(\text{EM})/E_{\text{jet}}(\text{truth})$ as a function of the jet pseudo-rapidity before applying the correction for the event vertex shown separately for various truth jet energies. The result is shown for MC simulation based on PYTHIA inclusive jet sample for different calorimeter regions. The inverse of the response shown in each bin is equal to the average jet energy scale correction. Taken from [158].

These calibration factors are derived from the inclusive QCD jet events sample simulated using PYTHIA (Figure 5.8). They are calculated as the ratio $\mathcal{R}^{\text{jet}}_{\text{EM}} = E^{\text{jet}}_{\text{EM}}/E^{\text{jet}}_{\text{truth}}$ (the so called average EM scale jet energy response) of the energy of a jet calibrated at the EM+JES scale $E^{\text{jet}}_{\text{EM}}$ and the simulated energy $E^{\text{jet}}_{\text{truth}}$ of the generator level jet matched to it within a cone or $\Delta R = 0.3$. In addition, a correction based on the data to MC comparison of in-situ balance techniques ranging $2\% - 5\%$ is applied to energy scale of jets in data.

5.5.3 Jet Cleaning

Calibrated jets need to satisfy a set of jet quality criteria accounting for hardware problems in the calorimeter, non-collision LHC interactions (beam-gas and beam-halo), as well as cosmic-ray induced showers [159]. Jets suffering such problems are identified by a set of criteria included in the *Loose* jet selection summarised in Table 5.2 and explained below. In addition to improving the quality of the reconstructed jets, this procedure also contributes to the accuracy of the $E^{\text{miss}}_{\text{T}}$ reconstruction (see Section 5.6). In the analyses described in this thesis, jets identified as *Loose* are rejected from the event.

In case of *sporadic noise bursts* in the hadronic end cap calorimeter (HEC), almost the entire jet energy is deposited in a single noisy calorimeter cell. Such jets can be efficiently rejected by filtering out jets with a high fraction of energy in the HEC (f_{HEC}) and in cells with with poor signal shape quality ($f_{\text{HECquality}}$). Such jets are also characterised by apparent negative energy (E_{neg}) in neighbouring cells, due to the capacitive coupling between them.

Type	Loose selection
HEC spikes	($f_{\text{HEC}} > 0.5$ and $\lvert f_{\text{HECquality}}\rvert > 0.5$) or $\lvert E_{\text{neg}}\rvert > 60$ GeV
Coherent EM noise	$f_{\text{EM}} > 0.95$ and $f_{\text{quality}} > 0.8$ and $\lvert\eta\rvert < 2.8$
Non-collision background	$\lvert t_{\text{jet}}\rvert > 25$ ns or ($f_{\text{EM}} < 0.05$ and $f_{\text{ch}} < 0.05$ and $\lvert\eta\rvert < 2$) or ($f_{\text{EM}} < 0.05$ and $\lvert\eta\rvert \geqslant 2$) or ($f_{\text{max}} > 0.99$ and $\lvert\eta\rvert < 2$)

Table 5.2: Summary of variables used for electron identification [159]. For those for which requirements are E_{T}- and η-dependent no cut values are specified.

Similarly, fake jets can also arise from *rare coherent noise* signal in the electromagnetic calorimeter. Such jets are distinguished by a large fraction of electromagnetic energy (f_{EM}) and contribution of calorimeter cells with poor signal shape quality (f_{quality}).

Fake jets originating from cosmic rays or non-collision backgrounds can be suppressed by making use of their time (t_{jet}) calculated with respect to the event time recorded by the trigger, as such jets are normally not synchronised in time with the hard process. Such jets are further characterised by a low energy deposit close to the interaction point, low fraction of the track p_{T} relative to the jet p_{T} (f_{ch}) and a high fraction of energy deposit in a single calorimeter layer (f_{max}).

5.5.4 Pile-up suppression

Collision data collection at a high instantaneous luminosity and short bunch spacing time often suffers in-time (see Section 3.1.1 for an explanation). The presence of pile-up adds soft energy contribution to the event which can alter the calibration of the objects, but also create additional jets in the event.

The rejection of jets originating from pile-up relies on the so-called jet-vertex fraction (JVF) variable, which corresponds to a probability of the jet to not originate from pile-up interactions. This variable is constructed by inspecting the ID tracks associated the the jet, and determining the fraction of the p_{T} contribution coming from those tracks among them which arise from the PV of the event:

$$JVF = \sum_i p_{\text{T}}^i / \sum_j p_{\text{T}}^j. \tag{5.4}$$

Here, the index i runs over the tracks associated to the jets originating from the identified hard-scatter PV, and the index j runs over all tracks associated to the jet. Only jets with $\lvert\eta\rvert < 2.47$ and $20 < p_{\text{T}} < 50$ GeV are considered. Jets outside the fiducial tracking region, and jets without matched tracks are assigned a $JVF = -1$. The optimal working point for the analyses involving $t\bar{t}$-like event signatures using data from 2011 is found to be $\lvert JVF\rvert > 0.75$. Scale factors accounting for differences in pile-up rejection between MC and data are derived from $Z \to ee$ and $Z \to \mu\mu$ data using the T&P method.

5.5.5 Reconstruction Efficiency and Energy Resolution

The calorimeter jet reconstruction efficiency is defined relative to the jets formed from the ID tracks and measured using the T&P method [160]. To account for a disagreement in the efficiency between data and MC, a fraction of jets chosen randomly is discarded in the MC.

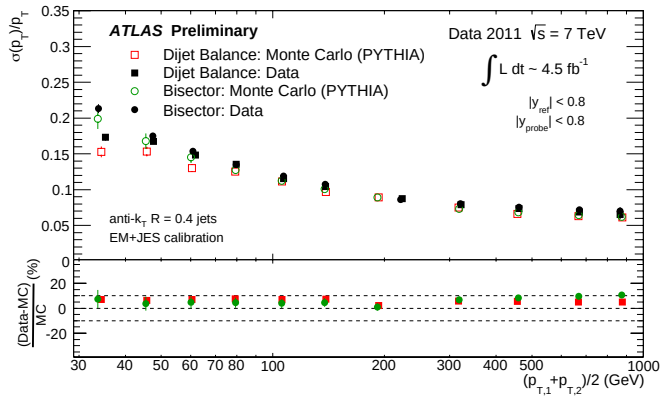


Figure 5.9: Fractional jet energy resolution as a function of the average jet transverse momenta measured with the di-jet balance (squares) and bi-sector (circles) in-situ techniques using the EM+JES calibration in anti-$k_\perp$ R=0.4 cluster jets. The bottom plot shows the relative difference between data results (black) and MC simulation for each method. The dotted lines indicate a relative difference of ±10%. Only statistical errors are shown. Taken from [160].

The jet energy resolution (JER) is measured using the di-jet balance and the bi-sector techniques and is found to agree between data and MC within uncertainties (Figure 5.9) [160].

5.5.6 Jet Selection

Jets used in the analyses presented in this thesis are reconstructed with the anti-k⊥ algorithm with the cone size of $R = 0.4$ and calibrated at the EM+JES scale, including the primary vertex correction. Each jet is required to have a corrected $p_{\mathrm{T}} > 25$ GeV and the be reconstructed within $|\eta| < 2.5$. In addition, in the VLT search the pile-up suppression is performed by requiring $|JVF| > 0.75$.

5.6 Missing Transverse Energy Reconstruction

The $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation used in the analyses presented in this thesis is `MET_RefFinal_em_tightpp`, based on the Top Group object reconstruction definitions [161]. The calculation includes the calorimeter cells calibrated according to the high-p_{T} objects they are associated to, such as electrons, jets, soft jets, and muons. Such calibration of the cells improves the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution. The remaining energy from cells not associated with high-p_{T} object is included as the so called Cell Out term, and is calibrated to the EM scale. The cells around a muon are removed, as they are already included in the MIP energy correction to the muon p_{T} (see Section 5.4). The $E_{\mathrm{T}}^{\mathrm{miss}}$ is then calculated as:

$$E_{\mathrm{T}}^{\mathrm{miss}} = \sqrt{\left(E_x^{\mathrm{miss}}\right)^2 + \left(E_y^{\mathrm{miss}}\right)^2}, \tag{5.5}$$

with:

$$-E_{x,y}^{\mathrm{miss}} = \sum E_{x,y}^{e} + \sum E_{x,y}^{\mu} + \sum E_{x,y}^{\mathrm{jet}} + \sum E_{x,y}^{\mathrm{soft\,jet}} + \sum E_{x,y}^{\mathrm{CellOut}}. \tag{5.6}$$

The electron term $\sum E_{x,y}^{e}$ in Equation 5.6 uses the energy of the electrons which pass the *tight++* selection described in Section 5.3 and satisfy $p_{\mathrm{T}} > 10$ GeV, including all correction factors except for the out-of-cluster correction. The muon term $\sum E_{x,y}^{\mu}$ is derived from the p_{T}

of `Muid` muons for the full MS acceptance ($|\eta| < 2.7$), and all combined muons within $|\eta| < 2.5$. In addition, the non-isolated muons and the muons with the track isolated from all anti-k$\perp$ jets with $R = 0.4$ within a cone of $\Delta R = 0.3$ are included as well. Jets entering the calculation are those reconstructed as described in Section 5.5, divided in two categories: jets with $p_{\mathrm{T}} > 20$ GeV at the EM+JES energy scale ($\sum E^{\mathrm{jet}}_{x,y}$), and soft jets with 7 GeV $< p_{\mathrm{T}} < 20$ GeV at the EM scale ($\sum E^{\mathrm{soft\,jet}}_{x,y}$).

The $E^{\mathrm{miss}}_{\mathrm{T}}$ resolution has been evaluated in MC for semi-leptonic $t\bar{t}$ events as a difference between the reconstructed value and that originating from truth level non-interacting particles (NonInt). The resolutions on $E^{\mathrm{miss}}_{\mathrm{x}}$ and $E^{\mathrm{miss}}_{\mathrm{y}}$ are found to be equivalent and are estimated from a Gaussian fit to the combined residual distribution $E^{\mathrm{miss}}_{\mathrm{x/y}} - E^{\mathrm{NonInt}}_{\mathrm{x/y}}$.

5.7 b-Tagging

Hadronisation of b-quarks results in the production of B-hadrons, i.e. mesons or baryons containing one or more b-quarks. These hadrons can be distinguished from those containing c-quarks or light quarks by exploiting their mass, long life time, and the high branching ratio for the decays containing leptons. Thus, the identification of B-hadrons, the so-called b-tagging, can be employed to accurately recognise events of the processes containing b-quarks in the final state. In particular, b-tagging strongly enforces separation of $T\bar{T}$ and $t\bar{t}$ events, where two b-jets are present, from background processes with no expected b-jets, such as W/Z+light jets, di-boson and multi-jet production. In this section, algorithms for b-jet reconstruction used in the VLT search (Chapter 9) are briefly explained.

5.7.1 b-Tagging Algorithms

The existing b-tagging algorithms exploit the lifetimes and the kinematics of the decay products of B-hadrons. Typical mean lifetimes of B-hadrons are $\tau \sim 10^{-12}$ s, corresponding to a flight length of $c\tau \sim 0.5$ mm before their decay. This leads to the presence of a characteristic pronounced secondary vertex (SV) within a b-jet, that is a jet resulting from a b-quark (Figures 5.10, 5.11). However, the presence of a SV is not sufficient to exclude the possibility that the jet is initiated by a charmed hadron, that is a hadron containing one c- or $\bar{c}$-quark. The mean lifetimes of charmed hadrons are by a factor ~ 2 shorter compared to those of B-hadrons, so that their decay vertex can occasionally be reconstructed as a displaced SV, similar to those in b-initiated jets. Nevertheless, being more than twice heavier, B-hadrons produce decay products with higher transverse momenta, so that charmed hadrons can be distinguished from B-hadrons by further exploiting the p_{T} spectrum of the corresponding decay products. In the ATLAS collaboration, three basic b-tagging algorithms are used, which are discussed in the following. More details about these algorithms can be found in [162, 163].

Impact parameter-based algorithms One class of b-tagging algorithms relies on the transverse and longitudinal impact parameters, d_0 and z_0 respectively, of the tracks associated to jets. As discussed above, the origin of a b-jet, i.e. the intersection point of the tracks matched to it, corresponds to a decay point of a B-hadron, which has a displacement of a few mm with respect to the PV of the hard process. Ideally, both the PV and the b-jet origin lie on the

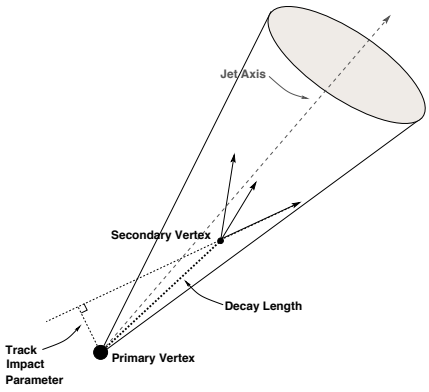


Figure 5.10: A schematic representation of a *b*-jet showing the secondary vertex within the jet, displaced from the primary vertex of the event by the decay length of the *B*-hadron produced from the hadronisation of the *b*-quark. In case the jet and the secondary vertex lie on the opposite sides of the IP, the track impact parameter and the decay length are considered negative.

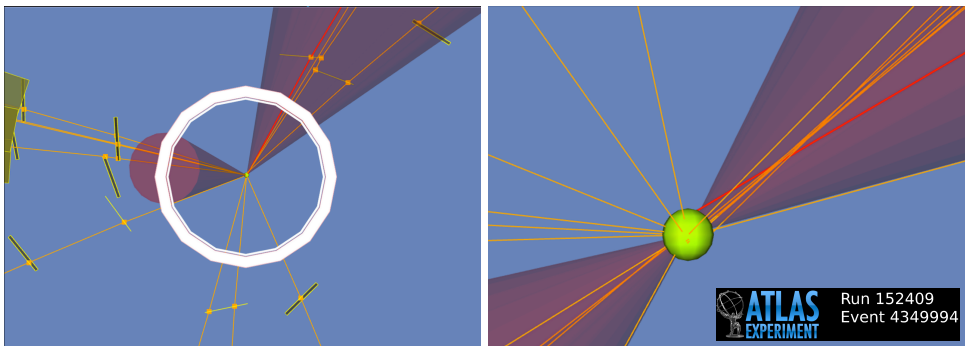


Figure 5.11: An event display of an ATLAS event with one *b*-tagged jet (to the left of the interaction point) magnified to show the Inner Detector region (left) and the secondary vertex displaced with respect to the primary vertex (right). The event display shows a jet with six *b*-tagging quality tracks (orange lines) including one muon (red line.)

direction defined by the jet axis. The IP3D algorithm makes use of this displacement expressed in terms of the impact parameter significances d_0/σ_{d_0} and z_0/σ_{z_0}. These terms are also signed: in case the intersection of the extrapolation of a given track and the jet direction lays at the same side of the PV as the jet, the IPs are signed as positive (which is a more likely signing for b/c-jets than for the light jets), and otherwise as negative. The IP3D algorithm computes a likelihood ratio based on smoothed and normalised two dimensional distributions of d_0/σ_{d_0} and z_0/σ_{z_0} produced from MC for *b*- and light jets separately to express the level of compatibility of a given jet with the *b*- and light jet hypotheses.

Secondary vertex-based algorithms Another class of *b*-tagging algorithms relies on the reconstruction of the inclusive vertex formed by the decay products of a *B*-hadron and eventual subsequent charmed hadron. The SV1 algorithm searches for all pairs of tracks which are associated to a jet, sufficiently far from the PV, and form a good vertex. Vertices compatible with V^0 decays[2] or material interaction are discarded. Tracks from all good vertices reconstructed in this way are combined into a single inclusive vertex (IV). In an iterative procedure, tracks are removed one by one until the best χ^2 of the IV is achieved. The fitted 3-dimensional position

[2] V^0 decays refer to decays with a "V" shape signature, i.e. decays of neutral particles into a pair of charged particles, such as $K^0_s \to \pi^+\pi^-$ or $\Lambda \to p\pi^-$.

of the IV is used to reconstruct the discriminating variables, such as the signed decay length significance $L_{3D}/\sigma_{L_{3D}}$ (with respect to the PV), the invariant mass of all tracks associated to IV, the ratio of total energy of the tracks associated to IV to the total energy of all tracks associated to the jet, and the space angle ΔR between the jet axis and the direction defined by the PV and the IV. These IV-specific variables are combined with the number of good two-track vertices using the likelihood ratio technique to compute the b-tagging jet weights.

Decay chain reconstruction based The JetFitter algorithm is based on exploiting the topology of weak decays of b- and c-quarks inside a jet. It uses a Kalman filter to reconstruct the flight of the b-hadron and subsequent c-hadron by fitting a track connecting the PV, the b- and the c-decay vertices. This fit allows for reconstruction of the variables used in the SV1 tagger. In the JetFitter, these variables, along with the flight length significances of the b- and c- vertices, are used in a likelihood method, to discriminate between b-, c- and light jets.

Combined multi-variate based Several b-tagging algorithms combine the outputs of the simpler algorithms to produce a multi-variate discriminator with a higher performance. The JetFitterCombNN algorithm uses the outputs of the IP3D and JetFitter algorithms to produce a discriminator based on an artificial neural network training. The JetFitterCombNNc algorithm is a similar algorithm specialised for distinguishing c-jets from b-jets. In the analysis described in Chapter 9, b-tagging is performed using the neural network based MV1-tagger [163] which combines the per-jet tagging weights of the IP3D, SV1 and JetFitterCombNN algorithms as well as p_{T} and η of jets, into a final b-tagging weight with a stronger discriminating performance compared to the individual methods (Figure 5.13). In these analyses the 70% efficiency working point is used, which in $t\bar{t}$ events ensures that 70% of b-jets are tagged, while the c- and light jets with $p_{\mathrm{T}} > 20\ GeV$ are suppressed by a factor of 5 and 137 respectively [164]. The power of a b-tagging algorithm can be demonstrated in terms of the rejection of non-b jets as a function of the tagging efficiency for b-jets. Such power curves for light and c-jet are shown in Figure 5.13 for various b-tagging algorithms, derived from simulated $t\bar{t}$ events using jets with $p_{\mathrm{T}}^{\mathrm{jet}} > 15$ GeV and $|\eta^{\mathrm{jet}}| < 2.5$. The plots in this figure show that among the general purpose tagging algorithms (that is regardless the specialised JetFitterCombNNc), the MV1 algorithm has the best rejection of light- and c-jets for the b-jet tagging efficiency of 70%.

5.7.2 b-Tagging Calibration

The efficiency with which jets originating from b-quarks are tagged as b-jets, as well as the rates with which jets originating from c- or a light quark are mistakenly tagged as b-jets (the so-called mis-tag rates), are estimated in the simulated $t\bar{t}$ events in bins of p_{T} and η, and measured in data events enriched in b-, c-, and light jets respectively in bins of p_{T}. Differences between data and simulation in the tagging efficiency and mis-tag rates are corrected in the simulation using the corresponding p_{T}-dependent SFs defined as $SF(p_{\mathrm{T}}, f) = \epsilon^{f}_{data}/\epsilon^{f}_{MC}$, for each flavour f individually (Figure 5.13).

The measurement of the b-tagging efficiency in data and the derivation of the calibration SFs is obtained as a combination of two methods, p_{T}^{rel} and System8, using the di-jet events in which one of the jets contains a muon resulting from a semi-leptonic decay of a heavy quark (b or c) [165]. The SFs as a function of the p_{T} of the tagged jet are shon in Figure 5.14.

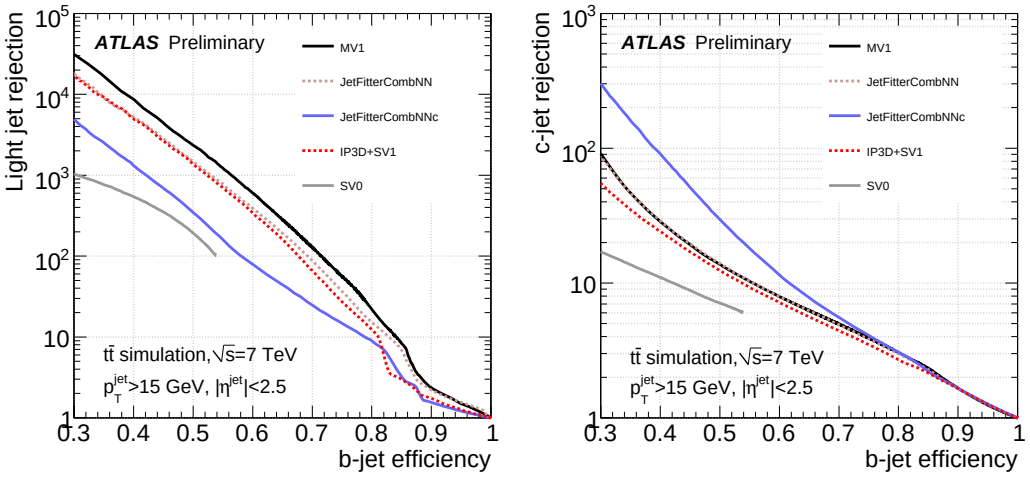


Figure 5.12: Light (left) and c-jet (right) rejection as a function of the b-jet efficiency for several tagging algorithms: MV1 (black line), JetFitterCombNN (grey dotted line), JetFitterCombNNc (blue line), IP3D+SV1 (red dotted line) and SV0 (grey line). The curves are derived from simulated $t\bar{t}$ events using jets with $p_{\mathrm{T}}^{\mathrm{jet}} > 15$ GeV and $|\eta^{\mathrm{jet}}| < 2.5$. Taken from [165].

The b-tagging efficiency on jets containing charm hadrons is measured on the $D^{*+} \to D^0(\to K^-\pi^+)\pi^+$ events by comparing the event yields of D^{*+} mesons before and after the b-tagging requirement [166]. The obtained efficiency as well as the corresponding SFs as a function of the p_{T} of the tagged jet are shown in Figure 5.15.

The mis-tag rates are measured using the so called negative tag method, that is by making use of the rate at which the secondary vertex of a jet lies at the opposite side of the PV with respect to the jet, i.e. has a so-called negative decay length [167]. The simulated and measured mis-tag rates as well as the corresponding calibration SFs are shown as a function of the p_{T} of the tagged jet in Figure 5.16.

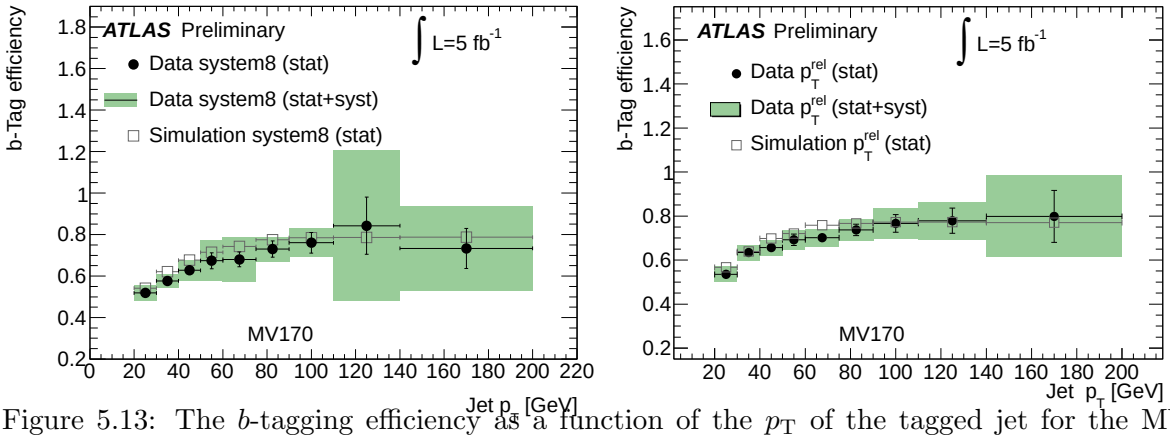


Figure 5.13: The b-tagging efficiency as a function of the p_{T} of the tagged jet for the MV1 algorithm at 70% efficiency, obtained by the System8 (left) and p_{T}^{rel} (right) methods. The data points are shown as black dots with statistical error bars. The simulation is shown as empty squares with green systematic error bands. Taken from [165].

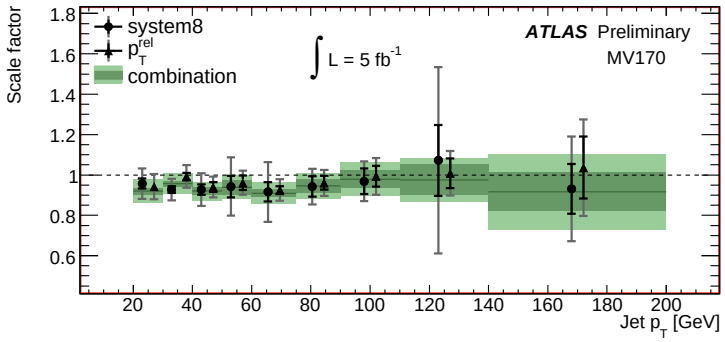


Figure 5.14: Data to simulation correction scale factors of the b=tagging efficiency estimated using the System8 and p_{T}^{rel} methods as well as their combination. The dark green band represents the statistical uncertainty of the combined scale factor while the light green band shows the total uncertainty. The data points showing the p_{T}^{rel} and system8 measurements have been separated a little along the p_{T}-axis to make the plot more readable. Taken from [165].

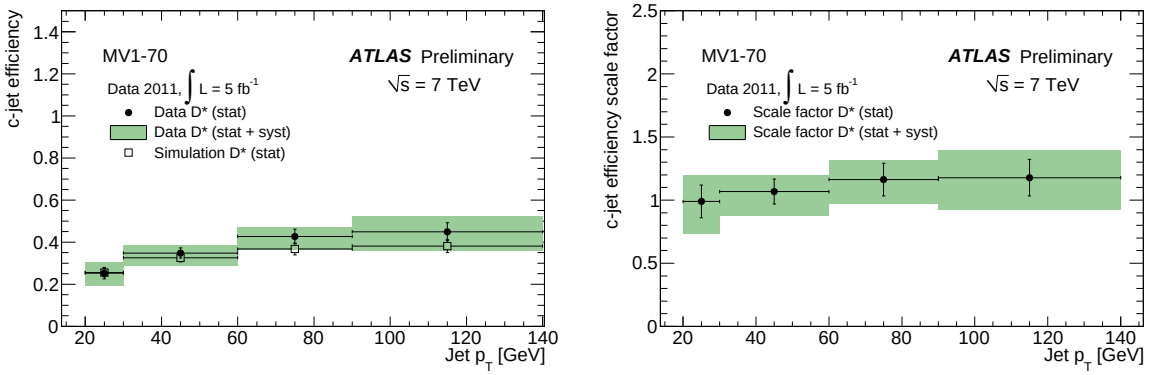


Figure 5.15: The b-tagging efficiency (left) and efficiency scale factors (right) as a function of jet p_{T} for the MV1 tagger at the 70% efficiency measured on the c-jets in $D^{*+} \to D^0(\to K^-\pi^+)\pi^+$ events. The black points show the data with, the error bars representing statistical uncertainties. The green bands indicate the total expected uncertainty. Taken from [166].

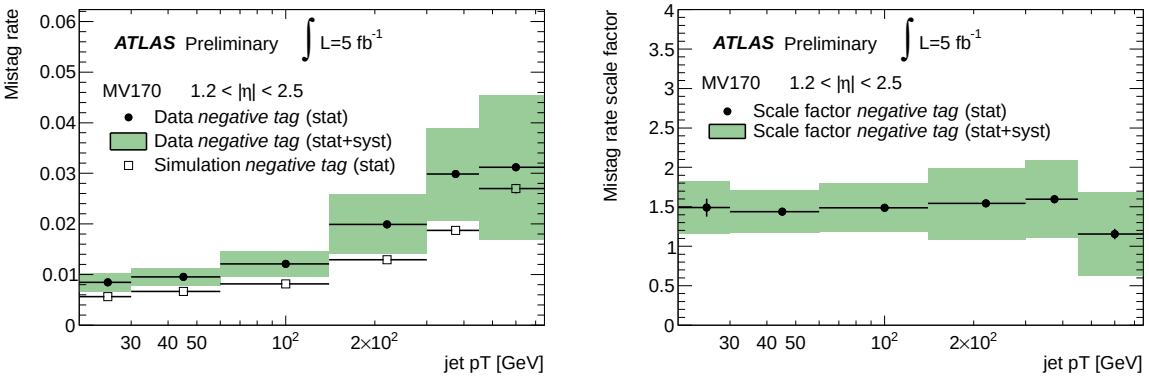


Figure 5.16: The mistag rates (left) and the corresponding SFs (right) as a function of jet p_{T} in data and simulation for MV1 algorithm at 70% measured using the negative tag method for jets with $1.2 < |\eta| < 2.5$. The black points show the data with, the error bars representing statistical uncertainties. The green bands indicate the total expected uncertainty. Taken from [167].

Chapter 6

Analysis Tools

In the epoch of an increasing demand for precision, along with an experimental accessibility to the processes that have decreasingly low production rates, the usage of sophisticated analysis tools for event reconstruction, selection optimisation or statistical hypothesis testing in high energy physics (HEP) data analyses is gradually becoming a necessity. The analysis tools employed in the searches presented in Chapters 8 and 9 of this thesis are presented in this chapter. The toolkit for kinematic likelihood fitting is described in Section 6.1. The framework for multi-variate analysis is presented in Section 6.2. The limit setting utility is explained in Section 6.3.

6.1 Kinematic Likelihood Fitter

Events resulting from the pair-production of top or heavy quarks decaying into a W-boson and a b-quark, in the final states where one W-boson decays leptonically and the other W-boson hadronically, contain four jets at the leading order - two of them originating from b-quarks and the other two from light quarks. At higher orders, further jets resulting from initial or final state radiation can be present, increasing the number of jets in the event. On the other hand, the merging of jets very close in space by a jet clustering algorithm, or the loss of jets due to detector effects, decrease the overall number of jets in the event, sometimes to less than four.

For any analysis in which the reconstruction of four-momenta of the initial quark pair is required, it is necessary to associate the reconstructed objects to one of the two initial quarks. In final states containing a single lepton, the lepton and the missing transverse energy clearly originate from the same quark, i.e. that where the W-boson decays leptonically. In order to fully reconstruct the quark from which this W-boson originates, it is necessary to find out which among the observed jets originates from the same quark, and to make an estimate of the longitudinal component of the neutrino momentum. Moreover, the observed energies and angles of the reconstructed objects result from a smearing of their initial values due to detector resolution, while the reconstructed masses are also subject to additional smearing due to the width of the particle in question. Thus, it is possible to make an estimation of the initial particle properties given their observed values and the best estimation for the particle widths and detector resolutions. One of possible approaches to this problem is kinematic fitting.

6.1.1 KLFitter Framework

Kinematic fitting is a procedure of assigning the reconstructed objects to the hard-scattering partons and finding the best estimate for the initial values of their kinematic properties, by finding an extremum of a given metric. That metric is constructed to express the compatibility of a given assignment to the signature in question, by making use of the kinematics of the underlying process and taking detector resolutions into account. One such algorithm is implemented in the Kinematic Likelihood Fitter (KLFitter) framework, a ROOT based C++ implementation of a kinematic fitting algorithm based on the maximum likelihood method. The KLFitter is fully equipped to reconstruct $t\bar{t}$ events in the lepton+jets channel, but designed as a flexible tool adaptable to other processes as well. The most recent version of the framework (which appeared later than the search using the KLFitter described in this thesis) is documented in [168]. The framework is based on the leading-order process assumption, not considering the presence of ISR or FSR, and considering all decay products of the event, i.e. a lepton, a neutrino and four jets detected. The detector resolutions are assumed to be known and modelled by the transfer functions (see Section 6.1.1.3).

The central ingredient of the framework is the likelihood function $\mathcal{L}$ described in detail below. With four jets in an event, and the event configuration being invariant under permutation of the light jets, 12 jet permutations are possible. For each permutation the likelihood is maximised in the space of its parameters by minimising the $-\log\mathcal{L}$ function. This procedure is performed by the Bayesian Analysis Toolkit (BAT) framework [169], employing three minimisation techniques: Marcov Chain Monte Carlo (MCMC) sampling [170], Simulated Annealing (SA) [171] and Minuit [172]. Global minima of the likelihood function are coarsely located by MCMC and then passed to Minuit for finer minimisation. Minuit thereby searches for the best estimates of the reconstructed heavy quark mass, and the p_z of the neutrino starting from the measured observables of the lepton and the four jets. In some rare cases Minuit fails to find a minimum, in which case the SA method is used. The best parameters found by SA are then passed again to Minuit for a more precise location of the minima. In this way, the best parameters are found for every jet permutation. The permutation which delivers the highest likelihood value is considered the best permutation and the parameters corresponding to it are used in the analysis.

6.1.1.1 Likelihood Function

The reconstruction of the four-momenta of the initial quark and anti-quark of the event relies on the measurement of energy, transverse momentum and angular coordinates of the observable objects resulting from their decay products, in this case a lepton, a neutrino and four jets, two of which need to originate from b-quarks. Thus, the mass reconstruction of initial quarks is subject to energy, momentum and angular detector resolutions, in addition to the natural particle width.

The likelihood function of the KLFitter is constructed as a conjunction of individual likelihood terms describing the kinematics of the decay, the properties of the particles involved as well as the detector energy resolutions. The other resolution effects have been ignored, being small compared with the energy resolution. The form of the likelihood function used in the search

presented in this thesis is given by:

$$\begin{aligned}\mathcal{L} \;=\; & BW\{m_{jj}|m_W,\Gamma_W\}\cdot BW\{m_{l\nu}|m_W,\Gamma_W\}\cdot \\ & BW\{m_{jjb_{\mathrm{had}}}|\tilde{m}_t,\Gamma_t\}\cdot BW\{m_{l\nu b_{\mathrm{lep}}}|\tilde{m}_t,\Gamma_t\}\cdot \\ & W(\tilde{E}_{\mathrm{jet}_1}|E_{b_{\mathrm{had}}})\cdot W(\tilde{E}_{\mathrm{jet}_2}|E_{b_{\mathrm{lep}}})\cdot W(\tilde{E}_{\mathrm{jet}_3}|E_{q_1})\cdot W(\tilde{E}_{\mathrm{jet}_4}|E_{q_2})\cdot \\ & W(\tilde{E}^{\mathrm{miss}}_{\mathrm{x}}|p^{\nu}_{\mathrm{x}})\cdot W(\tilde{E}^{\mathrm{miss}}_{\mathrm{y}}|p^{\nu}_{\mathrm{y}})\cdot \begin{cases} W(\tilde{E}_e|E_e) & \text{if single electron channel,} \\ W(\tilde{p}^{\mu}_{\mathrm{T}}|p^{\mu}_{\mathrm{T}}) & \text{if single muon channel.}\end{cases}\end{aligned} \tag{6.1}$$

As the pair of two light jets as well as the lepton-neutrino pair are assumed to originate from W-bosons, their invariant masses m_{jj} and $m_{l\nu}$ respectively need to be compatible with a Breit-Wigner distribution (called BW in Equation 6.1) around a mass $m_W = 80.4$ GeV within a width $\Gamma_W = 2.1$ GeV, which corresponds to the particle width of the W-boson. These two constraints are expressed by the first two terms of Equation 6.1.

Furthermore, the invariant mass of a reconstructed W-boson and the corresponding b-jet need to be compatible with the mass of the parent heavy quark. In analyses targeted to measurements of kinematic properties of $t\bar{t}$ events, this condition would be expressed as a constraint that the mass of the hadronically decaying top quark $m_{jjb_{\mathrm{had}}}$ and that of the leptonically decaying top quark $m_{l\nu b_{\mathrm{lep}}}$ need to be compatible with a Breit-Wigner distribution around a mass $m_t = 172.5$ GeV within a width $\Gamma_t = 1.5$ GeV. In the search presented in Chapter 8 of this thesis, pair-produced heavy t'-quarks of an unknown mass are searched for. Thus, the latter condition is modified to leave the t'-quark mass $m_{t'}$ as a free parameter, but to constrain the $m_{jjb_{\mathrm{had}}}$ and $m_{l\nu b_{\mathrm{lep}}}$ to be compatible with each other within a Breit-Wigner distribution with a width Γ_t. Strictly speaking, this width is dependent on the mass of the underlying heavy quark. However, the studies presented in Section 8.3 have shown that there are no significant differences in the fitter performance when using the particle width corresponding to a particle mass of 172.5 GeV and to that of 400 GeV, so the Γ_t is used. The latter constraint is expressed by the terms in the second line of Equation 6.1.

The measured energy and momentum resolutions of the objects involved in the fit, accounted for by the terms in the last two lines of Equation 6.1, are described by the transfer functions (called W in Equation 6.1) explained in Section 6.1.1.3.

6.1.1.2 Truth Matching

To derive the transfer functions, and also to study the performance of the Fitter, it is necessary to match the reconstructed jets with the corresponding generator level jets (i.e. the jets formed from generated particles of the particle shower). This match is established following a geometric requirement:

$$\Delta R = \sqrt{(\eta_{\mathrm{reco}} - \eta_{\mathrm{truth}})^2 + (\phi_{\mathrm{reco}} - \phi_{\mathrm{truth}})^2} < 0.3, \tag{6.2}$$

with η_{reco} and ϕ_{reco} being the azimuthal and polar angle of the reconstructed jet, and η_{truth} and ϕ_{truth} the corresponding coordinates of the matched generator level jet. Such matching is not necessarily one-to-one. A reconstructed jet can be matched to more than one generator level jets and vice versa. A reconstructed jet is considered *matched* only if it has a unique matching with a generator level jet. Lepton matching is established in an analogous way, with a difference that a cone size is required to be smaller than 0.1. An event is said to be matched if all four

jets and the charged lepton at the generator level have a unique match with corresponding reconstructed objects. An object is considered *identified* if the category to which it is assigned by the KLFitter is that to which it belongs at the generator level.

6.1.1.3 Derivation of the Transfer Functions

The transfer function (TF) of a quantity x is defined as the conditional probability $W(x_{\text{reco}}|x_{\text{true}})$ of measuring a response x_{reco} in a detector given a true value x_{true}.

In the KLFitter framework, the energy TFs are parametrised by double-Gaussian functions:

$$W(E_{\text{reco}}|E_{\text{true}}) = \frac{1}{\sqrt{2\pi}(p_2 + p_3 p_5)} \left(e^{-\frac{(\Delta E - p_1)^2}{2p_2^2}} + p_3 e^{-\frac{(\Delta E - p_4)^2}{2p_5^2}} \right), \tag{6.3}$$

with ΔE being the relative difference between the true and reconstructed energy:

$$\Delta E = \frac{E_{\text{true}} - E_{\text{reco}}}{E_{\text{true}}}, \tag{6.4}$$

and energies being expressed in GeV. The parameters p_i are functions of E_{true} given by:

$$p_i = a_i + b_i E_{\text{true}},\ i \neq 2, \tag{6.5}$$

$$p_2 = a_2/\sqrt{E_{\text{true}}} + b_2. \tag{6.6}$$

The momentum TFs (used for muons) are derived analogously to Equations 6.3-6.4, with the parameters p_i given by:

$$p_i = a_i + b_i p_{\text{T}}^{\text{true}}. \tag{6.7}$$

Parameters a_i and b_i have been derived from the events passing the standard $t\bar{t}$ selection of the nominal MC@NLO $t\bar{t}$ event sample (Section 7.4.2.2) as well as the corresponding event sample with no spin correlations. The derivation has been performed for each particle separately, in three bins of pseudo-rapidity, motivated by detector resolution and efficiency. Only the matched particles have been used. An initial set of fit parameters p_i has been obtained from fitting ΔE in bins of E_{true}. This set of parameters is then passed to a global fit performed using BAT for a multidimensional fit parameter optimisation via Markov Chain Monte Carlo.

6.2 Artificial Neural Network Method

The Artificial Neural Network (NN) method used in the VLT search (Chapter 9) is introduced in this section. The method explanation shown in this section is strongly inspired by the corresponding description from [173].

Generally speaking, a NN is a set of simulated interconnected neurones each of which has a certain response to a given signal. When an external signal is applied to some of the neurones, i.e the input neurones, it is propagated over the connections (synapses) to the other neurones following a given propagation rule so that the whole network is put into a defined state which can be measured from the response of one or more output neurones. In other words, a neural network is a mapping of a space of input variables $x_1, ..., x_{n_{in}}$ onto a space of output responses $y_1, ..., y_{n_{out}}$. In the case of a selection problem when processed data belongs to one of two categories: signal

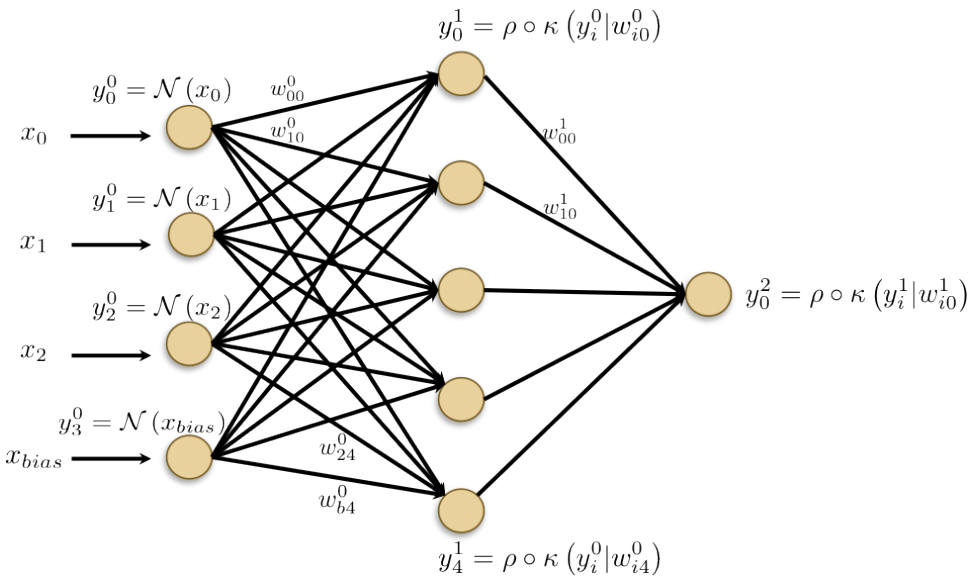


Figure 6.1: A schematic representation of a multi-layer perceptron architecture with a single hidden layer. The input and the hidden layer are shown as sets of neurones (yellow circles) arranged in vertical arrays. Since for the purpose of this analysis only two categories of events (signal and background) are distinguished, the output layer consists of a single neurone. Synapses are represented by black arrows pointing into the direction of the information flow. The synapse weights, the node values and the operations applied to them are indicated by the text in the scheme. Adapted from [173].

or background, a single output neurone is needed with the desired output close to 1 for the signal events and close to 0 for the background events. Among the available neural network implementations, this analysis employs the multi-layer perceptron (MLP) method contained in the TMVA package [173]. The underlying algorithm is extensively described in the following. The architecture of an MLP network and the signal propagation through it are described in Section 6.2.1. The learning procedure is explained in Section 6.2.2.

6.2.1 Network Architecture and Signal Propagation

The MLP method is based on the feed-forward multi-layer network construction, in which the neurones are organised in layers with synapses connecting only the neurones from the neighbouring layers as shown in Figure 6.1, and the signal propagates in only one direction - from the input layer towards the output layer. The input layer contains one neurone per input variable. All layers between the input and the output layer are called hidden layers. The number of hidden layers as well as the number of neurones in them is set by the user. Every layer is assigned an additional neurone introducing an independent bias to the network. The approximation theorem of Weierstrass [174] ensures that for a multilayer perceptron a single hidden layer with a sufficiently large number of neurones is sufficient to approximate a given continuous correlation function to any precision. The same performance is likely to be achieved with a larger number of hidden layers and potentially significantly smaller number of hidden neurones, which would also result in a shorter training time and a more robust network. Nevertheless, in this analysis only networks with a single hidden layer are considered.

The input variables x_i are all normalised to the same range before being provided to the input neurones. For a given event, a neurone from the input layer (denoted as layer 0) obtains the normalised value of the corresponding variable $y_i^0 = N(x_i)$, where N is a $\mathbb{R} \mapsto \mathbb{R}$ function. That signal is then transmitted to the outgoing synapses where it is weighted by the corresponding synapse weight. The set of synapse weights $\mathbf{w}$ is initialised randomly at the beginning of the training and is adjusted during the training procedure as described in 6.2.2.

For a neurone in a hidden layer, the signals carried by all incoming synapses are transformed into a single output by the neurone response function ρ which consists of a synapse function κ and a neurone activation function α. The $\mathbb{R}^n \mapsto \mathbb{R}$ synapse function is the summation rule for the incoming signals. The neurone activation function α is a $\mathbb{R} \mapsto \mathbb{R}$ function applied to the result of κ. For a neurone j in the layer l the response y_j^l equals:

$$y_j^l = \rho \circ \kappa \left(y_{ij}^{l-1} | w_{ij}^{l-1} \right), \tag{6.8}$$

where i loops over all incoming synapses of the neurone j. The available synapse and activation functions are shown in Equations 6.9 and 6.10.

$$\kappa : \left(y_1^l, ..., y_n^l | w_{0j}^l, ..., w_{nj}^l \right) \rightarrow \begin{cases} w_{bj}^l + \sum_{i=0}^n y_i^l w_{ij}^l & \text{Sum,} \\ w_{bj}^l + \sum_{i=0}^n \left(y_i^l w_{ij}^l \right)^2 & \text{Sum of squares,} \\ w_{bj}^l + \sum_{i=0}^n |y_i^l w_{ij}^l| & \text{Sum of absolutes.} \end{cases} \tag{6.9}$$

$$\alpha : x \rightarrow \begin{cases} x & \text{Linear,} \\ \frac{1}{1+e^{-kx}} & \text{Sigmoid,} \\ \frac{e^x - e^{-x}}{e^x + e^{-x}} & \text{Tanh,} \\ e^{-x^2/2} & \text{Radial.} \end{cases} \tag{6.10}$$

In this analysis, sum of squares is used as a summation rule and tanh is chosen for the neurone activation function, following the optimisation study presented below in Section 9.3.7.

6.2.2 Training Procedure

A neural network training represents a procedure of adjusting the synapse weights by minimising a chosen metric. The MLP uses so-called *bulk learning* which means that the training procedure is divided into training cycles in which the metric used is constructed from a number of events (not from individual events). The training methods implemented in the TMVA MLP belong to a family of supervised learning methods in which the desired outcome of the training events is known. During a training cycle c, N training events of the form $\mathbf{x}_e = (x_1, ..., x_{n_{var}})$ are provided to the network. Each event is characterised by the category indicator, $\hat{y}_e$, which equals 1 for signal events and 0 for background events. The network output, $\mathbf{x}_e$, is recorded for all events within one training cycle and compared to the corresponding desired outcome $\hat{y}_e$. The goodness of the network output within one training cycle is quantified in terms of an *error function* $E(\mathbf{x}_1, ..., \mathbf{x}_N | \mathbf{w}^c)$ defined as:

$$E(\mathbf{x}_1, ..., \mathbf{x}_N | \mathbf{w}^c) = \sum_{e=1}^{N} E_e(\mathbf{x}_e | \mathbf{w}^c), \tag{6.11}$$

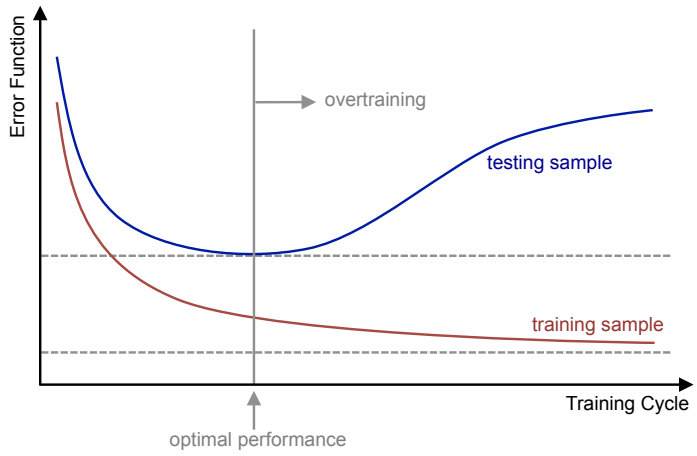


Figure 6.2: Example of the error function development for an over-trained network. The red (blue) line shows the development of the error function evaluated on the training (testing) sample over the course of the training epochs. When training beyond the point of the optimal performance marked by the vertical gray line, the error function continues decreasing for the training sample, while it starts increasing for the testing sample.

with the *estimator function* $E_e\left(\mathbf{x}_e|\mathbf{w}^c\right)$ of a form:

$$E_e\left(\mathbf{x}_e|\mathbf{w}^c\right) = \begin{cases} \frac{1}{2}\left(y_e^{nn} - \hat{y}_e\right)^2 & \text{Mean Square Estimator,} \\ -2\left(\hat{y}_e \log y_e^{nn} + (1-\hat{y}_e)\log(1-y_e^{nn})\right) & \text{Cross Entropy Estimator.} \end{cases} \tag{6.12}$$

At the end of a training cycle c, the current set of the synapse weights $\mathbf{w}^c$ is corrected in the manner which reduces the value of the estimator function. The MLP method offers two approaches to this problem as described below.

The value of the error function needs to be monitored carefully for the training and testing sample simultaneously at the end of each training cycle. At the beginning of the training procedure the most discriminative features of the training events drive the learning process while improving its conclusive power, which manifests itself as a diminution of the error function. During the first cycles the error function value normally makes rather large oscillations which then converge to a stable value as the training progresses. However, if the training continues in too many cycles, the network starts learning from very specific features of the events, like the variables with a weak separation, the bias or the statistical fluctuations of the variables provided rather than their general features, which is misleading. In that case the network learns to describe the events which it is trained on very accurately while at the same time its predictive power deteriorates. This phenomenon is known as *over-training* and is the most common problem of machine learning in general. Over-training can be easily detected from the error function evolution plot as illustrated in Figure 6.2. When the network starts describing the statistical fluctuations of the training sample, the error function as a function of the number of training cycles keeps shrinking for the training sample while it starts raising for the test sample. However, the training procedure should also not be interrupted too early, before reaching its optimal performance, which would then be a case of *under-training*.

The training procedure normally consists of at least 500 training cycles. In a stable training the error function is almost constant over the last ~ 100 training cycles in our training. The training used for the VLT search consists of 650 cycles as explained in Section 9.3.7.

6.2.2.1 Back-propagation

Back-propagation [175] is the most common method for training feed-forward neural networks. The set of the synapse weights $\mathbf{w}^c$ in the training cycle c is adjusted in the direction of the steepest decent of the error function in the $\mathbf{w}$ space:

$$\mathbf{w}^{c+1} = \mathbf{w}^c - \eta \nabla_{\mathbf{w}} E, \tag{6.13}$$

with the *learning rate* η adjustable by the user. The learning rate should be carefully chosen. The back-propagation procedure is in a danger of getting stuck in local minima of the error function if the learning rate is too small or falling into an oscillatory trap if it is too big. The weight of the outgoing synapse j of the neurone i of the layer l are then corrected by:

$$\Delta w_{ij}^l = -\eta \sum_{e=1}^{N} \frac{\partial E_e}{\partial w_{ij}^l}. \tag{6.14}$$

Given that the algorithm relies on the calculation of a gradient in the synapse space, the complexity of the algorithm (i.e. the number of operations) is $\mathcal{O}(N_s)$, with N_s being the number of synapses in the network.

6.2.2.2 BFGS Algorithm

Unlike the back propagation, the Broyden-Fletcher-Goldfarb-Shannon (BFGS) [176–179] method employs the second derivatives of the error function in the recalculation of the synapse weights. In general, the BFGS algorithm is an iterative method for solving unconstrained nonlinear optimisation problems based on the *hill-climbing* technique [180].The algorithm is an approximation of the Newton's method with the Hessian matrix being approximated by evaluating the gradient.

This algorithm is robust when a large number of variables is used. It approaches the minimum in fewer iterations than the back propagation algorithm. However, it is rather unfavourable in terms of the execution time, given that the complexity of the algorithm is $\mathcal{O}(N_s^2)$ while for back-propagation it is $\mathcal{O}(N_s)$.

6.3 Limit Setting Tool

Data is tested for the presence of the $T\bar{T}$ events using a Hybrid Bayesian-Frequentist approach based on the $\mathrm{CL_s}$ method implemented in McLimit [181] described in Section 6.3.1. The basis of the method is briefly described in this section. The treatment of the systematic uncertainties is discussed in 6.3.2.

6.3.1 CL_s Method

The method employs the log-likelihood ratio (LLR) as the test statistic:

$$LLR = -2 \ln \frac{\mathcal{L}\left(data|s{+}b\right)}{\mathcal{L}\left(data|b\right)}, \tag{6.15}$$

where $\mathcal{L}\left(data|H_i\right)$ represents the likelihood of data given the hypothesis H_i. In this case, the $s{+}b$ (signal+background) hypothesis assumes the presence of T events in addition to the SM background, while the b (background-only) hypothesis assumes the SM background only. The likelihood, given hypothesis H_i, is constructed as a product of the per-bin Poisson probabilities over the bins of the discriminant distribution. The log-likelihood is then given by:

$$-2 \ln \mathcal{L}(data|H_i) = -2 \ln \mathcal{L}(\vec{n}|R,\, \vec{s},\, \vec{b}) = -2 \sum_{j=1}^{N_{bins}} (n_j \ln \mu_j - \mu_j), \tag{6.16}$$

where n_j and μ_j denote the number of observed and expected events in the bin j respectively. Thereby μ_j equals to $Rs_j(\vec{\theta}) + b_j(\vec{\theta})$, with s_j and b_j standing for the expected number of signal and background events respectively and R being a scaling parameter applied to the signal to test the sensitivity of the search. In the case of the b hypothesis, the s_j component equals zero for all bins.

According to the CL_s method, for an observed LLR value, LLR_{obs}, the probabilities of the observation being compatible with the b hypothesis (p_b) and the $s{+}b$ hypothesis (p_{s+b}) respectively are defined as:

$$p_b = 1 - \mathrm{CL_b} = p(LLR \leqslant LLR_{obs}|b), \tag{6.17}$$

$$p_{s+b} = \mathrm{CL_{s+b}} = p(LLR \geqslant LLR_{obs}|s+b). \tag{6.18}$$

Equation 6.17 represents the probability that a background fluctuation produces a LLR value as $s{+}b$-like or more than LLR_{obs}. Equation 6.18 represents the probability that the $s{+}b$ sample produces LLR_{obs} by a downward fluctuation. The distributions $(LLR|s+b)$ and $(LLR|b)$ are generated from a set of pseudo-experiments assuming the presence and the absence of the signal, respectively.

Low $\mathrm{CL_{s+b}}$ values can be due either to a low compatibility of data with the $s+b$ hypothesis or to a low sensitivity of the analysis to that hypothesis. The quantity which is expressing the compatibility of data with the $s+b$ hypothesis by normalising out the insensitivity case is: $\mathrm{CL_s} = \mathrm{CL_{s+b}}/\mathrm{CL_b}$. In terms of that quantity, the $s+b$ hypothesis is deemed to be excluded at 95% C.L. if $\mathrm{CL_s} < 0.05$.

In this analysis, the $s+b$ hypothesis is defined in terms of the T mass m_T and the measured signal strength R representing the ratio of the measured and the theoretical $T\bar{T}$ production cross section given the mass m_T: $R = \sigma^{meas}(T\bar{T}|m_{t'})/\sigma^{th}(T\bar{T}|m_{t'})$. For each tested m_T the R_{95}^{obs} is quoted which provides $\mathrm{CL_s} < 0.05$. In addition, a corresponding R_{95}^{exp} is quoted. It is assessed in the analogous way to R_{95}^{obs} by substituting LLR_{obs} by the median of $(LLR|b)$ in the equations 6.17 and 6.18.

6.3.2 Pseudo Experiment Generation and Treatment of the Uncertainties

In each pseudo-experiment (PE), a distribution of the final discriminant M_{reco} has been produced by randomly generating the event yield of each bin, using the event yield of the corresponding bin of the nominal template as the central value, and allowing for a poissonian statistical variations, as well as a gaussian variations accounting for the systematic uncertainties regarding object reconstruction, flavour tagging, background normalisations and $t\bar{t}$ modelling. Correlations of various sources of uncertainty are accounted for across the bins and processes.

In the case of k sources of systematic uncertainty, each provided by an up and down fluctuation (which are in general asymmetric), the central value of the expected number of events in the bin j, $\mu_j^{central}$, needs to be corrected by allowing for fluctuations within the uncertainties. The up and down fluctuations for the bin j are evaluated in terms of the relative differences f_{jk}^{+} and f_{jk}^{-} with respect to $\mu_j^{central}$ for all k uncertainty sources. The fluctuated bin content μ_j can be retrieved by interpolating between the central value and the provided uncertainties. In this analysis quadratic vertical interpolation has been chosen:

$$\mu_j = \mu_j^{central} \prod_k \left(1 + s_k \left(\frac{f_{jk}^{+} - f_{jk}^{-}}{2} \right) + s_k^2 \left(\frac{f_{jk}^{+} + f_{jk}^{-}}{2} \right)^2 \right), \tag{6.19}$$

with s_k being the the strength factors for each uncertainty source retrieved from the Gaussian distribution with the mean value of 0 and the standard deviation of 1. Quadratic interpolation has been chosen in order to ensure a continuous interpolation function at 0. In case of asymmetric systematic uncertainties, linear interpolation often suffers discontinuities at 0.

Chapter 7

Common Analysis Elements

The analyses presented in Chapters 8 (the t' search) and 9 (the VLT search) are both searches for pair production of heavy quarks, generically called T in this chapter, decaying into a W-boson and a b-quark. The searches are performed in the channel in which the W-boson stemming from one heavy quark decays into a lepton-neutrino pair (lepton being e or μ or their respective anti-leptons), while the W-boson resulting from its counterpart decays into a pair of quarks from the first or second SM generation. This event topology is referred to as the *lepton+jets* signature. In this chapter, the common basis of the two searches is discussed. The topology of the sought for signal events, described in Section 7.1, is followed by the definition of the basic event selection introduced in Section 7.2, as well as the discussion of the SM background processes, presented in Section 7.3. The data and simulated event samples used in the searches are presented in Section 7.4. The estimation of particular background components using data-driven and semi-data-driven methods is explained in Section 7.5. Unless a difference between the t' and VLT search is explicitly specified, this discussion applies to both searches.

7.1 Signal Event Topology

The experimental signature of a $T\bar{T}$ event in the lepton+jets final states is characterised by the presence of a single isolated lepton (e or μ) with high transverse momentum (p_{T}), a significant amount of missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) arising from the neutrino, and four high p_{T} jets. The events in which a W-boson decays into a τ-lepton and the corresponding neutrino are not explicitly included in the searches. However, τ+jets events in which the τ decays leptonically result in an e/μ+jets final state, and are thus considered. In addition, a small contribution to the selected event sample can arise from $T\bar{T}$ events in the final states with two leptons, one of which is misidentified as a jet, or one lepton and a hadronically decaying τ. A significant fraction of $T\bar{T}$ events has fewer than four jets, as some jets can be produced outside of the detector acceptance, or some jets can be merged if they are produced very close to each other[1]. The latter case often results from a hadronic decay of a highly boosted W-boson. This particular feature was exploited in the VLT search. Events with exactly three jets can still be reconstructed and are used in the analyses presented in this thesis. Very often events have more than four jets

[1] Tangent of an angle θ between the directions of motion of two particles measured in the centre-of-mass frame shrinks by a factor of $\sim 1/\gamma$ when measured in the laboratory frame in which the system experiences a boost γ. Thus, the angles between objects in a highly boosted frame tend to be small.

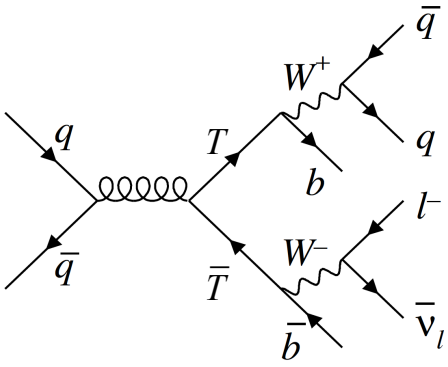


Figure 7.1: A Feynman diagram of a LO parton level $T\bar{T}$ event, produced in a $q\bar{q}$ initiated process, in the final state in which one of the W -boson decays into a $l\nu_l$ pair, while the other W-boson decays into a pair of light quarks. Two b-quarks originating from the decay of T-quarks are shown as well. Adopted from from [182].

due to ISR and FSR (see Section 4.1.3), gluon splitting, pile-up activity etc. In both analyses presented in this thesis, events with $\geqslant 3$ jets are used, whereby events with 3 jets and those with $\geqslant 4$ jets are reconstructed in different manners. In both analyses, events with the e+jets and μ+jets final states are analysed separately with a slightly different event selection. In the t' search events with 3 or $\geqslant 4$ jets are treated as separate channels in the statistical analysis, whereas in the T search the two signal regions are combined in the final discriminant.

7.2 Event Pre-selection

The same decay products are expected in the $T\bar{T}$ and the $t\bar{t}$ event for the decay mode considered. The differences between the two are expected in the event kinematics, due to the difference in the T- and t-quark masses. Thus, the basic $t\bar{t}$ event selection with small $T\bar{T}$ specific adjustments has been adopted. Although a significant improvement in object reconstruction has been achieved between the two analyses presented in this thesis, the basic $t\bar{t}$ event selection stayed similar, with minor differences which are pointed out below.

The basic selection of $T\bar{T}$ event consists of the following requirements:

- **Trigger requirement**: Each event is required to pass the lowest un-prescaled single lepton trigger requirement: the single electron trigger in the electron channel and the single muon trigger in the muon channel. The particular triggers used in each analysis are specified below, in the explanation of the data set used (Section 7.4.1).

- **Lepton requirement**: Each event must contain exactly one electron or muon fulfilling the criteria from Sections 5.3.3 and 5.4. The exclusion of events with two or more leptons reduces the contribution of the processes with two isolated leptons expected, such as Z+jets. In the t' search, the selected muon was not required to match with the object that triggered the event, as the inefficiency at L2 was found to bias the tails of the jet multiplicity distribution as well as the p_{T} distributions of the objects involved.

- **Jets multiplicity requirement**: Each event needs to have at least 3 jets satisfying the requirements explained in Section 5.5.6. Events containing at least one *Loose* jet (defined in Section 5.5.3) with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 4.5$ are rejected.

- **Leading jet p_{T} requirement**: Decay products of heavier particles (such as a T-quark) are produced with a higher momentum than those originating from lighter particles. Thus,

to suppress the $t\bar{t}$ and the other SM backgrounds, the highest p_{T} jet of the event is required to satisfy $p_{\mathrm{T}} > 60$ GeV.

- **$\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ and transverse W-boson mass requirements (triangular cut)**: A $T\bar{T}$ event in the lepton+jets final state contains one neutrino, which leads to a significantly higher $E_{\mathrm{T}}^{\mathrm{miss}}$ in the event than that caused by instrumentation effects in events with no real neutrinos. Using $E_{\mathrm{T}}^{\mathrm{miss}}$ and the reconstructed lepton, the transverse mass of the leptonically decaying W-boson can be calculated as $M_T^W = \sqrt{E_{\mathrm{T}}^l E_{\mathrm{T}}^{\mathrm{miss}} cos(\phi^l - \phi^{\mathrm{miss}})}$, where E_{T}^l and ϕ^l are the transverse energy and the azimuthal angle of the lepton, while the corresponding angle of the $E_{\mathrm{T}}^{\mathrm{miss}}$ can be computed as $\phi^{\mathrm{miss}} = \tan^{-1}(E_y^{\mathrm{miss}}/E_x^{\mathrm{miss}})$. As this variable represents the projection of the invariant W-boson mass, its distribution is expected to have a broad peak around the W mass, contrary to the steadily falling distribution expected from events in which either the lepton-$E_{\mathrm{T}}^{\mathrm{miss}}$ pair does not originate from a W, or the objects are misreconstructed. Thus, in the t' and VLT searches $M_T^W + E_{\mathrm{T}}^{\mathrm{miss}} > 60$ GeV is required. This so-called triangular cut is imposed rather than separate requirements on $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_T^W, as it ensures a higher signal acceptance, while maintaining similar levels of fake and non-prompt background rejection.

- **b-Tagging requirement**: To suppress the backgrounds with no b-quarks, especially W-boson production associated with light quarks, at least one selected jet needs to be b-tagged. In the t' (VLT) search, the JetFitterCombNN (MV1) algorithm is used for tagging. In both analyses the working point corresponding to a 70% efficiency is used. In the VLT search, the b-tagging requirement for the simulated samples is performed by applying the tag-rate-function (TRF) weights explained in Section 5.7, rather than by imposing an explicit requirement on the number of jets deemed sufficiently compatible to a b-jet according to b-tagging algorithm used.

Additional requirements are imposed to suppress non-collision backgrounds, and account for instrumentation and reconstruction issues.

- **Primary vertex requirement**: Each event is required to have at least one well-defined reconstructed primary vertex with at least five tracks of $p_{\mathrm{T}} > 400$ MeV associated to it, to ensure sufficient vertex position precision. This requirement further eliminates the non-collision backgrounds, such as proton collisions with beam gas. In the VLT search, only tracks with $p_{\mathrm{T}} > 400$ MeV are considered in this requirement, whereas all tracks above 100 MeV are considered in the t' search.

- **Object overlap removal**: As various object reconstruction algorithms often use the same detector signals, some signals can be shared between two or more objects, and thus counted more than once. To remove this effect, and also to reduce specific backgrounds, the following object removal procedures are performed:

  - Electron-muon overlap removal:
    Events in which an electron candidate and a muon candidate share an ID track are rejected.
  - Muon-jet overlap removal:
    The selected muon is removed if it is closer than $\Delta R(\mu, j) = 0.4$ to a selected jet, in order to reduce the contribution of the muons stemming from the decays of the heavy flavour hadrons.

- Jet-electron overlap removal:
  Each electron is also reconstructed as a jet, since the jet reconstruction is performed independently on the same energy deposits in the calorimeter as the electron reconstruction. To avoid double-counting, all jets closer than $\Delta R(e, j) = 0.2$ around a selected electron are removed; if there are further jets with $p_{\mathrm{T}} > 20$ GeV within $\Delta R(e, j) = 0.4$, the electron is discarded.

- **LAr noise and FEBs requirement**: ATLAS data taken in 2011 suffered from sporadic noise bursts in LAr calorimeters. These bursts affected individual events, which were rejected. Events in which at least one jet with corrected $p_{\mathrm{T}} > 20$ GeV had a direction within 0.1 in both η and ϕ around a noisy LAr front-end board were rejected as well. The latter was applied to both data and MC [183].

7.3 Backgrounds

The SM processes passing the basic selection presented above are backgrounds to the $T\bar{T}$ searches. The production cross sections of these background processes along with other representative SM processes measured on the ATLAS pp collision data at $\sqrt{s} = 7$ and 8 GeV are summarised in Figure 7.2. The computed $T\bar{T}$ production cross section for the masses examined in this thesis ranges between approximately 0.1 and 1.5 pb, which is several orders of magnitude lower than most of the SM backgrounds. The SM processes passing the pre-selection defined above and their general treatment are discussed in this section. The modelling of each individual process is discussed in Section 7.4.2.

7.3.1 $t\bar{t}$ background

Events of $t\bar{t}$ production contain the same final state particles as $T\bar{T}$ events, and thus represent an irreducible background (Figure 7.3 (a)). As the basic pre-selection described in the previous section is optimised to extract $t\bar{t}$ events, these events are the dominant contribution to the pre-selected sample, while the other backgrounds are suppressed. Thus, the selection optimisation beyond the pre-selection is mainly targeted towards separating $T\bar{T}$ signal from the $t\bar{t}$ background. The $t\bar{t}$ contribution is estimated from the simulated events (see Section 7.4.2.2).

7.3.2 W+jets background

Events with leptonically decaying W-boson produced in association with jets (e.g. Figure 7.3 (d), (e), and (f)) can have a high level of similarity with the $T\bar{T}$ signal, depending on the flavour of the additional jets. Events with two b-jets and additional light jets (W+$b\bar{b}$+jets) exactly match the signal topology, and are thus an irreducible background. However, this contribution has a low production rate. Events with one or two c-quarks (W+$c\bar{c}$+jets and W+c+jets) have a moderate chance of passing the pre-selection, as the high p_{T} c-jets have a relatively high probability of being tagged as b-jets (Figure 5.15). Events with only light jets in addition to a W-boson are efficiently suppressed by the b-tagging requirement, since the mis-tag efficiency is rather low (Figure 5.16). This component is, however, still dominant, due to a higher production rate compared to the other components of the W+jets background. The

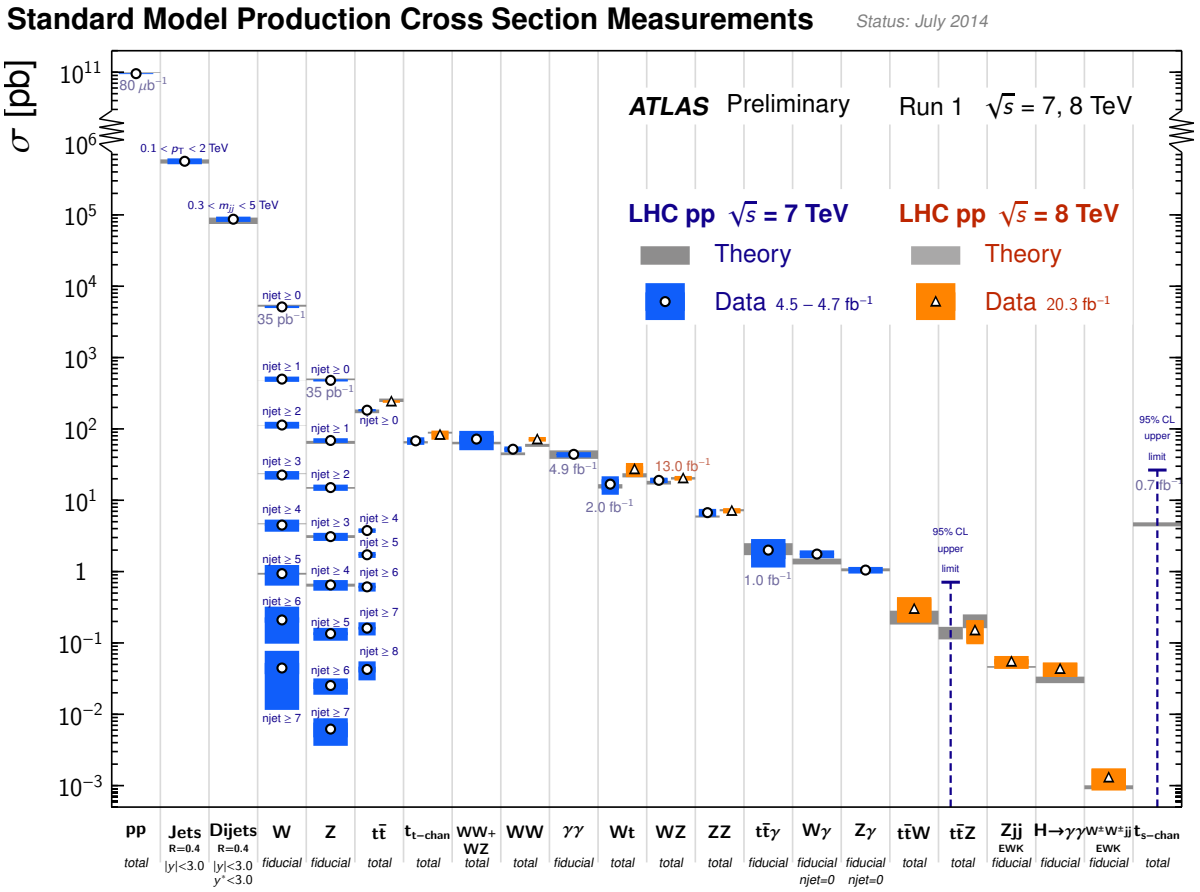


Figure 7.2: A summary plot of the total and fiducial production cross section measurements of representative SM processes performed with the ATLAS pp collision data at $\sqrt{s} = 7$ and 8 TeV. The cross sections are corrected for leptonic branching functions and compared to corresponding theoretical expectations. The data points are assigned total, and the theoretical expectations theoretical uncertainty bands. The integrated luminosity used is indicated for each measurement. Taken from [184].

contribution of the W+jets events is estimated using simulated distributions with a data driven normalisation that will be explained in Section 7.5.2.

7.3.3 Single top background

In pp collisions, single top quarks can be produced via the weak interaction in three possible final states, shown at LO in Figures 7.3 (j), (k), and (l). Single top events have on average fewer jets than $T\bar{T}$ events and can be selected only if the jets resulting from higher order contributions or from pile-up are sufficiently hard. It is also necessary that the W-boson in t- and s-channels, or at least one W-boson in the Wt-channel, decay leptonically, which is additionally suppressed by the branching ratio. The single top contribution is estimated from MC simulations (see Section 7.4.2.2).

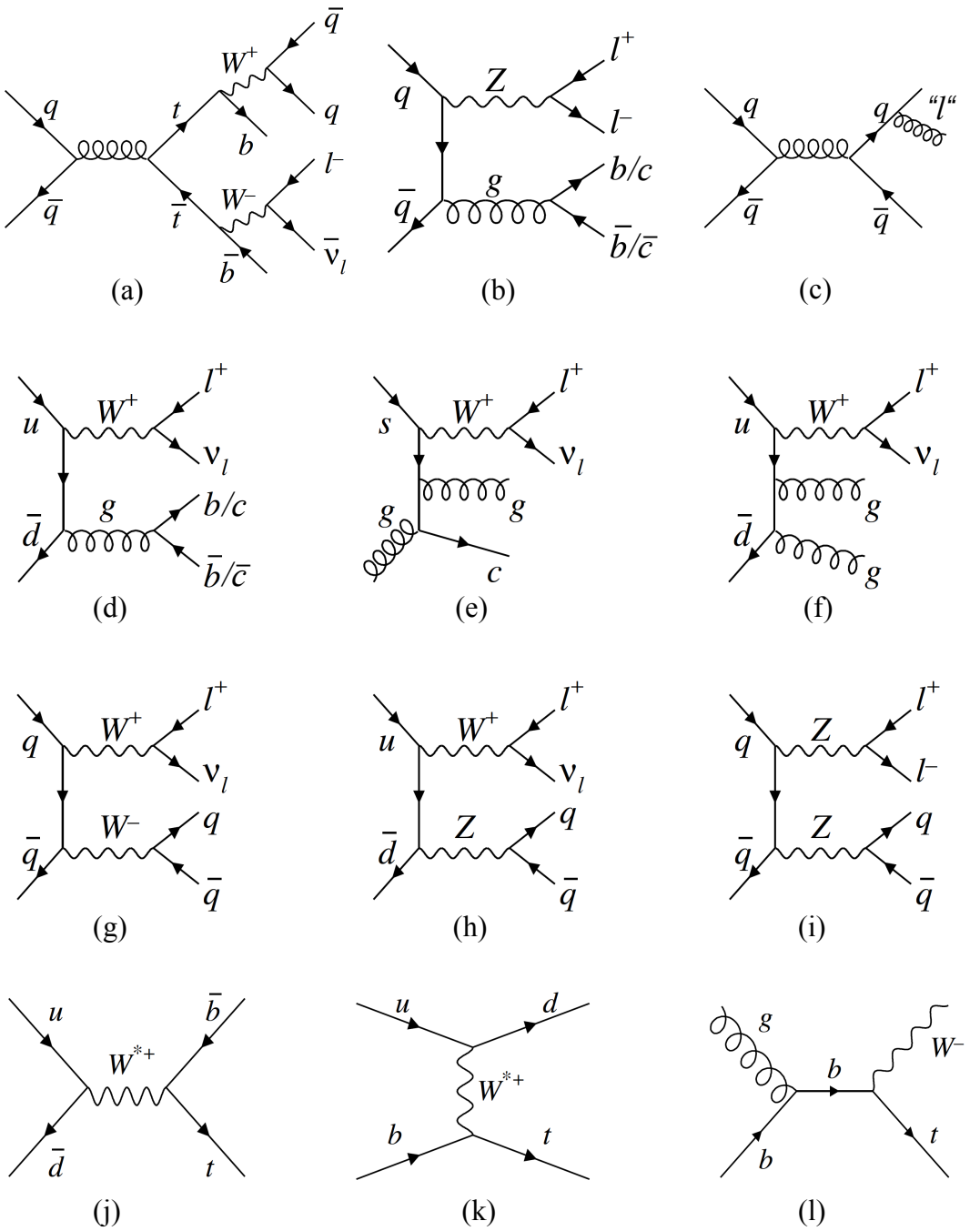


Figure 7.3: Representative Feynman diagrams of the SM backgrounds passing the pre-selection: $t\bar{t}$ production (a), $Z + b\bar{b}/c\bar{c}$ production (b), QCD multi-jet production (c), $W + b\bar{b}/c\bar{c}$ production (d), W+c+light jets production (e), W+light jets production (f), WW production (g), WZ production (h), ZZ production (i), single top production in the s-channel (j), single top production in the t-channel (k), and single top production in the Wt-channel (l). Additional light jets in the event can arise from gluon radiation off gluons or quarks. Taken from [182,185].

7.3.4 Fake and non-prompt lepton background

Multi-jet production via QCD processes (Figure 7.3 (c)) has a rather high cross section, which decreases as increasingly higher jet multiplicity is required. These events pass the pre-selection in case they contain one isolated fake or non-prompt lepton, that is a lepton which does not originate from the hard process but from some later stage of the event development. Fake leptons are usually jets mis-identified as leptons. Non-prompt leptons can originate from semi-leptonic decays or heavy-flavour bound states (e.g. B-mesons) contained in a jet. Electrons can also stem from photon conversions, while muons can be produced in the weak decays of charged pions and kaons. The triangular cut (see Section 7.2) and b-tagging requirement remove the majority of multi-jet events which have passed the other pre-selection requirements. Due to complexity and large variety of processes involved, the QCD multi-jet background simulation is extracted in a data-driven manner, as explained in Section 7.5.1.

7.3.5 Z+jets background

Z-boson production in association with jets (Figure 7.3 (b)) can pass the pre-selection if the Z-boson decays into a $\ell^+\ell^-$ pair, where one of the leptons is either outside of the acceptance or mis-identified. Apart from a small fraction of events in which two b-jets are produced in the event ($Z + b\bar{b}$+jets), the events containing light jets can be efficiently rejected by the b-tagging requirement. Since the Z+jets events do not have a real source of neutrinos, the triangular cut (see Section 7.2) is also efficient in eliminating them (the system of a lepton and a fake $E_{\mathrm{T}}^{\mathrm{miss}}$ is not expected to have a mass similar to the W boson mass.). The Z+jets contribution to the pre-selected sample is estimated entirely from simulation (see Section 7.4.2.2).

7.3.6 Di-boson background

A small contribution to the pre-selected sample comes from di-boson production, WW, WZ and ZZ, shown in Figure 7.3 (g), (h) and (i). Such events pass the pre-selection if exactly one prompt, non-prompt or fake lepton passes the lepton selection requirements, and if there are additional sufficiently hard jets in the event. The majority of events which fulfill these criteria fail to pass the b-tagging requirement. This background is estimated from simulated events (see Section 7.4.2.2).

7.4 Event Samples

All samples for a given analysis are produced using a specific Athena [186] release. An Athena release comprises a fixed set of the event generator and reconstruction software versions, which is used for both simulated and data event samples. The t' analysis is based on release 16 of the Athena software, while the VLT search is based on release 17, which incorporated a significantly improved performance of the detector simulation. The data and MC samples used in these two searches are discussed in this section.

7.4.1 Data

The searches presented in this thesis (Chapters 8 and 9) are based on the pp collision data at $\sqrt{s} = 7$ TeV delivered by the LHC and collected by the ATLAS detector in 2011. Both analyses are divided in two analyses channels: the e+jets and the μ+jets channel. For these two channels events are selected in different trigger streams: in the `Egamma` stream for the e+jets channel, and the `Muon` stream for the μ+jets channel. For both searches, the single lepton triggers with the lowest un-prescaled p_{T} threshold were used. The effect of the overlap between the two streams is minimised by imposing the offline selection described in Section 7.2. Only those luminosity blocks, that is units in which data is split within one run, are used that were collected with stable proton beams and with all subsystems of ATLAS fully operational.

7.4.1.1 Data for the t' Search

The search for the fourth generation t'-quark (Chapter 8) is performed using data collected from March until June 2011 in the periods B through G5. The early 2011 runs in the period B started with only 138 bunches per beam and a bunch spacing of 75 ns yielding a maximal instantaneous luminosity of $\sim 0.2 \cdot 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, to reach 1318 bunches per beam at the bunch spacing of 50 ns, resulting in a maximal instantaneous luminosity of $\sim 1.3 \cdot 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ in the period H. The data with a quality sufficient for the analysis corresponds to an integrated luminosity of 1.04 fb^{-1} with an uncertainty of 3.7% [187]. A problem with the lost LAr calorimeter front-end boards, affecting approximately 0.4% of the calorimeter coverage, was corrected. The number of interactions per bunch crossing averaged over individual luminosity blocks reaches up to approximately 6−9 in the periods considered. The `EF_e20_medium` and the `EF_mu18` triggers were used for the Egamma and the Muon stream, respectively.

7.4.1.2 Data for the VLT Search

The data used for the VLT search (Chapter 9) consists of full 2011 ATLAS data taking campaign. From March until October a total integrated luminosity of 5.2 fb^{-1} was collected, before selecting the good runs fulfilling the detector integrity requirements. Compared to the first half of the year, the maximal instantaneous luminosity was increased to $\sim 3.5 \cdot 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. In the last physics data taking period (M10), some test runs were performed with the bunch spacing of 25 ns and 1842 bunches per beam yielding instantaneous luminosities up to $\sim 3.8 \cdot 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and mean number of interactions per bunch crossing of up to 18. In the electron stream, the `EF_e20_medium` trigger was used for periods B-H while the `EF_e22_medium` was used for the periods I-K. For periods L-M, a logical OR of `EF_e22vh_medium1` and `EF_e45_medium1` was employed. In the muon stream, the `EF_mu18` trigger was used for periods B-I and `EF_mu18_medium` for the later periods. The data fulfilling these requirements, taking the correction for the lost LAr cells into account, corresponds to an overall integrated luminosity of 4.7 fb^{-1} with an uncertainty of 3.9% [187]. The run ranges and the separate integrated luminosities of the sub-periods are summarised in Table 7.1.

Periods	B-D	E-H	I	J	K	L-M
Run Range	177986 - 180481	180614 - 184169	185353 - 186493	186516 - 186755	186873 - 187815	188902 - 191933
Luminosity [pb^{-1}]	173.798	924.67	328.61	220.38	575.156	2368.38

Table 7.1: Data periods and the corresponding integrated luminosities for the VLT search (Chapter 9).

7.4.2 Monte Carlo Samples

To optimise the accuracy of the event simulation, the choice of the event generator, the PDF set and the corresponding generator tunes, is adapted individually to each process individually, as explained below. In both analyses, pile-up is modeled by overlaying minimum bias events generated using Pythia6 [130] in the simulation. For the VLT search, pile-up contribution is generated by the Pythia 6.4 generator [130] using the ATLAS AMBT2B tune [188] and the CTEQ6L1 [189] PDF set. The simulated distribution of the mean number of interactions per bunch crossing is reweighted such that it matches the data. Except for several samples specialised for studies of systematic uncertainties and processed by the fast simulation, other samples are processed by the full Geant4 [137, 138] based detector simulation (see Section 4.2 for an explanation). Between the two searches, the simulation tools have undergone improvements which are not explained in detail in this thesis. In the following, simulated signal and background samples used in the t' and VLT search are discussed.

7.4.2.1 Signal Modeling

Samples of strongly produced t' pairs generated by Pythia 6.421 with the MRST LO* [190] PDF set are used as signal in the t' and VLT search. The elements of the extended CKM matrix used in the event generation correspond to the central values of the fit provided by the CKMFitter group [191]:

$$\begin{bmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| & |V_{ub'}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| & |V_{cb'}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| & |V_{tb'}| \\ |V_{t'd}| & |V_{t's}| & |V_{t'b}| & |V_{t'b'}| \end{bmatrix} = \begin{bmatrix} 0.97417 & 0.22530 & 0.00341 & 0.01460 \\ 0.22443 & 0.972711 & 0.04117 & 0.04200 \\ 0.00642 & 0.03769 & 0.99413 & 0.10125 \\ 0.02393 & 0.04062 & 0.10003 & 0.99387 \end{bmatrix}, \tag{7.1}$$

where t' and b' designate the chiral fourth generation quarks. For both searches presented in this thesis, the only assumed decay is $t'/T \to W^+b$ (which is why these samples can be used for the VLT search presented in this thesis). For the t' search, samples were generated for masses between 200 and 500 GeV in steps of 50 GeV. The mass reach for the VLT search was extended, so samples were generated for masses between 400 and 750 GeV in steps of 50 GeV. In both cases, events that at generator level contain at least one lepton with $p_{\mathrm{T}} > 10$ GeV are used. The corresponding theoretical cross sections for all mass points are computed at approximate next-to-next-to-leading order (NNLO) using the Hathor tool [192] with the CTEQ66 [193] PDF set, following the ATLAS conventions [194].

The uncertainties on the cross sections arise from the renormalisation and factorisation scale variations, as well as PDF uncertainties. The scale uncertainty account for the missing higher-

t' mass	$\sigma_{t'\bar{t}'}$ (pb)	μ_F and μ_R uncertainty (pb)	PDF+α_s uncertainty (pb)
200	75.73	$+1.64/-4.29$	$+6.16/-5.45$
250	22.52	$+0.55/-1.29$	$+2.30/-1.96$
300	8.02	$+0.22/-0.46$	$+0.96/-0.80$
350	3.24	$+0.10/-0.19$	$+0.44/-0.36$
400	1.44	$+0.04/-0.08$	$+0.22/-0.17$
450	0.68	$+0.02/-0.04$	$+0.11/-0.09$
500	0.34	$+0.01/-0.02$	$+0.06/-0.04$

Table 7.2: Theoretical production cross sections with the corresponding scale and PDF uncertainties for all t' mass points considered in the t' search.

T mass	$\sigma_{T\bar{T}}$ (pb)	μ_F and μ_R uncertainty (pb)	PDF+α_s uncertainty (pb)	total uncertainty (pb)
400	1.406	$+0.045/-0.083$	$+0.176/-0.138$	$+0.182/-0.161$
450	0.6623	$+0.0227/-0.0395$	$+0.0865/-0.0647$	$+0.0894/-0.0758$
500	0.3296	$+0.0119/-0.0198$	$+0.0448/-0.0318$	$+0.0463/-0.0375$
550	0.1714	$+0.0065/-0.0104$	$+0.0242/-0.0167$	$+0.0251/-0.0197$
600	0.09227	$+0.00364/-0.00566$	$+0.01355/-0.00918$	$+0.01404/-0.01078$
650	0.05113	$+0.00210/-0.00317$	$+0.00781/-0.00522$	$+0.00809/-0.00610$
700	0.02901	$+0.00123/-0.00182$	$+0.00464/-0.00306$	$+0.00481/-0.00356$
750	0.01679	$+0.00074/-0.00107$	$+0.00287/-0.00184$	$+0.00296/-0.00212$

Table 7.3: Theoretical production cross sections with the corresponding scale, PDF and total uncertainties for all T mass points considered in the VLT search.

order perturbative corrections and are estimated by evaluating the cross section with the scales varied up and down by a factor of 2, as explained in Section 4.1.1. The central values for the renormalisation and factorisation scales used thereby are $\mu_R = \mu_F = m_T$ [194]. The scales are varied both simultaneously and one at the time, never differing from each other by more than a factor of 2. The larger variations from the central value are taken as the common scale uncertainties. The PDF uncertainties are evaluated as the maximal difference between the cross section values obtained using different PDF sets and their error sets. As various PDFs are derived using different values of α_s, a combined PDF+α_s uncertainty is used, which is derived based on the PDF4LHC procedure [195]. The cross section and uncertainty computations used in the VLT search have been slightly improved with respect to those used in the t' search. The calculated cross sections and the corresponding uncertainties used in both searches are presented in Tables 7.2 and 7.3.

Samples of 75k events for each mass point were provided by the central MC production of the ATLAS collaboration. For the purpose of the NN training, larger signal samples with $\sim$ 1M events have been generated privately for masses of 500 and 650 GeV. The need for high statistics signal samples for the training purpose is elaborated in Appendix C.3.

7.4.2.2 Background Modeling

Except for the QCD multi-jet background, which is extracted from data (see Section 7.5.1), the shapes of all background processes are estimated from simulation. The MC samples for

the background processes are adopted from the ATLAS top working group. A brief description of these samples is presented below. In cases where different versions of tools or different normalisations are used between the two searches, the information corresponding to the t' search is quoted in the brackets.

The top quark samples The dominant $t\bar{t}$ background is generated using the MC@NLO v4.01 (v3.41) [133] event generator with the CT10 [196] (CTEQ66 [193]) PDF set. The parton shower and the underlying event are simulated using HERWIG v6.520 (v6.510) [131] and JIMMY v4.31 [132] generators with the ATLAS AUET2 [197] (AUET1 [198]) tunes. The generated $t\bar{t}$ sample is normalised to the approximate next-to-next-to leading order (NNLO) production cross section of $166.78^{+16.5}_{-17.8}$ pb (164.6 pb) computed with HATHOR [192]. Sample of 15 M generated events are used in each analysis, corresponding to $\sim 150\text{fb}^{-1}$. The single top sample in the s-channel and Wt-channel is generated with the same tools as the $t\bar{t}$ sample. Due to the modeling deficiencies of the MC@NLO and HERWIG setup, the t-channel samples are generated with the ACERMC v3.8 generator [134] interfaced with PYTHIA v6.425 and using MRST LO** PDF set [190] and the ATLAS AUET2B [188]. The samples are normalised to the NNLO cross sections of 64.6 pb [199], 4.6 pb [200] and 15.7 pb [201] for the t-, s-, and Wt-channel respectively. The reference cross sections for $t\bar{t}$ and single top samples are evaluated using MSTW2008 NNLO PDF sets [123]. The statistical power of the samples is 300k, 1M and 900k events for the s-, t- and Wt-channel, respectively.

The vector boson samples The production of the vector bosons W and Z in association with up to 5 partons is simulated using the ALPGEN generator [135] interfaced to HERWIG [131] and JIMMY [132] generators tuned using the ATLAS AUET2 tunes [197]. The CTEQ6.1 PDF set [202] has been used for both matrix element calculation and the parton shower evolution. The subprocesses involving various composition of the additional partons are generated separately. The considered processes comprise the vector boson production in association with light partons only (W/Z+light jets) as well as heavy + light partons ($W+c$+jets, $W+c\bar{c}$+jets, $W+b\bar{b}$+jets, $Z+c\bar{c}$+jets and $Z+b\bar{b}$+jets) produced in the matrix element. The double counting of the heavy quark production is taken care of by the heavy flavour overlap removal tool [203]. The W+jets sample is normalised using a data driven method explained in 7.5.2, while the Z+jets samples are normalised to the corresponding NNLO cross section predictions [204]. The W/Z+jets samples are generated with a statistics that corresponds to $\sim 100\text{fb}^{-1}$. The production of di-boson events (WW, WZ and ZZ) is simulated using HERWIG and normalised to the NNLO prediction [204].

7.4.2.3 Summary of MC Samples

A short overview of all simulated samples used in the VLT analysis is presented in Table 7.4. For completeness, the samples used for estimating the $t\bar{t}$ modelling uncertainties discussed in 9.6.4.1 are also included in this table. The details of the generation are provided for each sample in the text above.

Process	MC Generator	PDF set	Detector sim.	MC Stat
$t't'$	Pythia	MRST LO*	full	75k/1M
$t\bar{t}$ nominal	MC@NLO+Herwig/Jimmy	CT10	full	15M
singletop t-chann	AcerMC+Pythia	MRST LO**	full	4.8M
singletop s- / Wt-chann	MC@NLO+Herwig/Jimmy	CTEQ6.6	full	
W+jets	Alpgen+Herwig (/Jimmy)	CTEQ6L1	full	46M
Z+jets	Alpgen+Herwig (/Jimmy)	CTEQ6L1	full	34 M
$WW/WZ/ZZ$	Herwig	CTEQ6L1	full	4M
$t\bar{t}$ nominal AFII	MC@NLO+Herwig/Jimmy	CT10	AFII	15M
$t\bar{t}$ more/less PS	AcerMC+Pythia	CTEQ6L1	AFII	10M
$t\bar{t}$ PowHg	Powheg+Herwig	CTEQ6L1	AFII	10M
$t\bar{t}$ PowPy	Powheg+Pythia	CTEQ6L1	AFII	3M

Table 7.4: Summary of the MC generation details of the simulated event samples used.

7.5 Data Driven Background Estimation

7.5.1 QCD Multi-jet Background Estimation

The QCD multi-jet background for the analyses presented in this thesis is fully extracted from data using the so-called matrix method technique. The method relies on differences in the selection efficiency between the prompt leptons (coming from W or Z decays) and the non-prompt (coming from the heavy flavour hadrons, photon conversions etc.) or misidentified (so-called fake) leptons. The method uses two data samples: one called "tight", that is a subset of the other called "loose". The tight selection includes the final lepton identification, while the loose selection has no lepton isolation requirement, and employs loosened identification requirements in the electron channel, in order to enrich the sample in fake leptons. The number of events for the selections N^{loose} and N^{tight} can then be expressed as:

$$N^{\text{loose}} = N^{\text{loose}}_{\text{real}} + N^{\text{loose}}_{\text{fake}}, \tag{7.2}$$

$$N^{\text{tight}} = \epsilon_{\text{real}} N^{\text{loose}}_{\text{real}} + \epsilon_{\text{fake}} N^{\text{loose}}_{\text{fake}}, \tag{7.3}$$

where $N^{\text{loose}}_{\text{real}}$ and $N^{\text{loose}}_{\text{fake}}$ signify the number of real and fake leptons passing the loose selection respectively, while $\epsilon_{\text{real}} = N^{\text{tight}}_{\text{real}}/N^{\text{loose}}_{\text{real}}$ and $\epsilon_{\text{fake}} = N^{\text{tight}}_{\text{fake}}/N^{\text{loose}}_{\text{fake}}$ stand for the efficiency of the real/fake leptons passing the loose selection to also pass the tight selection. Then, the number of fake leptons passing the tight selection requirements can be expressed as:

$$N^{\text{tight}} = \frac{\epsilon_{\text{fake}}}{\epsilon_{\text{real}} - \epsilon_{\text{fake}}} \left(N^{\text{loose}} \epsilon_{\text{real}} - N^{\text{tight}} \right). \tag{7.4}$$

The efficiencies ϵ_{real} are typically measured using the T&P method on $Z \to \ell\ell$ events: ϵ_{real} is measured in the signal region and ϵ_{fake} is measured in a control region having a richer fake lepton contribution.

7.5.2 W+jets Background Normalisation and Heavy Flavour Decomposition

The W+jets background is a composite sample consisting of Wbb+jets,Wcc+jets,Wc+jets and W+jets components with up to 5 additional jets, where b and c stand for corresponding quarks

and j signifies light jets originating from u, d or s-quarks. The theoretical production cross sections of the individual components, as well as their relative fractions (expressed by the heavy-flavour (HF) scale factors), are relatively poorly known, which makes it difficult to accurately predict the overall W+jets normalisation. Therefore, a combined simulation and data driven approach is applied.

7.5.2.1 Normalisation

The technique used for measuring the W+jets normalisation exploits the asymmetric production cross sections for W^+- and W^--boson at the LHC, due to the relative differences of the PDFs of the proton valence quarks (Figure 4.4). The production rate of W^+-bosons scales as the product $u(x_u)\bar{d}(x_{\bar{d}})$, where $u(x_u)$ stands for the PDFs of u- or c-quarks at the corresponding proton momentum fraction x_u (at the currently accessible $\sqrt{s}$, excitations of $b\bar{b}$ pairs in the quark see are negligible in pair production), and $\bar{d}(x_{\bar{d}})$ stands for the PDF of d or s flavoured anti-quark at the corresponding $x_{\bar{d}}$. The production rate of W^--boson, in turn, scales as $d(x_d)\bar{u}(x_{\bar{u}})$, which has a different distribution than $u(x_u)\bar{d}(x_{\bar{d}})$, because at high values of x, $u(x_u)$ is higher than $d(x_d)$ due to the fact that a proton has two u and one d valence quarks. Knowing the ratio of theoretical cross sections $r^{MC}_{+/-} \equiv \frac{\sigma(pp\to W^+)}{\sigma(pp\to W^-)}$ to within a few percent at the LHC energies, the overall W+jets normalisation $N_{W^+} + N_{W^-}$ can be extracted as:

$$N_{W^+} + N_{W^-} = \frac{r^{MC}_{+/-} + 1}{r^{MC}_{+/-} - 1} \left(N_{W^+} - N_{W^-}\right), \tag{7.5}$$

by measuring the charge asymmetry $N_{W^+} - N_{W^-}$ in the production of W-bosons. The observed charge asymmetry originates predominantly from W+jets events given that the other background processes containing a real or mis-reconstructed W-boson, such as $t\bar{t}$, multi-jet, and Z+jets processes, are charge symmetric. Being a function of the parton momentum fraction x, the cross section ratio $r^{MC}_{+/-}$, and thus also the charge asymmetry $N_{W^+} - N_{W^-}$, need to be estimated separately in different bins of jet multiplicity, N_j. The measurement is performed on the sample selected by the pre-selection ignoring the b-tagging (un-tagged selection) requirement to allow for a higher statistical power.

From the estimated un-tagged event yield of W+jets background in each N_j bin, a combined data/MC method is used to estimate the fraction of W+jets events expected to contain at least one b-tagged jet. This estimation is performed in the bin with $N_j = 2$, since that bin provides the smallest systematic uncertainties. The number of events of W-boson production associated with n jets, whereby at least one among them is b-tagged, $N^{\geqslant 1\,\mathrm{tag}}_{W,\,n}$, is estimated as:

$$N^{\geqslant 1\,\mathrm{tag}}_{W,\,n} = N^{\mathrm{pretag}}_{W,\,n} \cdot f^{2j}_{\mathrm{tag}} \cdot f^{2\to n}_{\mathrm{tag}}, \tag{7.6}$$

with $N^{\mathrm{pretag}}_{W,\,n}$ being the number of $W + n$ jets events passing the un-tagged selection, the tagging fraction for the 2 jet bin defined as $f^{2j}_{\mathrm{tag}} = N^{\geqslant 1\,\mathrm{tag}}_{W,\,n}/N^{0\,\mathrm{tag}}_{W,\,n}$, and the ratio of the tagging fractions in the n jet bin (n=2, 3, $\geqslant$ 3, 4, $\geqslant$ 4, and $\geqslant$ 5) and the 2 jet bin $f^{2\to n}_{\mathrm{tag}} = f^{nj}_{\mathrm{tag}}/f^{2j}_{\mathrm{tag}}$. Factors f^{2j}_{tag} and $f^{2\to n}_{\mathrm{tag}}$ are estimated from the un-tagged pre-selected data sample after subtracting all backgrounds other than W+jets.

7.5.2.2 Heavy Flavour Composition

To decompose the event yields $N_W^{\geqslant 1\,\mathrm{tag}}$ according to the flavour composition, the so-called HF scale factors K_{bb}, K_{light} and K_c are introduced, which reflect the relative abundance of the flavour components. The derivation of these scale factors is discussed in this section.

The W+jets event yield in the bin i of the un-tagged pre-selected sample can be decomposed as:

$$N_{W,\,i}^{pretag} = N_{bb,\,i}^{pretag} + N_{cc,\,i}^{pretag} + N_{c,\,i}^{pretag} + N_{light,\,i}^{pretag}. \tag{7.7}$$

into contributions of the $W + b\bar{b}$+jets, $W + c\bar{c}$+jets, $W + c$+jets and W+light jets. Here, each $N_{x,\,i}^{pretag}$ ($x = bb$, cc, c, or light) can be expressed as $N_{W,\,i}^{pretag} \cdot F_{x,\,i}$, using the flavour fractions $F_{x,\,i}$. The flavour fractions in each bin sum up to unity: $\sum_x F_{x,\,i} = 1$. To derive the corresponding event yield in the sample with at least one b-tagged jet, the tagging probabilities $P_{x,\,i}$ derived from the simulation need to be taken into account:

$$N_{W,\,i}^{tag} = N_{W,\,i}^{pretag} \sum_x F_{x,\,i} P_{x,\,i}. \tag{7.8}$$

	$N_j = 3$	$N_j = 4$	$N_j \geqslant 5$
$W + b\bar{b}$+jets	2915.7	898.5	305.6
$W + c\bar{c}$+jets	5673.6	1491.9	495.1
$W + c/\bar{c}$+jets	9391.0	1918.5	458.6
W+ light jets	39804.5	8341.8	2090.4

Table 7.5: The W+jets event yields for the un-tagged electron channel event pre-selection, as a function of the W+jets component and the reconstructed jet multiplicity.

	$N_j = 3$	$N_j = 4$	$N_j \geqslant 5$
$W + b\bar{b}$+jets	5354.3	1670.5	582.0
$W + c\bar{c}$+jets	10826.8	2827.1	873.6
$W + c/\bar{c}$+jets	16860.4	3416.9	833.3
W+ light jets	76083.3	15836.9	3920.9

Table 7.6: The W+jets event yields for the un-tagged muon channel event pre-selection, as a function of the W+jets component and the reconstructed jet multiplicity.

The fractions $F_{x,\,i}$ can from here be derived by demanding each measured contribution to be equal to the corresponding contribution estimated in the simulation:

$$N_{W,2}^{pretag} F_{x,2}^{data} = N_{W,2}^{pretag,\,MC} F_{x,2}^{MC}, \tag{7.9}$$

from where the HF scale factors can be calculated as:

$$K_{x,2} \equiv \frac{F_{x,2}^{data}}{F_{x,2}^{MC}} = \frac{N_{W,2}^{pretag,\,MC}}{N_{W,2}^{pretag}}. \tag{7.10}$$

These scale factors multiplied by an appropriate normalisation factor $1/A_i$ are applied in each bin i. The normalisation factor A_i is computed by requiring the sum of measured HF fractions to add up to unity:

$$\frac{1}{A_i}\sum_x K_{x,2}F_{x,i}^{MC} = 1. \tag{7.11}$$

These HF scale factors are derived for events with a positively and a negatively charged prompt lepton separately. They are applied to the measured event yields to compute the final normalisation for each flavour component in each bin of jet multiplicity. The resulting un-tagged yields are presented in Tables 7.5 and 7.6 for the electron and the muon channels separately.

Chapter 8

Search for Fourth Generation Top-Like Quark t'

The analysis summarised in this chapter was the first direct search for the fourth sequential SM generation t'-quark performed by the ATLAS Collaboration. It was performed using 1.04 fb^{-1} of LHC data at $\sqrt{s} = 7$ TeV in event topologies including one electron or muon, at least three jets and significant missing transverse energy. This event topology is discussed in Section 7.1. The basic analysis infrastructure comprising the event pre-selection, the estimation of the SM backgrounds as well as the data and simulated samples used, has been presented in Chapter 7. The reconstruction of the t' and $\bar{t}'$ quarks has been performed employing different methods for events with three jets, and those with at least four jets. In the latter case, the kinematic likelihood fitter KLFitter, introduced in Section 6.1, was employed. The contribution to this search provided by the author of this thesis is a set of KLFitter performance studies, which led to establishing the optimal fitter configuration for this analysis. The search is published in [205]. This chapter is organised as follows. The observed and expected event yields obtained after the pre-selection described in Section 7.2 are summarised in Section 8.1. The KLFitter performance studies carried out by the author are elaborated in Section 8.3. The treatment of the systematic uncertainties, the statistical analysis and the final result are discussed in Section 8.4. Conclusions are presented in Section 8.5.

8.1 Pre-selection Event Yields

Events that pass the pre-selection explained in Section 7.2 are categorised according to the number of selected jets. The search is performed with events containing either exactly 3, or at least 4 jets. Events with lower jet multiplicities are used for validation of the background predictions, particularly the W+jets background. The event yields of the pre-selected samples with no b-tagging requirement are summarised in Table 8.1 for the e+jets channel, and 8.2 for the μ+jets channel, as a function of jet multiplicity. The corresponding numbers for the sample with at least one b-tagged jet are summarised in Tables 8.3 and 8.4. The tables show that there is no evidence of a significant excess of data compared with the expected SM background in these two pre-selected samples. A quantification of the compatibility between the observation and the expectation in terms of shapes of the relevant distributions, however, requires a dedicated event reconstruction and statistical analysis, which are presented in the following sections.

Event category	= 1 jet	= 2 jets	= 3 jets	$\geqslant$ 4 jets
$t'\bar{t}'(400)$	0.30 ± 0.11	6.20 ± 0.54	22.71 ± 1.02	108.39 ±2.25
$t\bar{t}$	135.75 ± 5.55	853.67 ± 13.84	1854.38 ± 20.57	3422.05 ±28.27
QCD	4599.67 ± 48.32	3081.21 ± 39.25	1280.68 ± 25.76	668.19 ±19.48
W+jets	31575.32 ± 346.99	19186.99 ± 133.62	7196.52 ± 68.37	3027.89 ±40.60
Z+jets	3891.23 ± 33.26	2324.32 ± 25.69	1000.22 ± 16.86	539.32 ±12.19
Single top	254.05 ± 5.90	526.65 ± 7.83	332.43 ± 5.45	212.23 ±3.91
Di-bosons	231.86 ± 5.58	281.65 ± 6.02	127.61 ± 3.93	46.62 ±2.34
Total prediction	40687.88 ±352.05	26254.49 ±142.63	11791.85 ±78.04	7916.30 ±54.74
Data	41638	26227	11971	8075

Table 8.1: The event yields and the corresponding statistical uncertainties as a function of the jet multiplicity in the pre-selected pre-tag sample for the e+jets channel. The event yields are listed for the expected signal corresponding to $m_{t'} = 400$ GeV, the expected SM backgrounds, as well as the data.

Event category	= 1 jet	= 2 jets	= 3 jets	$\geqslant$ 4 jets
$t'\bar{t}'(400)$	0.49 ± 0.14	5.76 ± 0.47	22.98 ± 0.93	104.22 ±1.94
$t\bar{t}$	177.07 ± 5.53	1053.21 ± 13.72	2423.18 ± 20.78	4475.38 ±28.64
QCD	3465.36 ± 47.39	2612.78 ± 39.75	1231.17 ± 26.60	601.31 ±18.42
W+jets	53203.86 ± 397.21	33282.01 ± 153.82	11645.09 ± 80.82	4342.49 ±45.89
Z+jets	2381.51 ± 23.79	1987.08 ± 21.53	898.51 ± 14.27	460.41 ±9.91
Single top	420.08 ± 6.83	778.38 ± 8.60	470.19 ± 5.82	261.03 ±3.92
Di-bosons	349.15 ± 6.22	430.22 ± 6.72	180.86 ± 4.28	64.80 ±2.56
Total prediction	59997.04 ±400.88	40143.68 ±161.28	16848.99 ±89.04	10205.41 ±58.18
Data	60217	40000	16889	10235

Table 8.2: The event yields and the corresponding statistical uncertainties as a function of the jet multiplicity in the pre-selected pre-tag sample for the μ+jets channel. The event yields are listed for the expected signal corresponding to $m_{t'} = 400$ GeV, the expected SM backgrounds, as well as the data.

8.2 Event Reconstruction

In this analysis, the data are tested for the presence of the $t'\bar{t}'$ events by exploring the reconstructed t'-quark mass distribution. A simple estimation of the reconstructed t'-quark mass can be made by assuming that the three leading jets in the event originate from the hadronically decaying t'- or $\bar{t}'$-quark. With that assumption, the t'-quark mass can be approximated by the reconstructed invariant mass of the leading p_{T} three-jet system, m_{jjj}. This method provides a good estimate for the top mass reconstruction. However, for heavy quarks with a mass of 300 GeV or higher, this method is strongly biased towards lower masses (Figures 8.1 (a) and (b)), which demonstrates the inaccuracy of the assumption. A study relying on the information from the generator level jets matched to the reconstructed jets used in the t'-quark reconstruction has shown that in $\sim 20\%$ of events in the pre-selected t' sample with a mass of 400 GeV, the

Event category	= 1 jet	= 2 jets	= 3 jets	$\geqslant$ 4 jets
$t'\bar{t}'(400)$	0.22 ± 0.10	4.31 ± 0.44	15.06 ± 0.77	80.73 ±1.84
$t\bar{t}$	48.55 ± 3.09	456.31 ± 9.64	1181.43 ± 15.56	2318.96 ±22.23
QCD	361.42 ± 17.51	415.24 ± 20.06	220.22 ± 17.28	126.07 ±18.02
W+jets	683.07 ± 35.90	856.26 ± 24.74	510.45 ± 16.87	301.99 ±11.68
Z+jets	7.80 ± 1.42	34.49 ± 3.01	28.50 ± 2.72	23.69 ±2.47
Single top	93.81 ± 3.55	251.36 ± 5.10	175.29 ± 3.78	123.98 ±2.83
Di-bosons	7.73 ± 0.85	16.58 ± 1.21	9.69 ± 0.87	4.18 ±0.58
Total prediction	1202.38 ±40.26	2030.24 ±33.82	2125.58 ±29.12	2898.87 ±31.14
Data	1136	1953	2184	3216

Table 8.3: The Event yields and the corresponding statistical uncertainties as a function of the jet multiplicity in the pre-selected sample with $\geqslant$1 b-tags for the e+jets channel. The event yields are listed for the expected signal corresponding to $m_{t'} = 400$ GeV, the expected SM backgrounds, as well as the data.

Event category	= 1 jet	= 2 jets	= 3 jets	$\geqslant$ 4 jets
$t'\bar{t}'(400)$	0.14 ± 0.06	3.97 ± 0.37	16.12 ± 0.73	78.19 ±1.59
$t\bar{t}$	71.99 ± 3.38	567.08 ± 9.54	1513.90 ± 15.63	3067.49 ±22.65
QCD	304.51 ± 12.29	407.87 ± 13.96	244.15 ± 10.83	142.39 ±8.57
W+jets	1035.41 ± 28.84	1441.46 ± 27.46	775.93 ± 19.11	420.14 ±12.90
Z+jets	24.06 ± 2.28	51.04 ± 3.38	40.02 ± 2.96	25.53 ±2.27
Single top	168.55 ± 4.17	378.94 ± 5.71	257.57 ± 4.12	149.22 ±2.84
Di-bosons	12.35 ± 0.97	26.64 ± 1.46	12.85 ± 0.98	6.09 ±0.71
Total prediction	1616.87 ±31.90	2873.04 ±32.95	2844.43 ±27.45	3810.86 ±27.69
Data	1681	2930	3003	3987

Table 8.4: The event yields and the corresponding statistical uncertainties as a function of the jet multiplicity in the pre-selected sample with $\geqslant$1 b-tags for the μ+jets channel. The event yields are listed for the expected signal corresponding to $m_{t'} = 400$ GeV, the expected SM backgrounds, as well as the data.

leading p_{T} jet does not originate from any of the hard scattering partons. Furthermore, in approximately 66% of events the b-jet originating from the leptonic side of the decay is one of the three hardest jets of the event. In both cases (which might overlap), the m_{jjj} is constructed from the incorrect set of jets. However, for events with only three jets, this method still provides a better estimate than the alternative option, i.e. the reconstruction of the leptonic side of the decay from the lepton, $E_{\mathrm{T}}^{\mathrm{miss}}$ and one jet, with no information on the longitudinal component of the neutrino momentum. Thus, the thee-jet invariant mass method is used for the t'-quark mass reconstruction in the sample with exactly three selected jets.

On the other hand, if four or more jets are present in the event, both sides of the decay, i.e. the hadronically and the leptonically decaying quark, can be reconstructed: the hadronically decaying quark is reconstructed by combining the 4-vectors of three jets, while the leptonically decaying quark is reconstructed by combining one jet, the lepton and the $E_{\mathrm{T}}^{\mathrm{miss}}$, whereby the

longitudinal component of the neutrino momentum is retrieved by requiring the reconstructed invariant mass of the lepton-neutrino pair to be compatible with the W-boson mass within the corresponding particle width and detector resolution. To assign the reconstructed jets to one of the quarks, various approaches can be followed. In this analysis, events are reconstructed employing the kinematic fitting KLFitter, introduced in Section 6.1. Provided with four jets, a lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$, KLFitter tests the compatibility of each permutation of the reconstructed objects to the real final state of the hard process, by minimising the underlying log likelihood function explained in Section 6.1.1.1. This function incorporates the particle width constraints on the W-boson and the t'-quark, as well as the modelling of object energy resolution. Assuming that the four leading p_{T} jets of the event originate from the hard scattering, the t'-quark mass provided by KLFitter, $m_{t'}^{par}$, is reconstructed with a significantly better resolution compared to the m_{jjj} method, as can be seen by comparing Figures 8.1 (**??**) and (**??**), or Figures 8.1 (**??**) and (**??**). The t' mass reconstruction performance of the fitter can be adjusted by tuning the user defined fit settings, which is discussed in the next section.

8.3 KLFitter Performance Studies

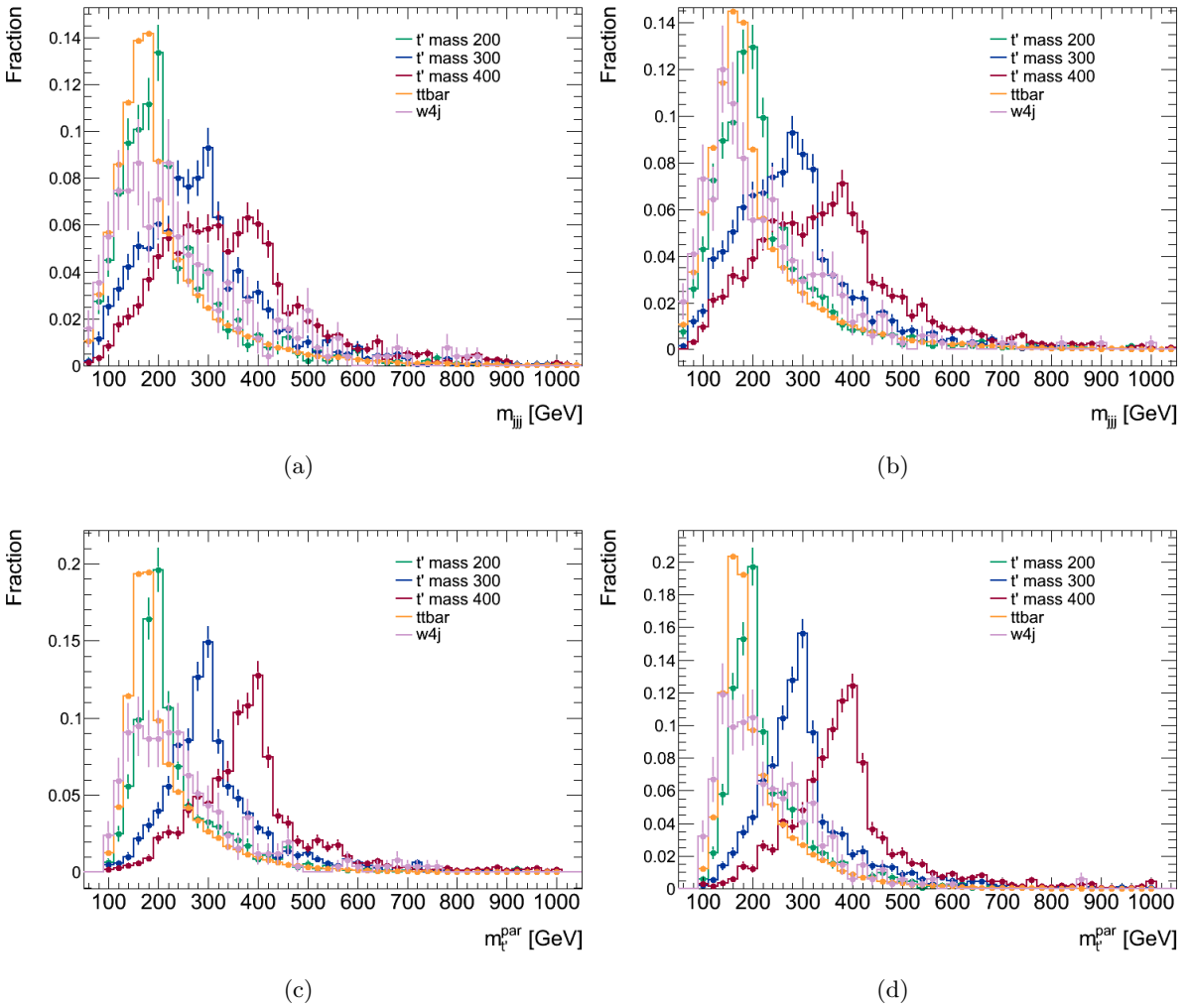
The KLFitter performance for a given final state is largely dependent on the user defined settings, such as the choice of the jets used in the event reconstruction, the mass of the heavy quark to be reconstructed, use of the b-tagging information in the fit, etc. In order to establish the set of the KLFitter settings that provides the optimal fit performance, i.e. the best resolution of the reconstructed t'-quark mass, a set of KLFitter performance studies was performed by the author of this thesis. These studies are presented in Sections 8.3.1 – 8.3.3. Additional studies aimed to improve the understanding of the KLFitter performance in the presence of more and less ISR and FSR, as well as its energy resolution, are presented in Section 8.3.4 and 8.3.5 respectively. Unless stated otherwise, all studies were performed using the t' sample with $m_{t'} = 400$ GeV as signal.

8.3.1 Number of permuted jets

The calculation of the KLFitter likelihood function requires the four-momenta of the reconstructed lepton and four jets (one for each final state parton), as well as the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$, along with its projections onto the x and y axes. If KLFitter is provided with n jets with $n > 4$, all $\binom{n}{4}$ combinations of four jets are tested. Each possible assignment of the reconstructed jets to the final state partons is tested. In case KLFitter does not use the per-jet b-tagging information, that is if reconstructed b-tagged jets are not restricted to be matched only to b-partons, 4!/2 permutations are examined for each combination.[1] Thus, if n is larger than five, the computation time becomes unacceptably long (as the number of the likelihood minimisation steps is multiplied by the number of the jet permutations).

The choice of the jets used in the fitting procedure is user defined. The most common choice is to permute only the four hardest jets. However, the four leading p_{T} jets are not necessarily those coming from the hard process of interest. In roughly 30% of events in the preselected nominal t' sample (with t'-quark mass of 400 GeV) containing at least five selected jets, one of

[1] The number of tested permutations is reduced by the factor 1/2 since the reconstructed t'-quark mass is invariant under the permutation of the two light jets.



(a) (b)

(c) (d)

Figure 8.1: Comparison of the reconstructed mass of the three leading jets ((a) and (b)) and the t'-quark mass fitted using KLFitter where four leading jets are assumed to originate from the hard process ((c) and (d)). Figures (a) and (c) correspond to the e+jets, while Figures (b) and (d) correspond to the μ+jets channel. The distributions are shown for three assumed t'-quark masses (200, 300 and 400 GeV) as well as for the two most abundant backgrounds, $t\bar{t}$ and W+jets.

the hard process partons matches the fifth leading jet. Considering five jets in the KLFitter thus increases the chance of using the right jets in the event reconstruction, but also the number of the possible permutations (and thus the processing time) by the factor of $\binom{5}{4}$. However, it turns out that, when the reconstructed t' mass is treated as a fit parameter, the resulting fitted t' mass distribution has a strong bias towards low masses when the five leading p_{T} jets are permuted. In Fig. 8.2, distributions of the fitted t' mass $m_{t'}^{par}$ produced by permuting four (red line) and five (blue line) leading p_{T} jets are shown for events with at least five jets in the electron (left) and muon (right) channels.

The origin of such behaviour has been studied using the available generator level information relying on the truth matching described in Section 6.1.1.2. Figure 8.3 shows the decomposi-

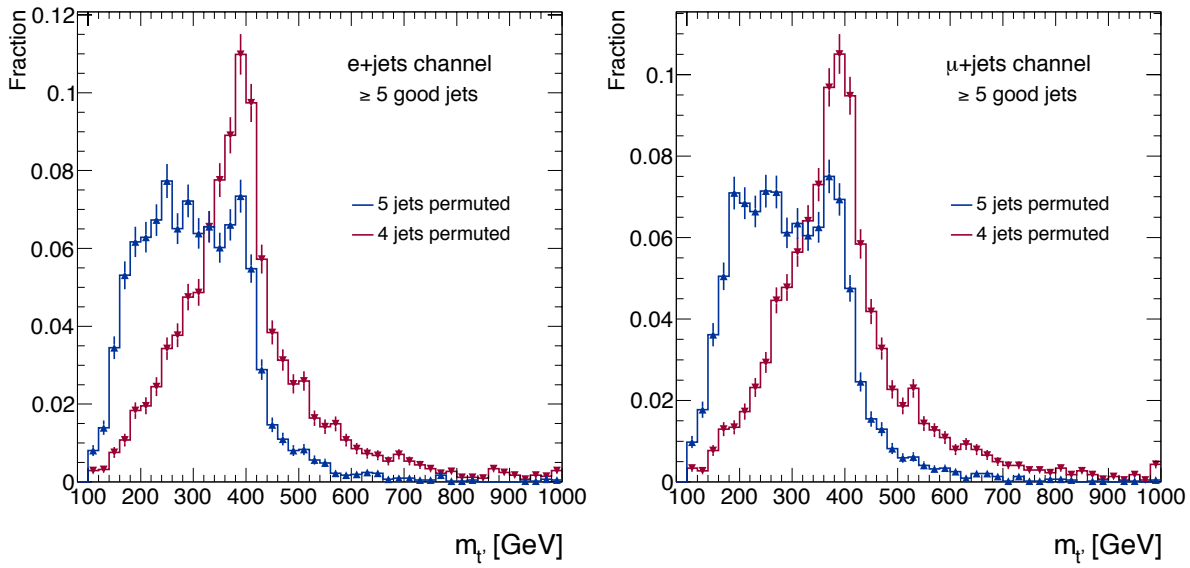


Figure 8.2: Reconstructed t'-quark mass determined in events with $\geqslant 5$ jets, by finding the best permutation of either the four jets (red line) or five (blue line) leading jets in the event. The distributions correspond to the nominal t' mass of 400 GeV in the electron (left) and muon (right) channel separately.

tion of the fitted $m_{t'}$ distribution according to different truth matching scenarios. All relevant numbers deduced from this figure are summarised in Table 8.5 for the e+jets and μ+jets channels separately, and briefly discussed in the following. Figures 8.3 (a) and (b) shows that in approximately 83% of preselected events with $\geqslant 5$ reconstructed (so-called *good*) jets, the fifth hardest jet enters the mass reconstruction, if five leading jets are permuted. However, in only about 30% of events the fifth leading jet is actually matched to a final state jet at the generator level. In almost all these events the fifth jet is used in the fit, that is identified as one of the four tree level partons. Nonetheless, in almost 60% of events, at least one generator level final state jet has no match among the selected reconstructed jets (orange line in Figures 8.3 (c) and (d)). Among the remaining $\sim 40\%$ with four so-called *matched* jets, in only one quarter of events ($\sim 10\%$ in total) all four final state jets are properly recognised by the fit, i.e. *correctly identified* (regardless the permutation of the light jets), as shown by the blue distributions in the middle row plots. Finally, in only about 5% of events, one among four properly identified jets is the fifth leading jet (green distributions in Figures 8.3 (a) – (d)). For a comparison, in case only four leading jets are permuted in the same event sample, in only about 25% of events all four final state generator level jets are matched to reconstructed jets (orange distribution in Figures 8.3 (e) and (f)), while the fraction of events with four properly identified jets increases to approximately 16% (blue distributions in Figures 8.3 (e) and (f)). This suggests that in the sample with at least 5 reconstructed jets, better performance is achieved when four leading jets are fitted. A corresponding set of plots shown in Figure B.4 of Appendix B shows that in a similar fraction of $t\bar{t}$ events the fifth leading jet is used in the fit, while the percentage of events in which it has a valid match to a generator level final state jet increases to approximately 40%.

The fitted $m_{t'}$ distribution can further be decomposed according to the number of matched objects, number of properly identified jets or the type of mis-identifications (e.g. if b-jets are swapped etc.). The identification efficiency of the KLFitter, in the sense of the fraction of events

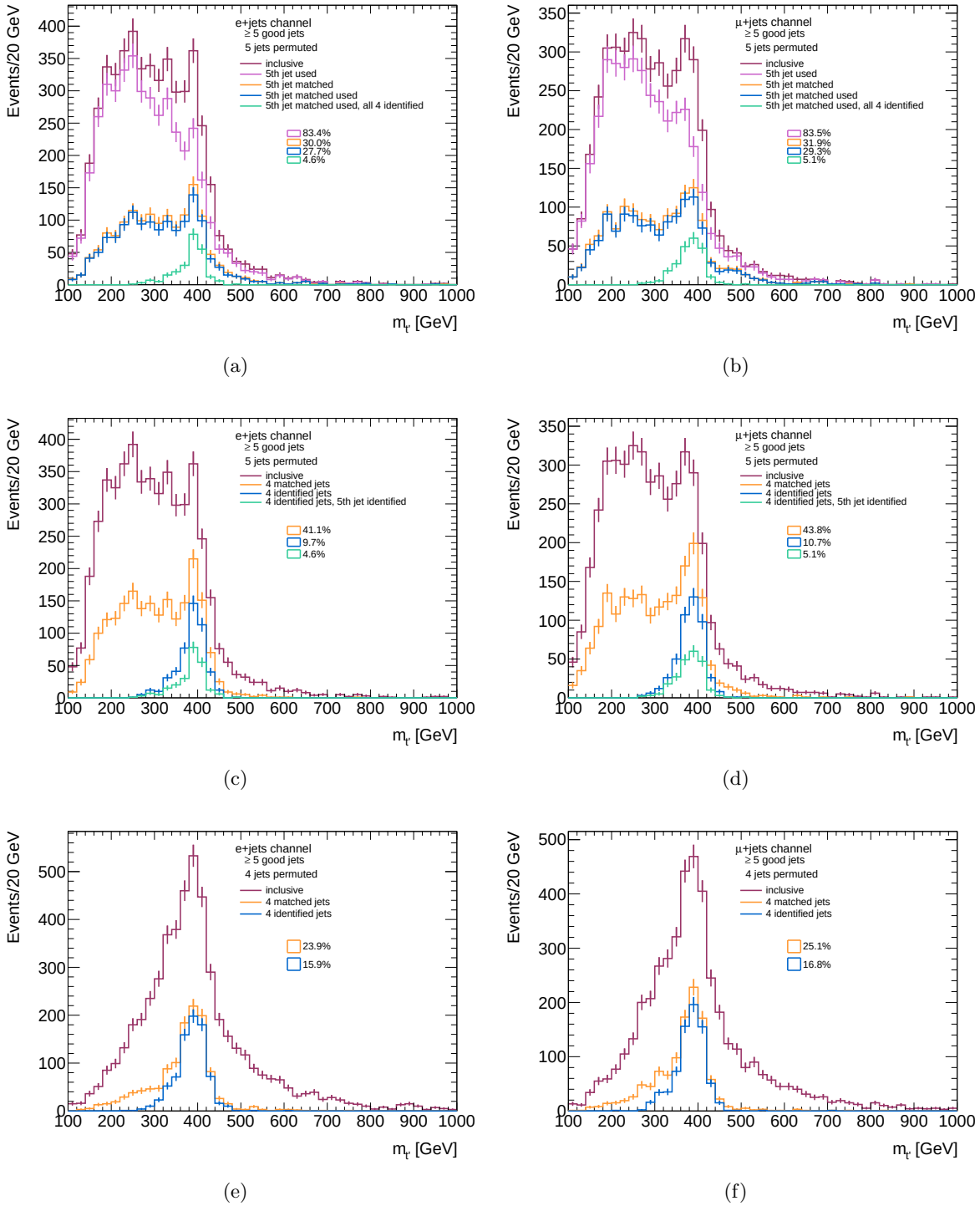


Figure 8.3: Decomposition of the fitted t'-quark mass ($m_{t'}$) distributions in the e+jets (left) and μ+jets (right) channels for the events with $\geqslant 5$ reconstructed jets obtained by permuting five (top and middle) or four (bottom) leading jets. Figures (a) and (b) show the fractions of events in which the fifth leading jet is: used in the mass reconstruction (pink line), matched to a generator level final state jet (orange line), matched and used in the mass reconstruction (blue line), matched and used in events with all four jets properly identified. Figures (c) – (f) show the fractions of events in which: four reconstructed jets are matched to generator level final state jets (orange line) or properly identified (blue line). In addition, Figures (c) – (d) show the fraction of events with four identified jets among which is also the fifth hardest reconstructed jet.

Channel	Configuration	5^{th} jet used	5^{th} jet match.	5^{th} jet match., used	4 match. jets	4 id. jets	4 id. 5^{th} id.
			t'				
e+jets	$\geqslant 5$ good, 5 fit	83.4%	30.0%	27.7%	41.1%	9.7%	4.6%
	$\geqslant 5$ good, 4 fit	–	–	–	23.9%	15.9%	–
μ+jets	$\geqslant 5$ good, 5 fit	83.5%	31.9%	29.3%	43.8%	10.7%	5.1%
	$\geqslant 5$ good, 4 fit	–	–	–	25.1%	16.8%	–
			t				
e+jets	$\geqslant 5$ good, 5 fit	84.5%	39.7%	36.3%	35.8%	8.9%	5.9%
	$\geqslant 5$ good, 4 fit	–	–	–	13.4%	6.4%	–
μ+jets	$\geqslant 5$ good, 5 fit	84.5%	40.4%	37.0%	37.2%	8.5%	5.8%
	$\geqslant 5$ good, 4 fit	–	–	–	13.9%	6.5%	–

Table 8.5: Truth matching summary for e+jets and μ+jets channels. The fraction of events in which the fifth leading p_{T} jet is: used for the mass reconstruction, matched to a final state jet, or both matched and used are presented. Furthermore, the fractions of events in which four jets are matched to the final state jets and properly identified, as well as those with four identifications among which is also the fifth leading jet, are shown.

as a function of the number of properly identified jets, can be plausibly measured only on the fully matched events, i.e. those with four jets and the lepton matched to the corresponding hard process final state objects at the generator level. The efficiency measurement as well as various sample decompositions are shown in Figures 8.4 – 8.6 for the preselected e+jets sample and three different fitting configurations:

1. $\geqslant 4$ selected jets in the event, 4 leading jets permuted in the fit,
2. $\geqslant 5$ selected jets in the event, 4 leading jets permuted in the fit, and
3. $\geqslant 5$ selected jets in the event, 5 leading jets permuted in the fit.

The first category comprises all events available for the analysis in the channel with $\geqslant 4$ jets and employs the most common choice of jets provided to the fit, i.e. four leading p_{T} jets. The other two categories are restricted to the sub-sample with $\geqslant 5$ selected jets. The results in these two categories are compared, in order to study the effect of the inclusion of the fifth leading p_{T}jet in the permutations. The plots are shown for the e+jets channels only. The corresponding plots for the μ+jets channel are shown in Figures B.1–B.3, while the plots produced using the generated $t\bar{t}$ sample for both channels are presented in Figures B.5–B.10 of Appendix B.

8.3.1.1 Number of Truth Matches and Identification Efficiency

The truth matching efficiencies achieved for the preselected nominal t' ($t\bar{t}$) sample using various fit configurations introduced in the previous section, are summarised in Table 8.6 (8.7). The corresponding object identification efficiencies are presented in Table 8.8 (8.9). The numbers presented are briefly discussed in the following.

Configuration	4 jets matched	3 jets matched	2 jets matched	1 jet matched	0 jets matched	lepton not matched	event not matched
			e+jets				
$\geqslant$ 4 good, 4 fit	31.4%	35.8%	14.9%	2.6%	0.22%	0.22%	14.9%
$\geqslant$ 5 good, 4 fit	23.9%	42.1%	17.9%	3.3%	0.30%	0.34%	12.4%
$\geqslant$ 5 good, 5 fit	41.1%	32.9%	11.3%	1.8%	0.18%	0.34%	12.4%
			μ+jets				
$\geqslant$ 4 good, 4 fit	32.8%	35.3%	14.7%	2.7%	0.25%	0.22%	14.2%
$\geqslant$ 5 good, 4 fit	25.0%	42.0%	17.6%	3.6%	0.30%	0.27%	11.5%
$\geqslant$ 5 good, 5 fit	43.6%	31.4%	10.9%	2.0%	0.14%	0.27%	11.5%

Table 8.6: Truth matching summary for e+jets and μ+jets channels for the nominal t' sample with $m_{t'} = 400$ GeV. The fraction of events for each number of matched jets bin is shown.

According to Figure 8.4, about 15% of the events of the inclusive analysis sample ($\geqslant 4$ selected jets per event) in the e+jets channel, have no semi-leptonic final state at the generator level (yellow contribution in Figure 8.4 (a)). About 1/3 of events in the same sample are fully matched (dark purple contribution in the same plot). The KLFitter efficiency of correctly identifying all four jets is 68.2% for these fully matched events (dark purple contribution in Figure 8.4 (b)). In most of remaining fully matched events (26.2%), two jets are properly identified (blue contribution in the same plot). In the sub-sample with $\geqslant 5$ good jets, the fraction of fully matched events increases from $\sim$ 24% when four leading jets are considered (dark purple contribution in Figure 8.4 (c)) to $\sim$ 40% when five leading jets are considered (dark purple contribution in Figure 8.4 (e)). However, while the efficiency of the full event identification remains high, i.e around 67%, for the case of four leading jets permuted (dark purple contribution in Figure 8.4 (c)), this efficiency drops to only about 24% when five leading jets are considered (dark purple contribution Figure 8.4 (f)). In the former case, the majority of remaining events ($\sim$ 27%) has two identified jets (blue contribution in the middle row right plot), whereas in the latter case a significant fraction of events has three or one identifications, while the fraction of events with no identifications increases by almost a factor of 4 (comparison between the orange contributions in Figures 8.4 (d) and (f)). At the same time, the enhancement of the low tail of the fitted t' mass distribution $m_{t'}$, which is mostly populated by the events with two identifications when four leading jets are permuted, obtains a significant increase due to the events with three, one and zero identifications in the case where five leading jets are permuted (see the same two plots). Nevertheless, in the latter case the high tail of the distribution is less pronounced. The fraction of events with no semi-leptonic signature at the generator level is about 12% in the subsample with $\geqslant 5$ jets (yellow contribution in Figures 8.4 (c) and (e)). The situation is similar for the t' sample in the μ+jets channel (Figure B.1), as well as for the dominant $t\bar{t}$ background sample in both channels (Figures B.5-B.6).

The plots in Figure 8.5 show the shapes of the contributions corresponding to different levels of identification and different mismatch scenarios. The plots show that, in general, with a decreased number of truth matches per event, the mean value of the t' distribution decreases. The bias is significantly stronger if five leading p_{T} jets are permuted (Figures 8.5 (e) and (f)). While the distributions with less than three successful identifications peak around 350 GeV if four leading jets are permuted, (Figures 8.5 (a) and (c)), those produced by permuting five

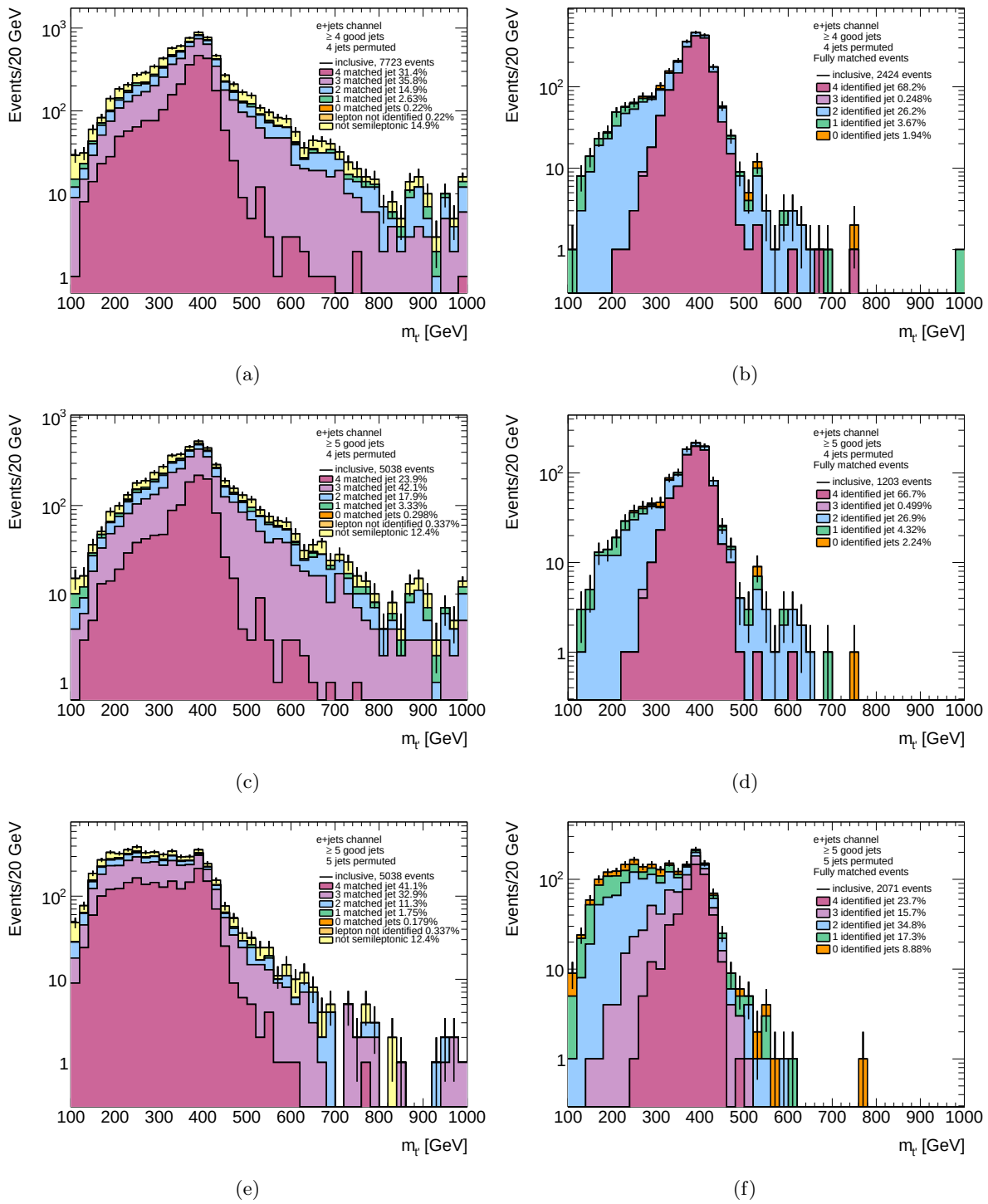


Figure 8.4: Distribution of the fitted t'-quark mass decomposed according to: the number of truth matches ((a), (c), and (e)) and the number of proper identifications in events with four truth matches ((b), (d), and (f)) shown in stacked plots. The distributions are shown for the samples with $\geqslant 4$ ((a) and (b)) and $\geqslant 5$ ((c) – (f)) reconstructed jets in the event in the e+jets channel. The distributions have been produced by fitting four ((a) – (d)) or five ((e) and (f)) leading jets in the nominal t' sample with $m_{t'} = 400$ GeV.

Configuration	4 jets matched	3 jets matched	2 jets matched	1 jet matched	0 jets matched	lepton not matched	event not matched
			e+jets				
$\geqslant$ 4 good, 4 fit	22.6%	42.0%	16.9%	2.6%	0.15%	0.43%	15.3%
$\geqslant$ 5 good, 4 fit	13.3%	47.4%	22.5%	3.7%	0.24%	0.52%	12.3%
$\geqslant$ 5 good, 5 fit	35.7%	37.4%	12.1%	1.7%	0.09%	0.52%	12.3%
			μ+jets				
$\geqslant$ 4 good, 4 fit	24.0%	42.6%	16.5%	2.4%	0.14%	0.49%	13.8%
$\geqslant$ 5 good, 4 fit	13.8%	48.7%	22.2%	3.5%	0.20%	0.60%	11.2%
$\geqslant$ 5 good, 5 fit	37.0%	37.7%	11.7%	1.6%	0.09%	0.59%	11.1%

Table 8.7: Truth matching summary for e+jets and μ+jets channels for the $t\bar{t}$ sample. The fraction of events for each number of matched jets bin is shown.

Configuration	4 jets identified	3 jets identified	2 jets identified	1 jet identified	0 jets identified
		e+jets			
$\geqslant$ 4 good, 4 fit	68.2%	0.25%	26.2%	3.7%	1.9%
$\geqslant$ 5 good, 4 fit	66.7%	0.50%	26.9%	4.3%	2.2%
$\geqslant$ 5 good, 5 fit	23.7%	15.7%	34.8%	17.3%	8.9%
		μ+jets			
$\geqslant$ 4 good, 4 fit	69.8%	0.45%	24.6%	3.2%	2.2%
$\geqslant$ 5 good, 4 fit	66.8%	0.82%	27.3%	3.3%	2.5%
$\geqslant$ 5 good, 5 fit	24.5%	16.2%	34.3%	16.0%	9.4%

Table 8.8: KLFitter identification efficiency measured on the fully matched events for e+jets and μ+jets channels for the nominal t' sample with $m_{t'} = 400$ GeV.

Configuration	4 jets identified	3 jets identified	2 jets identified	1 jet identified	0 jets identified
		e+jets			
$\geqslant$ 4 good, 4 fit	47.7%	0.27%	35.9%	11.7%	4.4%
$\geqslant$ 5 good, 4 fit	48.1%	0.48%	36.0%	11.3%	4.16%
$\geqslant$ 5 good, 5 fit	24.0%	12.8%	30.5%	23.6%	9.3%
		μ+jets			
$\geqslant$ 4 good, 4 fit	46.8%	0.26%	36.1%	12.3%	4.6%
$\geqslant$ 5 good, 4 fit	47.1%	0.49%	36.1%	11.9%	4.5%
$\geqslant$ 5 good, 5 fit	22.8%	12.4%	30.6%	24.6%	9.5%

Table 8.9: KLFitter identification efficiency measured on the fully matched events for e+jets and μ+jets channels for the $t\bar{t}$ sample.

Configuration	4 jets identified	3 jets identified	2 jets identified	1 jet identified	0 jets identified	lep. not matched	evt. not matched
			e+jets				
$\geqslant$ 4 good, 4 fit	22.0%	11.3%	17.8%	19.6%	14.3%	0.22%	14.9%
$\geqslant$ 5 good, 4 fit	16.4%	12.9%	18.1%	22.4%	17.6%	0.34%	12.4%
$\geqslant$ 5 good, 5 fit	10.4%	10.4.9%	25.1%	22.0%	19.7%	0.34%	12.4%
			μ+jets				
$\geqslant$ 4 good, 4 fit	23.4%	11.2%	16.9%	19.1%	15.1%	0.22%	14.2%
$\geqslant$ 5 good, 4 fit	17.1%	13.0%	17.6%	22.2%	18.6%	0.27%	11.5%
$\geqslant$ 5 good, 5 fit	11.2%	10.8%	25.6%	22.1%	18.9%	0.27%	11.5%

Table 8.10: Truth matching summary for e+jets and μ+jets channels for the nominal t' sample with $m_{t'} = 400$ GeV. The fraction of events for each number of matched jets bin is shown.

leading jets have a mean value around 250 GeV (Figures 8.5 (e) and (f)).

In terms of mismatches, in a large fraction of events, i.e. approximately 20% (12%) when four (five) leading jets are permuted, the reconstructed jets matched to the generator level b-jets are swapped (light magenta distributions in Figures 8.5 (b), (d) and (d)). The second largest contribution comes from the events in which the b-jet from the leptonic side of the decay is identified as a light jet (green line in the plots). All kinds of mismatches pull the mean value of the fitted $m_{t'}$ distribution downwards, while at the same time increasing the high tail of the distribution. The corresponding plots for the t' sample in the μ+jets channel is presented in Figure B.2. The analogous plots for the $t\bar{t}$ background are presented in Figures B.7–B.8. The fractions of events corresponding to each kind of mismatch are summarised in Table 8.12 for the preselected nominal t' sample, and in Table 8.13 for the preselected $t\bar{t}$ sample.

The overall identification efficiency, i.e. that derived from all events regardless of the number of truth matches, is shown in Figure 8.6 for the preselected nominal t' sample in the e+jets channel. Different components of the fitted $m_{t'}$ distribution according to the number of identifications per event are shown in a stacked plot normalised to the integrated luminosity assuming the production cross section, and as individual shapes normalised to unity. The plots show that the overall identification efficiency is higher while the bias of the distribution towards the low values is less pronounced in case only four leading jets are permuted (Figures 8.6 (a) – (d)). The corresponding plots for the t' sample in the μ+ jets channel, as well as those for the $t\bar{t}$ background are shown in Figures B.3 and B.9-B.10 of Appendix B.

Conclusion The study presented in this section has shown that in events with 5 or more selected jets, the probability of including the jets originating from the hard process partons in the t' mass reconstruction is $\sim$ 1.7 times higher if five hardest jets are provided to the fit, compared to the case when only four hardest jets are used (Table 8.6). In the former case, however, the jet identification efficiency is significantly lower than in the latter case - the probability of identifying all four jets correctly is $\sim$ 1.6 times lower in the former case than in the latter (Table 8.10). In the inclusive sample with 4 or more selected jets, the fitted t'-quark mass distribution $m_{t'}$ produced from the preselected nominal t' sample, using the four hardest jets in the fit peaks at $\sim$ 380 GeV with a width of $\sim$ 120 GeV, while the corresponding

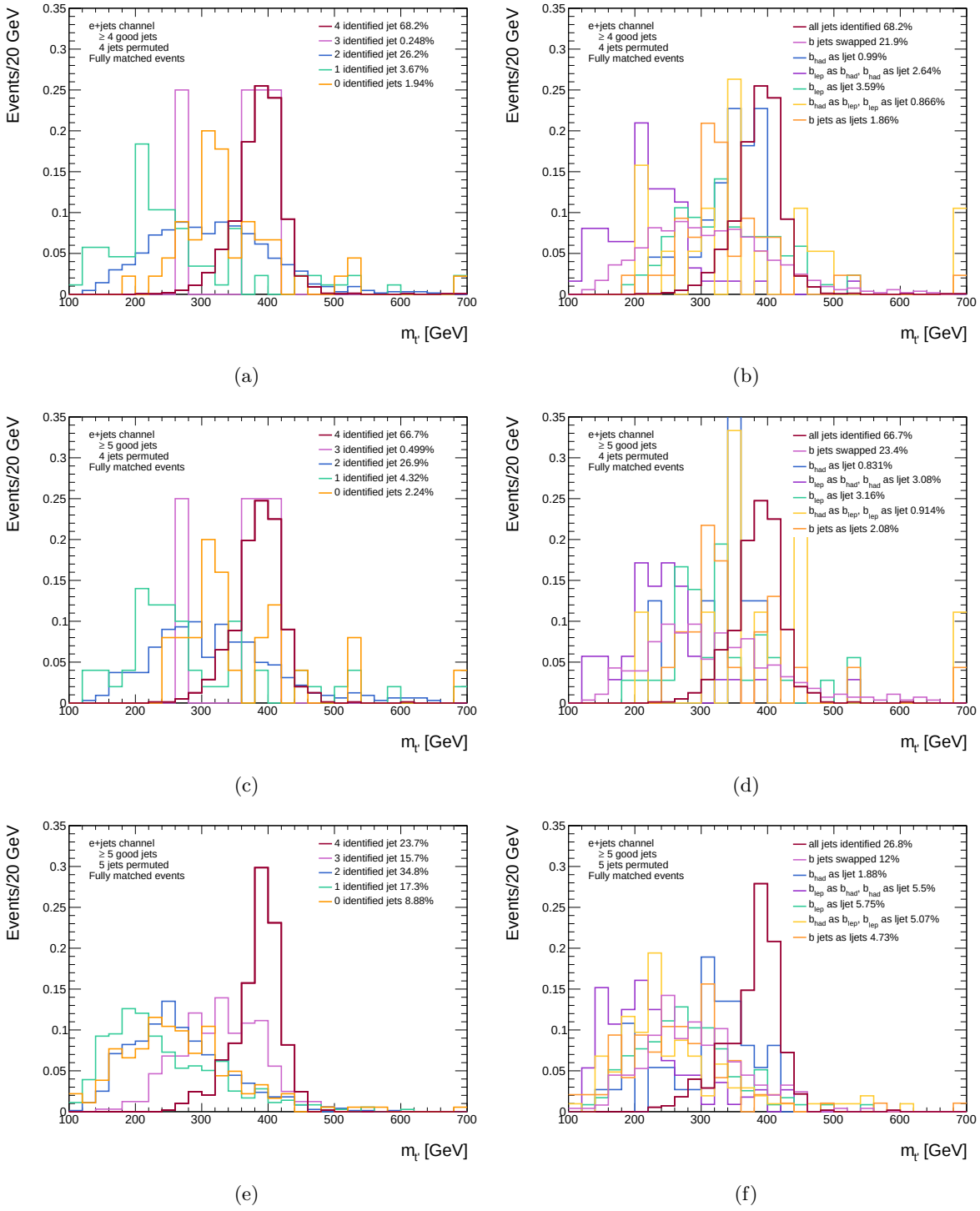


Figure 8.5: Shapes of the various components of the fitted t' mass distribution of the fully matched events decomposed according to: the number of identified jets ((a), (c), and (e)) and nature of the mis-identification ((b), (d), and (f)). All histograms are normalised to unity. The distributions are shown for the samples with $\geqslant 4$ ((a), (b)) and $\geqslant 5$ ((c) – (f)) reconstructed jets in the event in the e+jets channel. The distributions have been produced by fitting four ((a) – (d)) or five ((e), (f)) leading jets in the nominal t' sample with $m_{t'} = 400$ GeV.

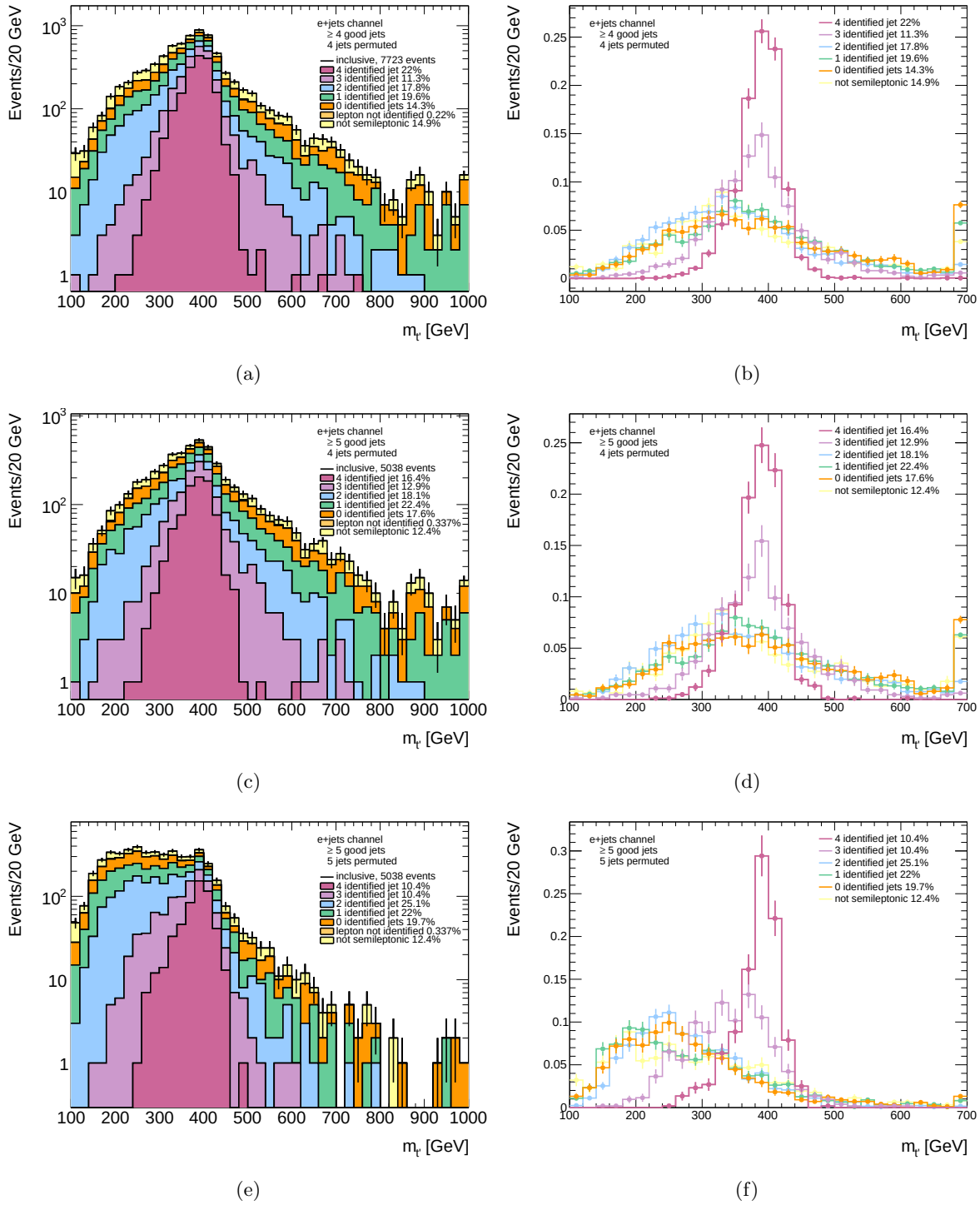


Figure 8.6: Distribution of the fitted t'-quark mass decomposed according to the number of identified reconstructed jets shown in a stacked plot (a, c, e) and as individual histograms normalised to unity. The distributions are shown for the samples with $\geqslant 4$ (a and b) and $\geqslant 5$ (c – f) reconstructed jets in the event in the e+jets channel. The distributions have been produced by fitting four (a – d) or five (e and f) leading jets in the nominal t' sample with $m_{t'} = 400$ GeV.

Configuration	4 jets identified	3 jets identified	2 jets identified	1 jet identified	0 jets identified	lep. not matched	evt. not matched
			e+jets				
$\geqslant$ 4 good, 4 fit	22.6%	42.0%	16.9%	2.64%	0.15%	0.43%	15.3%
$\geqslant$ 5 good, 4 fit	13.3%	47.4%	22.5%	3.7%	0.24%	0.52%	12.3%
$\geqslant$ 5 good, 5 fit	35.7%	37.4%	12.1%	1.8%	0.09%	0.52%	12.3%
			μ+jets				
$\geqslant$ 4 good, 4 fit	24.0%	42.6%	16.5%	2.4%	0.14%	0.49%	13.8%
$\geqslant$ 5 good, 4 fit	13.8%	48.7%	22.2%	3.5%	0.20%	0.60%	11.2%
$\geqslant$ 5 good, 5 fit	37.0%	37.7%	11.7%	1.6%	0.09%	0.59%	11.1%

Table 8.11: Truth matching summary for e+jets and μ+jets channels for the nominal t' sample with $m_{t'} = 400$ GeV. The fraction of events for each number of matched jets bin is shown.

Configuration	all jets identified	b jets swapped	b_{had} as light	b_{lap} as b_{had} b_{had} as light	b_{lep} as light	b_{had} as b_{lep} b_{lep} as light	both b-jets as light
			e+jets				
$\geqslant$ 4 good, 4 fit	68.2%	21.9%	1.0%	2.6%	3.6%	0.9%	1.9%
$\geqslant$ 5 good, 4 fit	66.7%	23.4%	0.8%	3.1%	3.2%	0.9%	2.1%
$\geqslant$ 5 good, 5 fit	26.8%	12.0%	1.9%	5.5%	5.8%	5.1%	4.7%
			μ+jets				
$\geqslant$ 4 good, 4 fit	70.1%	20.5%	0.8%	2.4%	3.6%	0.8%	2.1%
$\geqslant$ 5 good, 4 fit	67.2%	23.2%	1.1%	2.5%	3.5%	0.8%	2.2%
$\geqslant$ 5 good, 5 fit	27.7%	11.6%	2.0%	5.1%	4.9%	5.0%	4.0%

Table 8.12: Nature of the mismatches measured on the fully matched events for e+jets and μ+jets channels for the nominal t' sample with $m_{t'} = 400$ GeV.

Configuration	all jets identified	b jets swapped	b_{had} as light	b_{lap} as b_{had} b_{had} as light	b_{lep} as light	b_{had} as b_{lep} b_{lep} as light	both b-jets as light
			e+jets				
$\geqslant$ 4 good, 4 fit	47.9%	8.2%	22.3%	5.1%	5.3%	6.2%	4.2%
$\geqslant$ 5 good, 4 fit	48.3%	7.5%	23.8%	4.6%	4.5%	5.9%	3.8%
$\geqslant$ 5 good, 5 fit	26.9%	4.3%	17.8%	4.9%	4.3%	4.6%	4.6%
			μ+jets				
$\geqslant$ 4 good, 4 fit	46.9%	8.7%	22.1%	5.4%	5.2%	6.5%	4.4%
$\geqslant$ 5 good, 4 fit	47.3%	7.8%	23.5%	5.1%	4.7%	5.9%	4.1%
$\geqslant$ 5 good, 5 fit	25.7%	4.6%	17.0%	5.3%	4.5%	5.0%	4.7%

Table 8.13: Nature of the mismatches measured on the fully matched events for e+jets and μ+jets channels for the nominal $t\bar{t}$ sample.

distribution produced using five leading p_{T} jets peaks at approximately ~ 300 GeV with a width of ~ 220 GeV, resulting in a reduced high tail for the t' signal sample as well as the dominant $t\bar{t}$ background. The broadening and lowering of the the $m_{t'}$ spectrum in the latter case occurs due to an increased rate of jet mis-identifications by the KLFitter. Although the latter case provides a better controlled high tail of the $t\bar{t}$ background, the background contribution is still ~ 40 times higher than the expected signal contribution at its mass peak. When systematic error bands on the SM backgrounds expectation are included, such a broad peak is covered by the uncertainty range. A narrow peak produced by permuting only the four hardest jets of the event has a better separation from the background, and therefore leads to a better sensitivity of the analysis based on the fitted $m_{t'}$ variable. Consequently, it was decided to use only the hardest four jets for the event reconstruction performed using kinematic fitting.

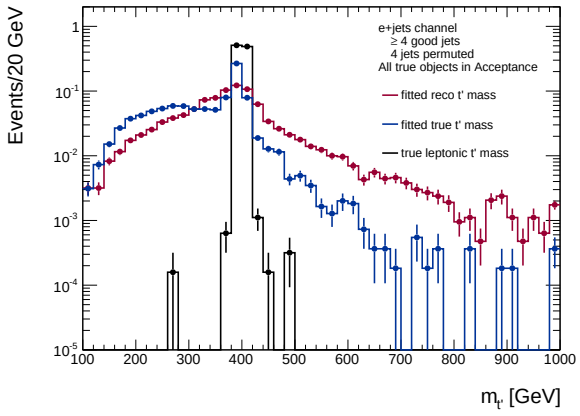
8.3.1.2 Fitting the t'-quark Mass at the Generator Level

The jet energy resolution, within which the measured jet energies are allowed to vary in the fit, is described by a set of transfer functions explained in Section 6.1.1.3. To entangle the effect of these jet energy variations governed by the transfer functions from the reconstruction effects, the t'-quark mass was fitted using the generator level objects in the events in which all objects of interest were within the detector acceptance. This generator level fitted mass distribution was compared to the reconstructed level fitted mass distribution, as well as the parton level mass of the leptonically decaying heavy quark, reconstructed as the invariant mass of the object composed of the generated level b-quark, lepton and neutrino. This test has been performed with the nominal t' sample as well as with the $t\bar{t}$ sample using usual KLFitter configurations: $\geqslant 4$ or $\geqslant 5$ reconstructed jets per event, and four or five leading p_{T} jets used in permutations (Figure 8.7).

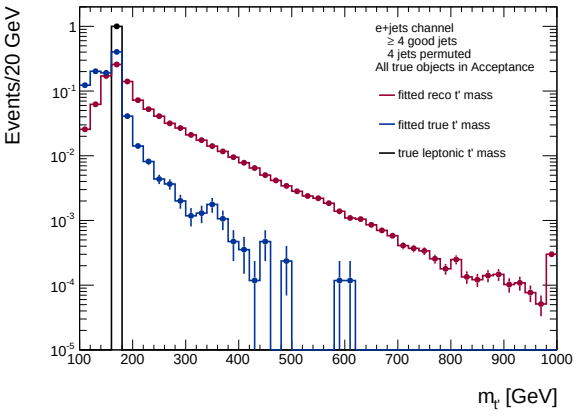
It turns out that the mass fitted from the generator level information exhibits the usual broadening of the distribution, and enhancement at low masses. Furthermore, for events with $\geqslant 5$ jets, the generator level fitted mass is more alike the reconstructed level mass fitted using five leading jets than that fitted with four leading jets. This suggests that these jet permutations leading to the low masses are indeed preferred by the fit, i.e. that the bias towards the lower masses results from the intrinsic KLFitter features (the likelihood function definition and the transfer functions used) rather than the reconstruction effects. To understand in which way the variation of jet energy due to the fitting procedure contributes to this enhancement, a study of jet energy resolution has been undertaken, as reported in Section 8.3.5.

8.3.1.3 Log Likelihood Function Correlations

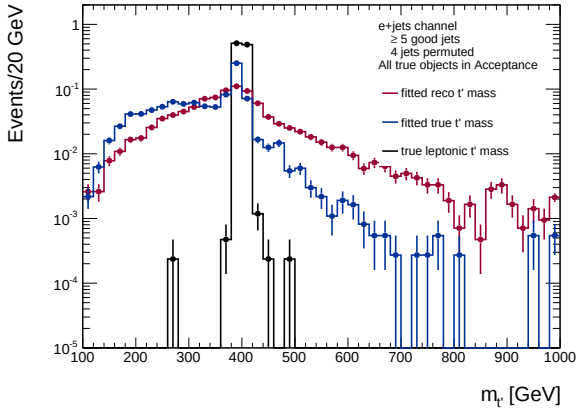
To get a closer insight into the decision process of the KLFitter, the correlations of its metric, the log likelihood function, and the representative observables of the fitted final state have been studied. Figures 8.8 (a) and (b) show the correlation of the log likelihood function with the fitted t'-quark mass for the best permutation (i.e. that with the highest likelihood), for events having $\geqslant 5$ jets per event. The plots have been produced by permuting four (8.8 (a), (c), and (e)) and five (8.8 (b), (d), and (f)) hardest jets of the event. Both sets of plots show a slight anti-correlation between the log likelihood and the fitted t'-quark mass. The plots suggest that the most populated regions are those with high log likelihood values. In the case of five permuted jets, a high fraction of events with a high log likelihood corresponds to the masses



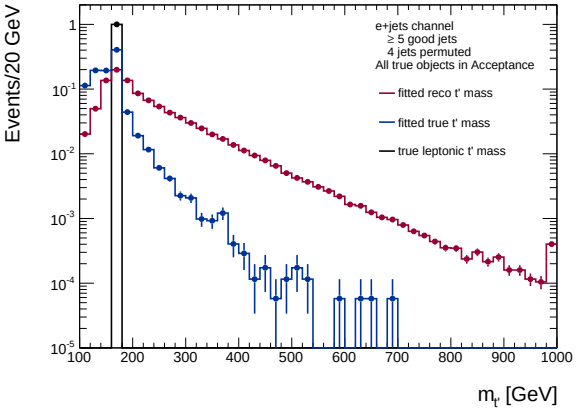
(a)



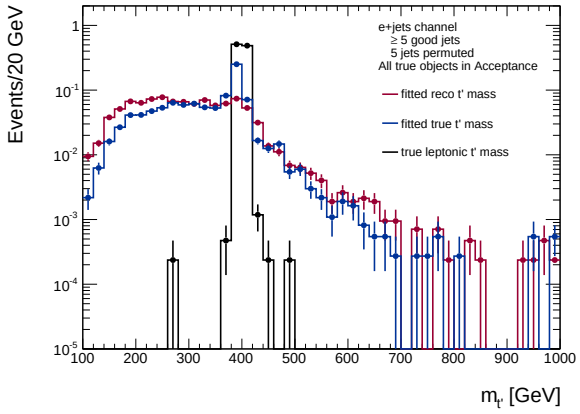
(b)



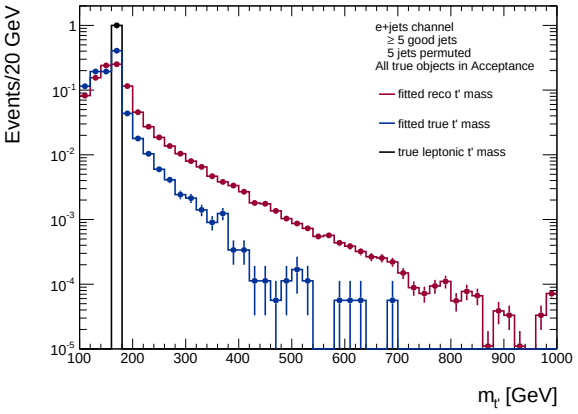
(c)



(d)



(e)



(f)

Figure 8.7: Distribution of the t' mass $m_{t'}$ fitted using the KLFitter with the reconstructed (red line) and the generator level (blue line) input, as well as the generator level invariant mass of the leptonically decaying t'-quark (black line). The distributions produced with the simulated t' ((a), (c), and (a)) and $t\bar{t}$ ((b), (d), and (f)) events are shown for the sample with $\geqslant 4$ selected jets ((a) and (b)) as well as with $\geqslant 5$ selected jets, using four ((c) and (d)) or five ((e) and (f)) leading p_{T}jets for constructing the permutations.

between 150 and 400 GeV. This shows that the bias towards low masses in case of five permuted jets does not correspond to poorly reconstructed events with a low likelihood, but are rather preferred by the fit. This implies that the KLFitter definition of the likelihood is not well suited to correctly reconstruct the heavy quark mass in the case when it is fitted as a free parameter. The fact that the most frequent mis-reconstruction scenario is the b-jets swap, suggests that the introduction of a spatial constraint (e.g. a space angle within which the decay products of a t'-quark need to be situated) into the definition of the likelihood function could be beneficial.

Figures 8.8 (c) and (d) further show a slight correlation between the number of correctly identified jets and the log likelihood of the best permutation. The identification multiplicity bin marked by the value -1 corresponds to the events with no semi-leptonic signature at the generator level. While the most preferred region in the case of four permuted leading jets corresponds to four identified jets, in the case of five permuted jets the high likelihood region is spread over several identification multiplicity bins with a pronounced maximum in the bin corresponding to two identified jets. Figures 8.8 (e) and (f) demonstrate that in case of five permuted jets the majority of events with less than four identified jets contributes to the mass range $150 - 350$ GeV, in which the fitted t' mass distribution exhibits a plateau.

8.3.1.4 KLFitter Mass Resolution

The fitted mass difference between a t' and the corresponding $\bar{t}'$-quark of the event, i.e the difference between the fitted mass of the hadronically and leptonically decaying quark in the event, is shown as a function of the corresponding log likelihood value (Figure 8.9) and the fitted t'-quark mass form the hadronic side of the decay (Figure 8.10). These correlation plots have been produced using the preselected event sample with $\geqslant 5$ jets of the nominal t' sample, and with a varied number of permuted jets: four and five leading p_{T} jets. The resulting mass differences turn out to be very small, of order of tens of MeV, except for few events with the mass difference of several hundreds of GeV.

Figure 8.9 shows that in case of four permuted jets, the mass differences for the high log likelihood values reach up to $\sim$ 50 MeV, while for the lower log likelihood values the mass differences shrink below 20 MeV. In case of five permuted jets the mass differences are slightly higher, reaching up to 80 MeV for high log likelihood region. In both cases the majority of events have a mass difference up to 20 MeV.

The plots in the upper row of Figure 8.10 show that the high mass differences correspond to low fitted t'-quark mass. The mass resolution in the mass region of interest (around 400 GeV) is around 30 MeV. The bottom row plots show that, while in the case of five permuted jets the mass differences are mostly very low, about 7% of events reconstructed by permuting four leading jets have very high mass difference (up to 300 GeV), reaching in few cases as high as 600 GeV. This shows that the fit reaches a better mass agreement between the two sides of the decay if it is provided five leading jets, as softer jets can more easily be combined into two objects with similar masses than harder jets.

8.3.2 Top Mass Study

As this analysis is a search for a hypothetical particle of an unknown mass, the t'-quark mass cannot be fixed, but rather fitted for, as explained in Section 6.1.1.1. However, the mass of the

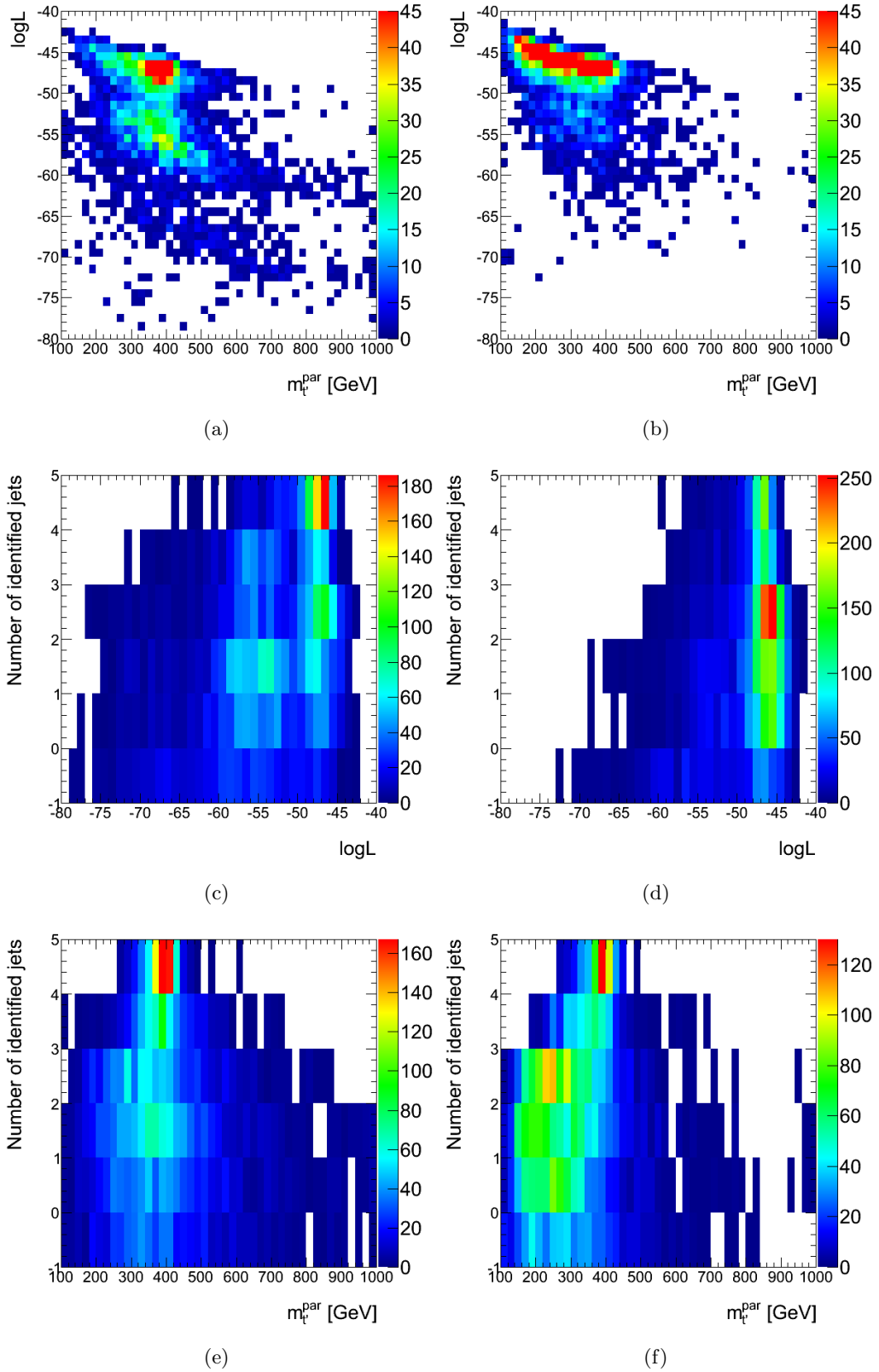


Figure 8.8: The KLFitter log likelihood of the best permutation as a function of the fitted t' mass ((a) and (b)), the number of correctly identified jets as a function of the KLFitter log likelihood value ((c) and (d)) and the fitted t' mass ((e) and (f)). All figures are shown for the nominal t' sample with $m_{t'} = 400$ GeV in the e+jets channel, in case of four ((a), (c), and (e)) and five ((b), (d), and (f)) hardest jets used in the fit.

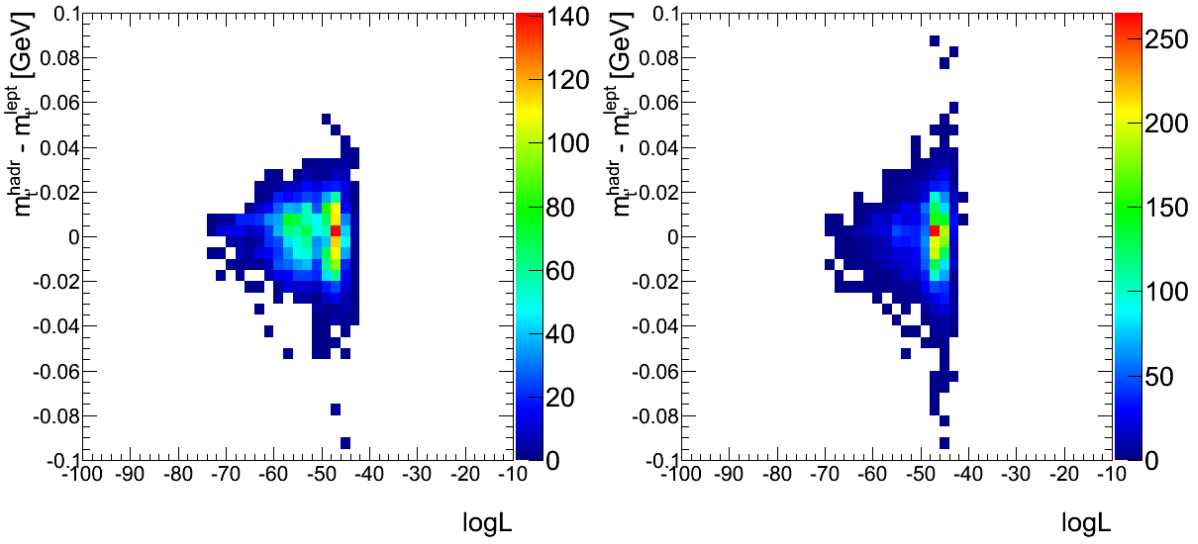


Figure 8.9: The mass difference between the heavy quark and its anti-partner reconstructed by the KLFitter as a function of the log likelihood value for the best permutation for the preselected events of the nominal t' sample with $m_{t'} = 400$ GeV in the e+jets channel, in case of four (left) and five (right) hardest jets used in the fit.

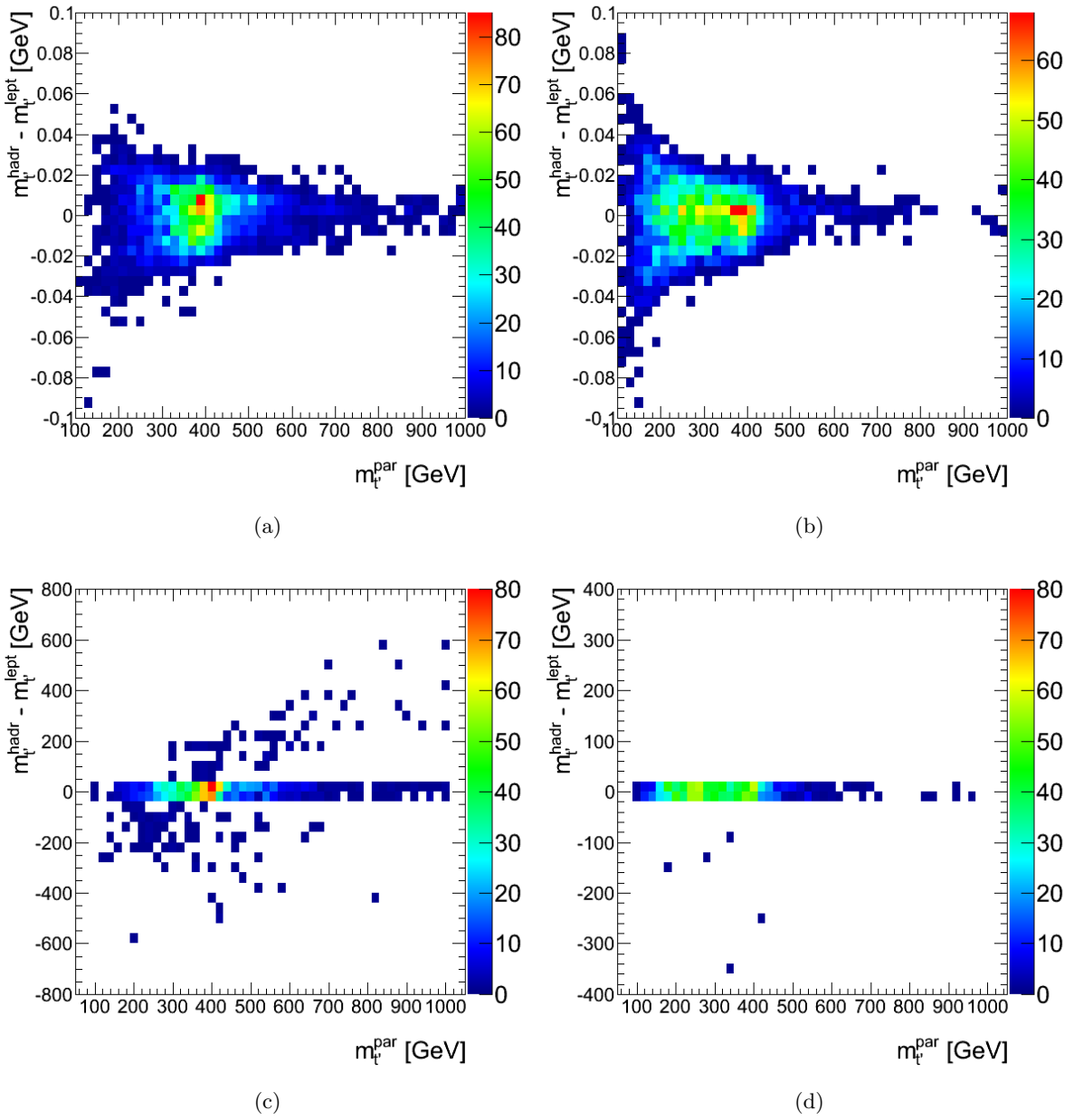
target quark enters the central likelihood function (Equation 6.1) indirectly through the quark width, present in the likelihood term constraining the quark and anti-quark to have compatible masses. The KLFitter performance was tested with the quark mass set to the nominal top quark mass of 172.5 GeV, as well as the benchmark t'-quark mass of 400 GeV. Since the final observables have not shown a noticeable sensitivity to a modification of the quark mass parameter in the top width calculation, it has been decided to keep the ATLAS-wide nominal value of 172.5 GeV.

8.3.3 B-tagging Option Study

KLFitter has three options for treating the per-jet b-tagging information:

- no per-jet b-tagging information used (the 'kNotag' option)
- the permutations where a b-tagged jet is fitted to be a light jet or vice-versa are rejected (the 'kVeto' option)
- permutations are assigned a probability calculated using the per-jet b-tagging efficiency and rejection weight (the 'kWorkingPoint' option).

Figure 8.11 shows the fitted mass $m_{t'}$ produced from preselected simulated t' event using three different b-tagging options in the fit. The plots show that, while the appliance of the 'kNotag' and the 'kWorkingPoint' options provide very similar results, the 'kVeto' option brings to a slight increases of the high tail of the fitted t' mass distribution. Between the two options with the similar performance, that with the lower computational demand and less affected by the systematic uncertainties has been chosen, i.e. it has been decided not to make use of the per jet b-tagging information in the fit.



(a) (b)

(c) (d)

Figure 8.10: The mass difference between the heavy quark and its anti-partner reconstructed by the KLFitter as a function of the fitted t'-quark mass for the preselected events of the nominal t' sample with $m_{t'} = 400$ GeV in the e+jets channel, in case of four ((a) and (c)) and five ((b) and (d)) hardest jets used in the fit. The top row plots show the most populated region while the bottom row plots show a larger range of the mass difference.

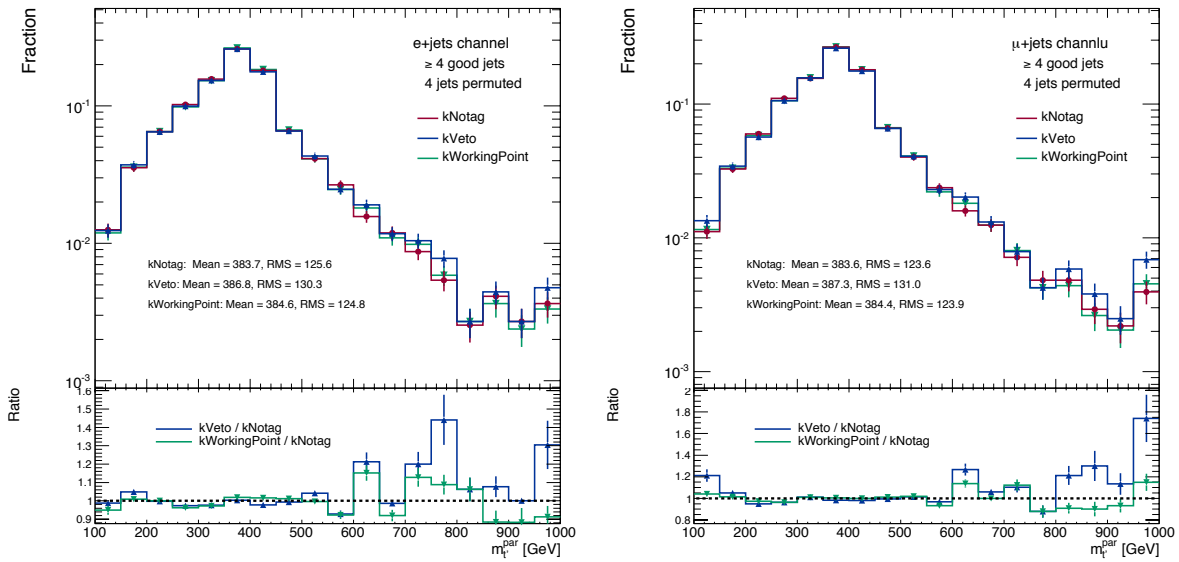


Figure 8.11: Expected reconstructed mass of the t' quark (with a mass of 400 GeV) obtained from KLFitter using one of three possible b-tagging options: kNotag (red), kVeto (blue) and kWorkingPoint (green). The left plot shows the e+jets channel, the right plot shows the μ+jets channel. The sub-plots show the ratio with respect to the kNotag option.

8.3.4 Effect of ISR and FSR

The effect of the increased and decreased amount of initial and final state radiation on the overall KLFitter identification efficiency has been studied for the t' sample assuming $m_{t'} = 500$ GeV. Plots in Figure 8.12 as well as the numbers in Tables 8.14 – 8.15 show that increased and decreased radiation rate do not affect either the shapes of the $m_{t'}$ distribution or the object identification efficiency significantly. While the shapes of the distributions vary only slightly, the identification efficiencies vary within a few percent.

8.3.5 Energy Resolution Studies

The jet energy resolution is a quantity which enters the KLFitter likelihood and is fitted within the resolution given by the transfer functions explained in Section 6.1.1.3. In this section, the jet energy emerging from the fit has been compared to the jet energy which enters the fit, to understand the magnitude of the jet energy modification in the fitting procedure.

Figure 8.13 shows the difference between the reconstructed and fitted energy of the jets identified as particular partons. The distributions are shown as stacked plots divided into components regarding the number of correctly identified jets per event. The b-jet distributions (Figures 8.13 (a) and (b)) have quite an asymmetric shape. In addition to an approximately Gaussian core, the distributions have a pronounced enhancement on the negative side corresponding to the events in which the fitted energy exceeds the reconstructed energy. This enhancement is dominated by the events including misidentified objects, although it is also present in events with all jets properly identified. The light jet resolutions (Figures 8.13 (c) and (d)) have a shape which is closer to a double Gaussian, although also in this case the fit tends to vary the

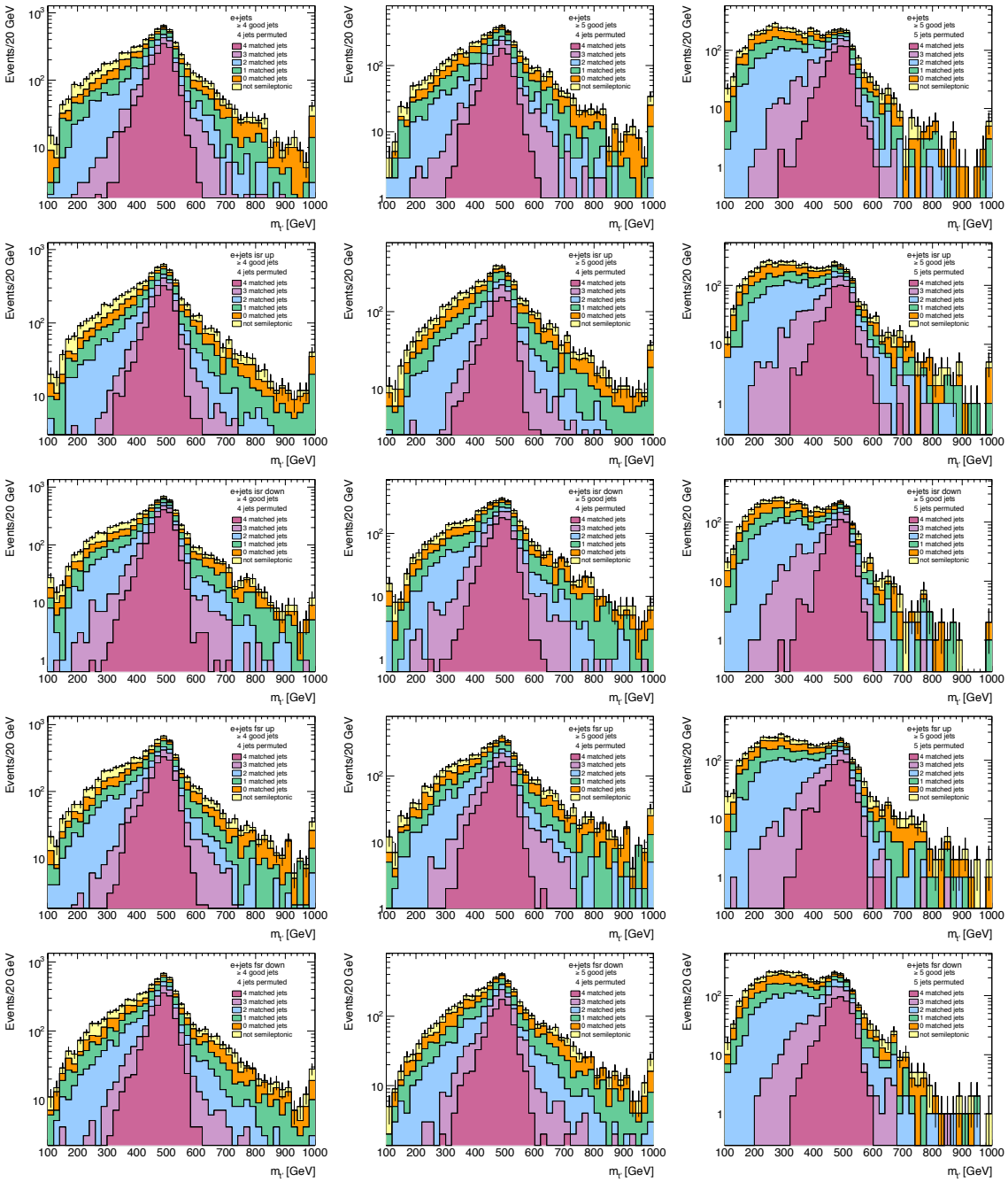


Figure 8.12: Distribution of the t' mass fitted by permuting four leading jets in events with $\geqslant 4$ (left column) and $\geqslant 5$ (middle column), or five leading jets in events with $\geqslant 5$ (right column). The distribution is shown for the nominal t' samples with the mass of 500 GeV (top row) as well as the samples with varied ISR and FSR. From top to bottom: more (second row) and less (third row) ISR, and more (fourth row) and less (bottom row) FSR. The distributions are decomposed in various levels of truth matches.

Configuration	4 jets matched	3 jets matched	2 jets matched	1 jet matched	0 jets matched	event not matched	total events
			$\geqslant$ 4 good, 4 fit				
nominal	23.3%	10.5%	16.5%	20.0%	14.3%	15.5%	6631
ISR up	21.8%	10.5%	16.8%	19.9%	14.7%	16.3%	6603
ISR down	26.4%	10.7%	16.4%	18.1%	13.4%	15.0%	6519
FSR up	22.9%	10.9%	16.8%	19.1%	14.5%	15.8%	6663
FSR down	23.2%	10.6%	17.0%	19.3%	14.9%	14.9%	6647
			$\geqslant$ 5 good, 4 fit				
nominal	18.9%	12.2%	16.2%	22.7%	16.9%	13.0%	4445
ISR up	16.7%	11.6%	17.8%	22.7%	17.8%	13.5%	4485
ISR down	20.0%	12.5%	17.1%	21.4%	16.7%	12.3%	4050
FSR up	18.0%	12.4%	16.9%	21.8%	17.6%	13.2%	4312
FSR down	18.4%	12.1%	17.6%	21.7%	18.3%	12.0%	4387
			$\geqslant$ 5 good, 5 fit				
nominal	12.0%	10.0%	25.5%	19.9%	19.6%	13.0%	4445
ISR up	10.6%	10.3%	25.5%	19.7%	20.3%	13.5%	4485
ISR down	11.5%	9.4%	27.1%	19.6%	20.1%	12.3%	4050
FSR up	10.8%	9.6%	26.3%	20.6%	19.4%	13.2%	4312
FSR down	11.0%	10.4%	26.4%	20.2%	20.2%	12.0%	4387

Table 8.14: KLFitter identification efficiency measured on all events with $\geqslant$ 4 jets for the t' sample with $m_{t'} = 500$ GeV in the e+jets channels.

jet energy upwards more often than downwards.

Figure 8.14 shows the difference between the fitted and generator level jet energy as a function of the generator level jet energy. The left hand side plots correspond to the b-jet from the leptonic side of the decay, and the right hand side to the light jet with a higher p_{T}. The top row plots comprise all preselected events with $\geqslant$ 4 jets in the e+jets channel. The middle row plots show only those events for which the jet whose energy is plotted is correctly identified by the KLFitter. The bottom row plots correspond to the events with four successful jet identifications. The approximately triangular cutoff in the shape of the function $(E_{\mathrm{fit}} - E_{\mathrm{true}}) = -E_{\mathrm{fit}}$ obvious in the top row plots depicts the fact that E_{fit} is always greater than zero. The dominant structure in all plots is the rather Gaussian distribution around $(E_{\mathrm{fit}} - E_{\mathrm{true}}) = 0$. The plots also show that in the low true energy region (i.e. for $E_{\mathrm{true}} <\sim 300$ GeV), the fitted jet energy is more often higher than the true jet energy, whereas in the high true energy range the situation is reversed. While for b-jets this feature almost vanishes for a high jet identification efficiency (bottom left plot of Figure 8.14), for light jets it is always visible.

Figure 8.15 shows the correlation between the fitted (the fit output) and reconstructed (the fit input) energy of the b-jet from the leptonic side of the decay (left hand side plots) and the harder light jet (right hand side plots) in the same three cases as in the previous figure. The plots show that the majority of jets are distributed around the $E_{\mathrm{fit}} = E_{\mathrm{true}}$ line in a Gaussian manner. However, for the b-jet there is a smaller pronounced component approximately distributed along the line $E_{\mathrm{fit}} = 1.2 \cdot E_{\mathrm{true}}$. This component is present in the light jet plots as well, to a lower extent. In this case, another component, in which E_{fit} is slightly lower than E_{reco} (approximately along

Configuration	4 jets matched	3 jets matched	2 jets matched	1 jet matched	0 jets matched	event not matched	total events
			$\geqslant$ 4 good, 4 fit				
nominal	24.2%	10.3%	17.6%	19.7%	14.4%	13.8%	6638
ISR up	23.2%	10.8%	16.6%	19.7%	14.7%	15.0%	6629
ISR down	25.8%	10.4%	18.0%	19.4%	13.7%	12.7%	6627
FSR up	24.2%	11.1%	16.5%	20.9%	14.0%	13.3%	6625
FSR down	23.9%	10.5%	17.8%	20.0%	14.5%	13.4%	6524
			$\geqslant$ 5 good, 4 fit				
nominal	19.5%	12.2%	17.9%	22.3%	17.0%	11.1%	4439
ISR up	19.1%	12.2%	16.7%	22.5%	17.2%	12.4%	4533
ISR down	20.2%	12.3%	17.7%	22.5%	17.0%	10.4%	4208
FSR up	18.9%	12.7%	16.0%	24.1%	17.0%	11.3%	4328
FSR down	19.0%	11.8%	17.6%	23.3%	17.7%	10.7%	4300
			$\geqslant$ 5 good, 5 fit				
nominal	10.3%	10.0%	29.1%	19.3%	20.3%	11.1%	4439
ISR up	10.6%	10.2%	26.2%	20.6%	19.9%	12.4%	4533
ISR down	11.9%	10.1%	27.5%	20.1%	20.0%	10.4%	4208
FSR up	10.9%	10.0%	27.1%	20.4%	20.3%	11.3%	4328
FSR down	11.5%	10.0%	27.5%	20.2%	20.0%	10.7%	4300

Table 8.15: KLFitter identification efficiency measured on all events with $\geqslant$ 4 jets for the t' sample with $m_{t'} = 500$ GeV in the μ+jets channels.

the line $E_{\text{fit}} = 0.7 \cdot E_{\text{true}}$) is stronger. These off-diagonal components likely depict the maxima of the underlying transfer functions, which have double-Gaussian shape (Section 6.1.1.3). These plots suggest, that in case of the low jet identification efficiency, the energy of b-jets is rather pulled upward by the fit, while that of light jets is slightly more often pulled downward.

The findings shown in Figures 8.14 and 8.15 for one b-jet (light jet) apply for another jet of that kind in the event as well.

8.3.6 Conclusion of the KLFitter Performance Studies

The studies presented in Section 8.3.1 show that the optimal fit performance is achieved when four highest jets in the event are used in the fit. Although the probability of including all jets originating from the hard scattering is increased when the fifth leading p_{T} jet is provided to the fit, the resulting reconstructed mass spectrum is broadened and biased toward lower values in this case, as a consequence of the likelihood function and the transfer functions used. It has further been found that the heavy quark particle width corresponding to the mass of 172.5 GeV (Section 8.3.2), as well as no usage of per-jet b-tagging information (Section 8.3.3) are beneficial for the fit performance. In addition, the shape of the fitted mass distribution has been found to be rather stable under the ISR/FSR variations 8.3.4. The study of the jet energy variation presented in Section 8.3.5 has shown that the fit tends to scale the jet energy up, especially in cases of the permutations with a lower jet identification efficiency.

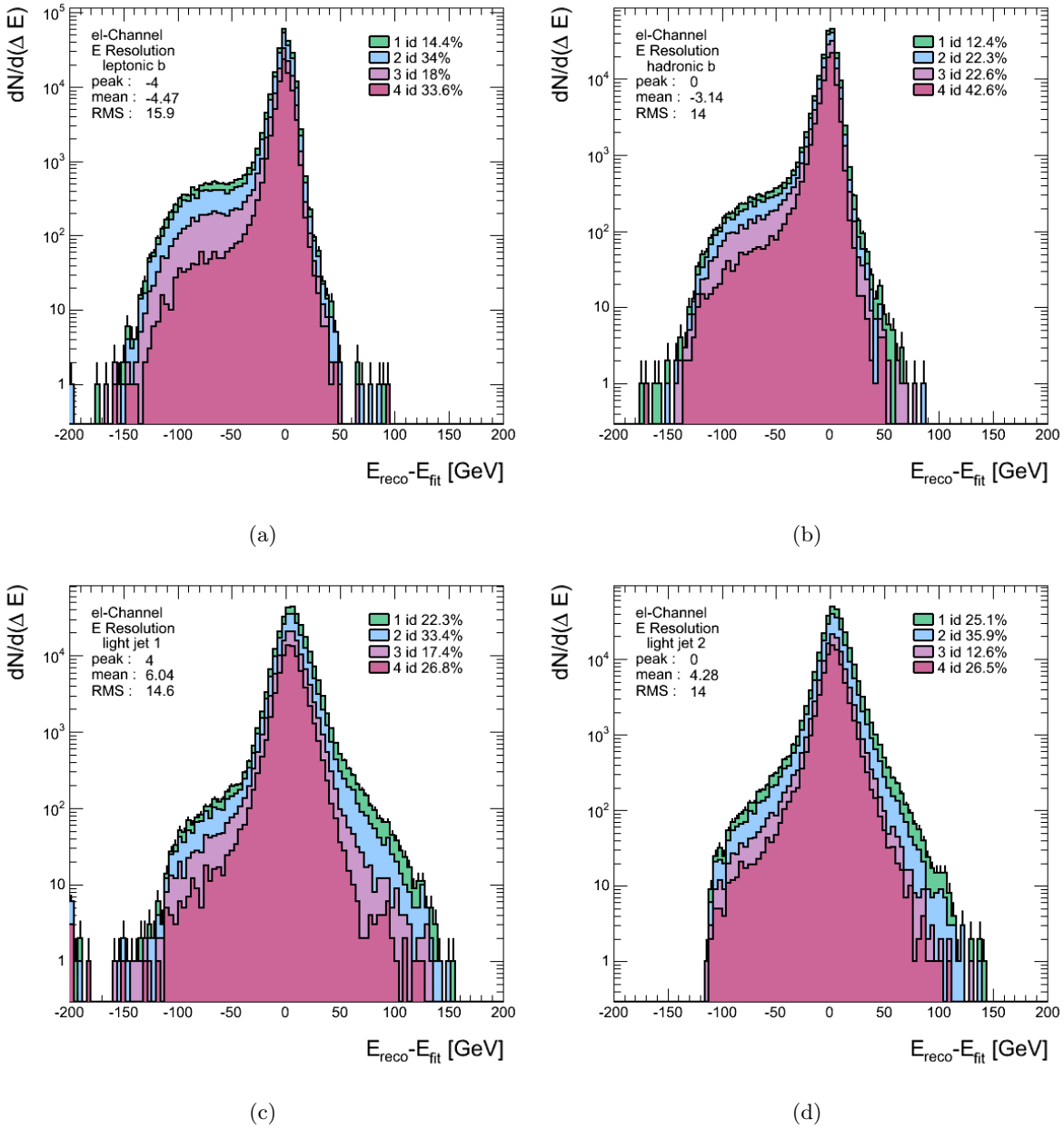


Figure 8.13: Difference between the fitted and true energy of the b-jet from the leptonic (a) and hadronic (b) side of the decay, as well as of the light jets (c and d) decomposed according to the number of identified jets in the events for the preselected events with $\geqslant 4$ jets of the $t\bar{t}$ sample in the e+jets channel.

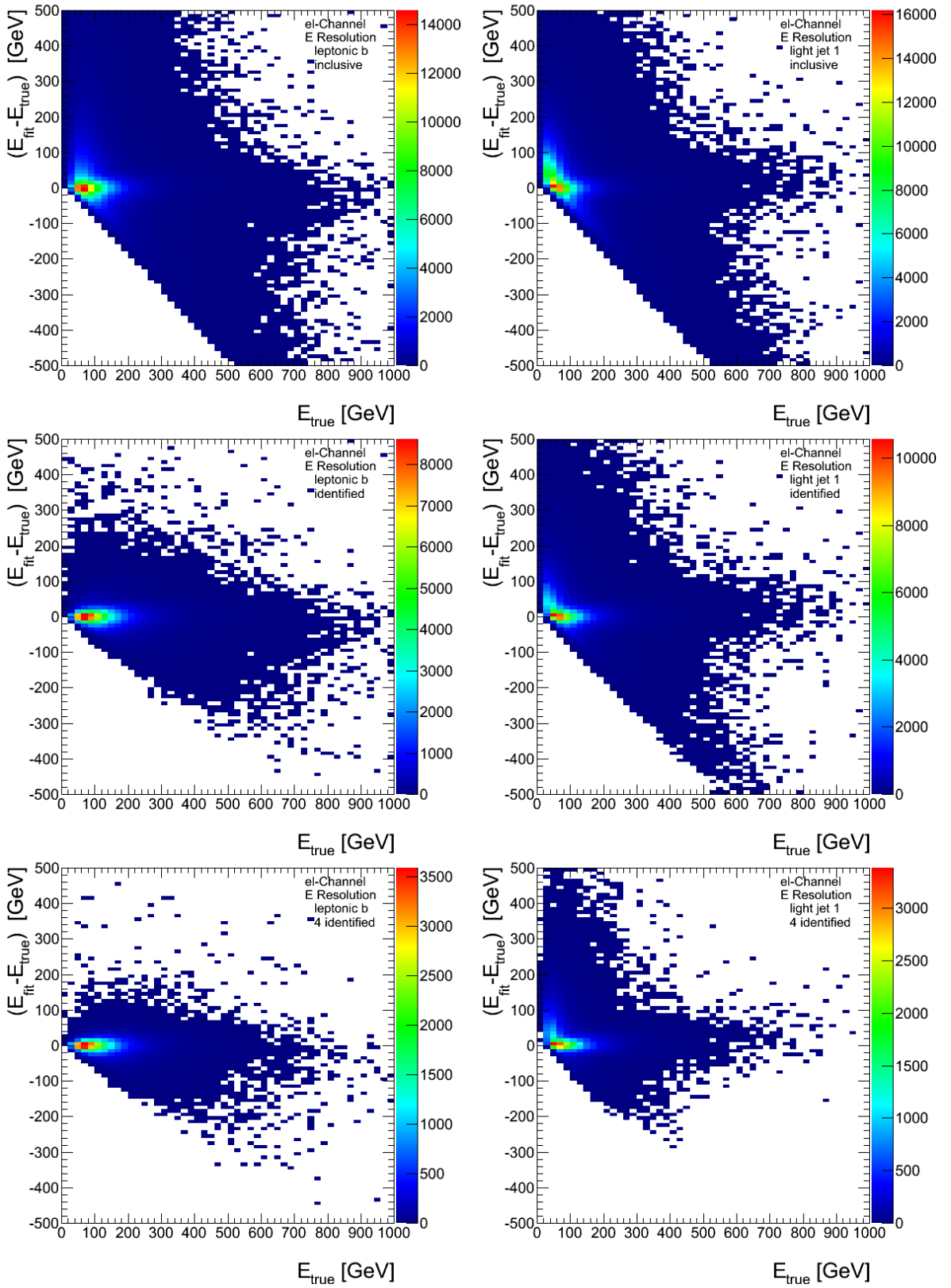


Figure 8.14: Difference between the fitted and true energy as a function of the true energy of the b-jet from the leptonic side of the decay (left), as well as one light jet (right) for the preselected events with $\geqslant 4$ jets of the $t\bar{t}$ sample in the e+jets channel, in three cases: inclusive (top row), the jet in question is identified by the KLFitter (middle row), all four jets are identified (bottom row).

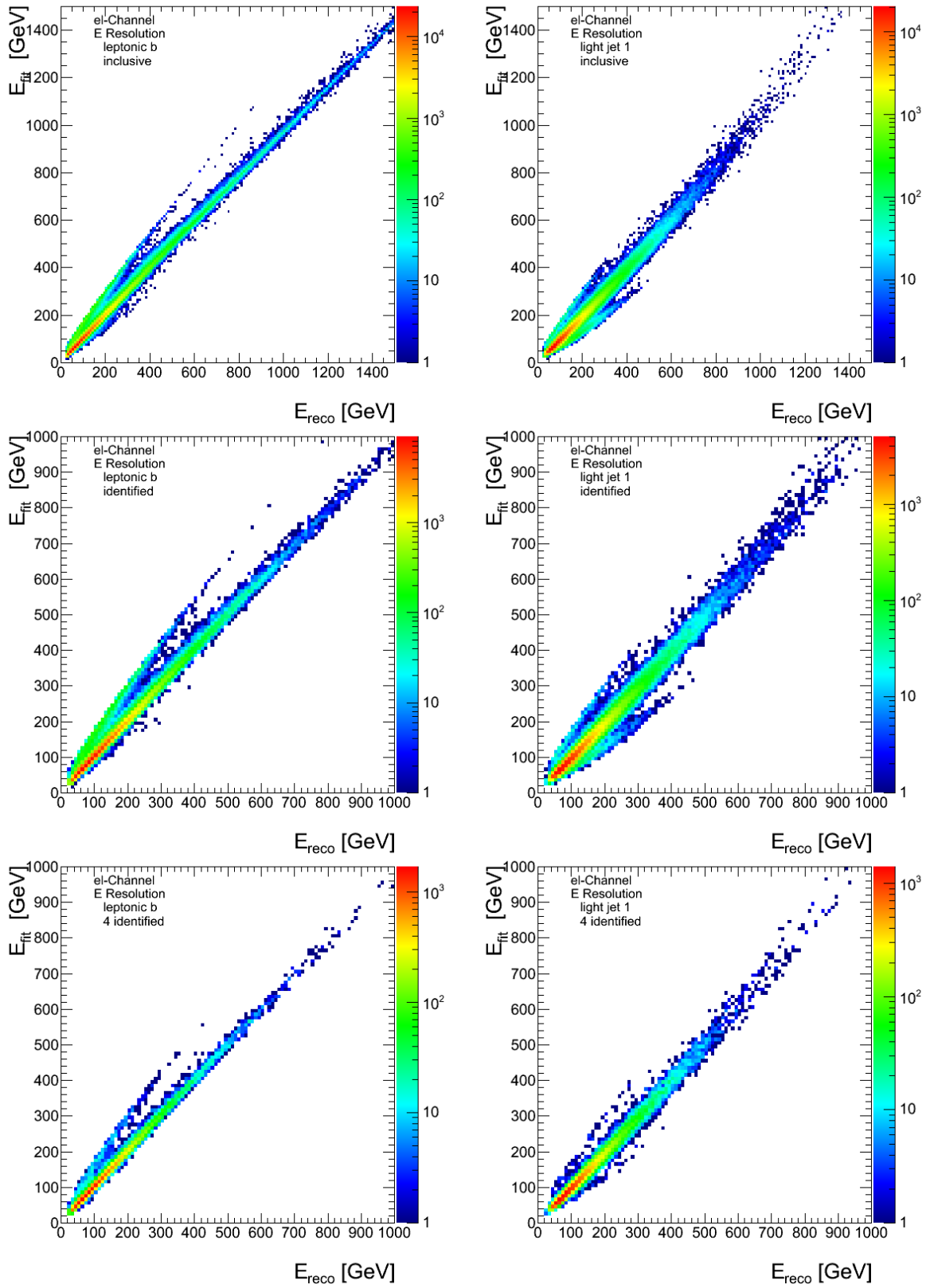


Figure 8.15: Correlation between the fitted and reconstructed energy of the b-jet from the leptonic side of the decay (left), as well as one light jet (right) for the preselected events with $\geqslant 4$ jets of the $t\bar{t}$ sample in the e+jets channel, in three cases: inclusive (top row), the jet in question is identified by the KLFitter (middle row), all four jets are identified (bottom row).

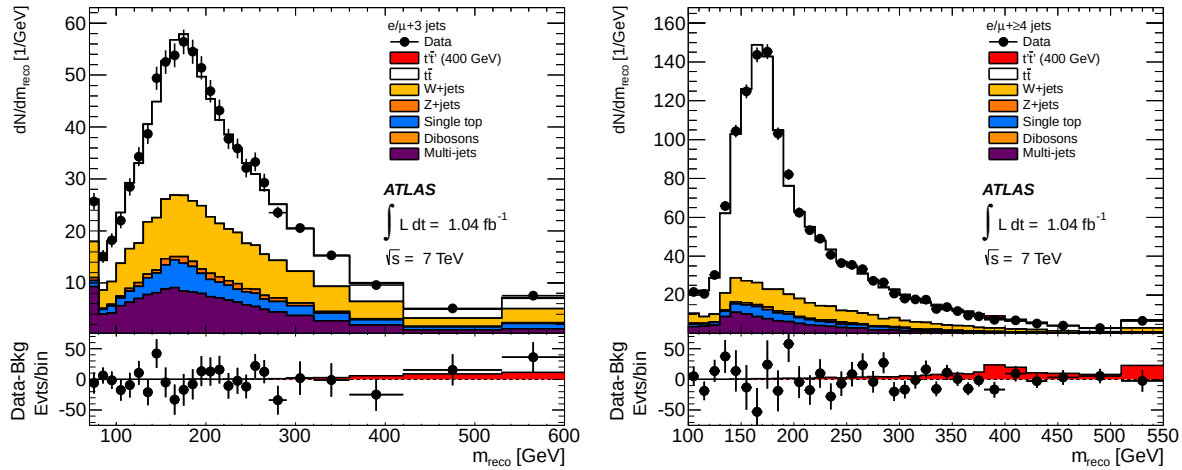


Figure 8.16: Observed (black points) and expected (stacked histograms) distributions of the reconstructed t'-quark mass in the $e/\mu + 3$ jets (left) and the $e/\mu+ \geqslant 4$ jets (right) channels. While the top panels show the absolute bin contents, the bottom panels show the absolute difference between the data and the expected backgrounds. The expected contribution from a signal with t' mass of 400 GeV is added to the stacked distributions. Taken from [205].

8.4 Statistical Interpretation and Final Results

The statistical analysis was performed on the full range of the reconstructed t' mass distribution in the event samples with $= 3$ and $\geqslant 4$ jets (Figure 8.16) using the log-likelihood ratio $LLR = -2\log(L_{\mathrm{s+b}}/L_{\mathrm{b}})$ explained in Section 6.3.1 as test-statistic. The bin-wise expectations have been parametrised in terms of 12 nuisance parameters, modelling the most pronounced sources of systematic uncertainties: jet energy scale, ISR/FSR estimation, as well as the normalisations of the $t\bar{t}$, W+jets and QCD multi-jet backgrounds. The systematic uncertainties are estimated as explained further below in Section 9.6. Both $L_{\mathrm{s+b}}$ and L_{b} have been maximised with respect to these nuisance parameters to reduce the effect of the systematic uncertainties on the sensitivity of the search. All nuisance parameters are subject to Gaussian constraints of their priors. Some of the systematic effects have been constrained through the simultaneous treatment of the 3-jets and $\geqslant$ 4-jets channels. No significant excess of data events over the SM expectation has been observed in this search, resulting in an upper limit placed on the $t'\bar{t}'$ production cross section at 95% C.L. assuming $t' \to W^+b$ as the only possible decay mode. This translates into an observed lower bound on the t'-quark mass at $m_{t'} > 404$ GeV, for an expected exclusion limit at $m_{t'} > 394$ GeV (Figure 8.17).

8.5 Conclusion

The analysis presented in this chapter is a search for the pair-production of a heavy quark decaying into a W-boson and a b-quark. The benchmark model used is the chiral fourth generation t'-quark with assumed $BR\,(t' \to W^+b) = 1$. The search is performed with the pp collision data at $\sqrt{s} = 7$ TeV collected by the ATLAS detector in 2011 with a total integrated

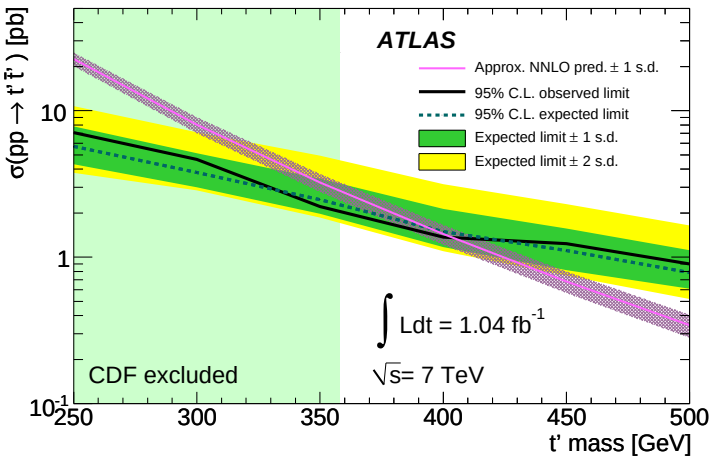


Figure 8.17: Observed (solid black line) and expected (dashed blue line) 95% C.L. upper limits on the $t'\bar{t}'$ cross section as a function of the t' mass. The blue bands correspond to the 1 and 2 standard deviations (s.d.) around the expected limit. The purple line shows the theoretical prediction including its 1 s.d. uncertainty band. The green shaded band is the mass region previously excluded by the CDF experiment [66]. Taken from [205].

luminosity of 1.04 fb^{-1}. Only events in the final states with one isolated high momentum electron or muon, at least three jets, and significant missing transverse energy are analysed. The full list of the selection requirements employed is presented in Section 7.2.

The reconstruction of initial t' and $\bar{t}'$-quarks employs different methods depending on the jet multiplicity. For the events with three jets, all jets are assumed to originate from the same hadronically decaying quark or antiquark. For the events with four or more jets, a likelihood based kinematic fit called KLFitter (see Section 6.1 for a description), has been used, to best assign the reconstructed objects to the initial quark or antiquark, and to fit the reconstructed invariant mass of the hadronically decaying quark, taking the jet energy resolution derived by comparing the generator and reconstruction level jet energy into account. The studies of the KLFitter performance carried out by the author of this thesis and presented in Section 8.3 have shown that the best resolution of the reconstructed mass is achieved when the fit is performed using the four leading p_{T} jets and the t' particle width corresponding to the mass of 172.5 GeV, while not taking the b-tagging information into account.

The SM backgrounds to this search are composed of $t\bar{t}$ production, single top production, W/Z-boson production in association with jets, QCD multi-jet production and diboson (WW, WZ, ZZ) production events, as explained in Section 7.3. The expected contributions of the $t\bar{t}$, single top, Z+jets and di-boson production are estimated from simulated events and normalised to the corresponding theoretical cross section calculated at NNLO. The W+jets contribution was estimated from simulation using a data-driven method for the normalisation estimation and a semi-data-driven method for the derivation of the flavour composition (see Section 7.5.2). The QCD multi-jet background is estimated from data using the matrix method (see Section 7.5.1).

The statistical analysis of the full reconstructed t' mass distribution has been performed using the log-likelihood ratio $LLR = -2\log(L_{\mathrm{s+b}}/L_{\mathrm{b}})$ as test statistics, whereby both $L_{\mathrm{s+b}}$ and L_{b} have been maximised with respect to the 12 nuisance parameters subject to Gaussian constraints of their priors, by which the sources of the largest systematic uncertainties are modeled. The

samples with 3 and $\geqslant 4$ jets have been treated simultaneously in order to constrain some of the systematic effects. As no significant excess of data events over the SM expectation has been observed, an upper limit placed on the $t'\bar{t}'$ production cross section at 95% C.L. translating into an observed lower bound on the t'-quark mass at $m_{t'} > 404$ GeV, for an expected exclusion limit at $m_{t'} > 394$ GeV [205].

Chapter 9

Search for the Vector-Like T-quark

The analysis presented in this chapter represents a search for the pair production of a heavy quark decaying into a W-boson and a b-quark, developed using the vector-like T (VLT) quark as a benchmark, and employing multi-variate analysis techniques. The search is performed in the lepton+jets final states with the full set of ATLAS pp collision data at $\sqrt{s} = 7$ TeV collected in 2011, corresponding to an integrated luminosity of 4.7 fb^{-1}. This analysis uses the basic analysis infrastructure adopted from the search published in [206], while the final selection of that search, which is composed of selection requirements on a set of kinematic variables, is replaced by a single requirement on a multi-variate discriminant, produced with the same variables using the Artificial Neural Network (NN) method. The improvement due to the NN-based selection is evaluated as the difference in the sensitivity achieved with these two approaches. The complete analysis presented in this chapter has been performed by the author of this thesis.

9.1 Motivation and Strategy

Multi-variate analysis techniques have proved to be powerful data analysis tools in many fields, and are gradually finding their place also in high energy physics analyses, where they often demonstrate an advantage in the selection efficiency over the more common event selection methods based on kinematic requirements. The focus of this study is the exploration of the analysis potential of a multi-variate NN method applied in the search for pair-produced VLT in the lepton+jets final states. The analysis potential is quantified as the sensitivity improvement obtained comparing two analyses, that differ only in the final selection. The reference analysis (published in [206]) uses selection requirements on eight kinematic variables to impose the final selection on the pre-selected event sample. In the tested analysis (developed by the author), the pre-selected event sample is used to produce a NN discriminator called MLP, trained on the same eight variables. The final selection of this analysis is performed by imposing a selection requirement on the MLP variable. In order to isolate the effect of the multi-variate selection on the analysis sensitivity, as much of the analysis infrastructure as possible is kept in common between the two analyses. Both analyses use the event pre-selection explained in Section 7.2. In events passing the pre-selection, T and $\bar{T}$ quarks are fully reconstructed as described in Section 9.2. The optimisation, validation and performance evaluation of the NN training is elaborated in Section 9.3. Section 9.4 demonstrates the level of agreement between data and

simulation in various control regions. In Section 9.5, the two mentioned final selections are defined. Systematic uncertainties affecting this search as well as their estimation are explained in Section 9.6. The results achieved are presented in Section 9.7 and the conclusion is provided in Section 9.8. The limitations of this analysis are explained and suggestions for its further refinement are provided in Section 9.9.

9.2 Event Reconstruction

In events passing the pre-selection defined in Section 7.2, T- and $\bar{T}$-quarks are reconstructed by making use of b-tagging information and the kinematic distinctiveness of $T\bar{T}$ events. In particular, kinematic differences between $T\bar{T}$ and $t\bar{t}$ events are exploited to maximally suppress this dominant background in the pre-selected sample. Being heavier than $t\bar{t}$ pairs, $T\bar{T}$ pairs are produced with weaker boost, i.e. more isotropic, allowing the direct decay products of T- and $\bar{T}$-quarks to be more separated in space compared to the decay products of the top quarks in $t\bar{t}$ events. Thus, the boost invariant space angle[1] ΔR between a b-quark and the corresponding W-boson is expected to be larger for the signal than for the background. On the other hand, as the T-quark is heavier than the top quark, W-bosons are expected to be more boosted in the signal events, so that their decay products are more collimated in space than for the $t\bar{t}$ background. Applying this consideration to the hadronically decaying W-boson, W_{had}, the two light jets in $T\bar{T}$ events are expected to be within a very small ΔR, which in some cases is smaller than 0.4, so that both light jets are contained within a single jet reconstructed by the anti-k_t algorithm with the size parameter $R = 0.4$. This inspires two methods of reconstructing W_{had}, and the analysed events are subdivided into two analysis sub-channels:

- Boosted channel, when the W_{had} is reconstructed as a single jet,
- Resolved channel, when the W_{had} is reconstructed as a pair of close-by jets.

A detailed description of the $T\bar{T}$ event reconstruction is explained below.

9.2.1 Identification of b jets

As explained in Section 7.2, pre-selected events are required to contain at least one jet tagged as originating from a b-quark by the MV1 algorithm with the tag probability of $\epsilon_b = 70\%$. For the purpose of reconstruction, two jets with the highest b-tagging probabilities, i.e. those with the highest MV1 weights, are considered to be the b jets of the event[2]. At this point it is not yet decided which b-jet corresponds to the leptonic and hadronic W-boson decays.

[1] The boost invariant space angle between two objects given by η and ϕ coordinates is defined as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$

[2] This b-jet identification differs from that of [206] in the order of the requirements. Contrary to this analysis, where only jets passing all quality criteria described in Section 5.5 are considered, the b-jets in [206] are defined before the p_{T} requirement on jets. It can therefore occur in that analysis that a jet with a p_{T} lower than 25 GeV becomes a b-jet of the event, whereas in this analysis it would be discarded.

9.2.2 Reconstruction of W_{had} in the BOOSTED channel

A jet that is not identified as a b-jet (see previous section) is considered a W-jet candidate, if it has a mass close to the W-boson mass, i.e. between 60 GeV and 110 GeV. Jet mass is computed using the jet energy and momentum measurements. Boosted W-bosons carry on average a higher p_{T} than a resolved W-bosons. Generator level studies performed for the reference analysis have shown that boosted W-bosons occur at transverse momenta larger than 250 GeV. Thus, a boosted W-boson candidate is also required to satisfy $p_{\text{T}} > 250$ GeV. In case more than one jet fulfils this requirement, that with a mass closest to the W-boson mass is identified as the W_{had}. The distribution of the reconstructed mass of the boosted W_{had} for the signal and the SM backgrounds is presented in Figures 9.1 (a) (normalised to unity) and 9.1 (b) (normalised to an integrated luminosity of 4.7 fb^{-1} assuming the production cross-section). The signal shows a pronounced peak at ~ 80 GeV, while backgrounds show a falling distribution. This clearly demonstrates that at $\sqrt{s} = 7$ TeV the W-bosons originating from particles lighter than or as massive as top quark cannot fulfil the BOOSTED channel requirement.

9.2.3 Reconstruction of W_{had} in the RESOLVED channel

Events without a boosted W_{had} (defined in the previous section) are tested for the presence of a pair of jets, that are not identified as b-jets that are separated by a space angle $\Delta R < 0.8$ and have a reconstructed invariant mass between 60 GeV and 110 GeV. If a di-jet system passing these two requirements also satisfies $p_{\text{T}} > 150$ GeV (which corresponds to a threshold above which such di-jet systems are found in the T sample at the generator level in the reference analysis), the system is identified as a resolved W_{had} candidate. If there is more than one W_{had} candidate in the event, that with the mass closest to the W-boson mass is identified as the W_{had} of the event. The distribution of the reconstructed mass of the resolved W_{had} for the signal and the SM backgrounds is presented Figures 9.1 (c) (normalised to unity) and 9.1 (d) (normalised to an integrated luminosity of 4.7 fb^{-1} assuming the production cross-section). Except for the T-quark signal, the $t\bar{t}$ background also sees a significant enhancement of the events around the W-boson mass above the falling distribution.

9.2.4 W_{had} Event Selection

A selection of the events containing a W_{had} as defined in the previous two sections, in the following called the W_{had}-selection, is defined by imposing the following requirement in addition to the pre-selection:

- Selected events are required to contain one[3] W_{had}, either boosted or resolved.

Events passing the W_{had}-selection are used in the NN training (see Section 9.3.1).

[3] As explained in the previous two sections, in case more than one W_{had} candidates fulfil the boosted or resolved W requirements, only one is defined as the W_{had} of the event following the mass consideration.

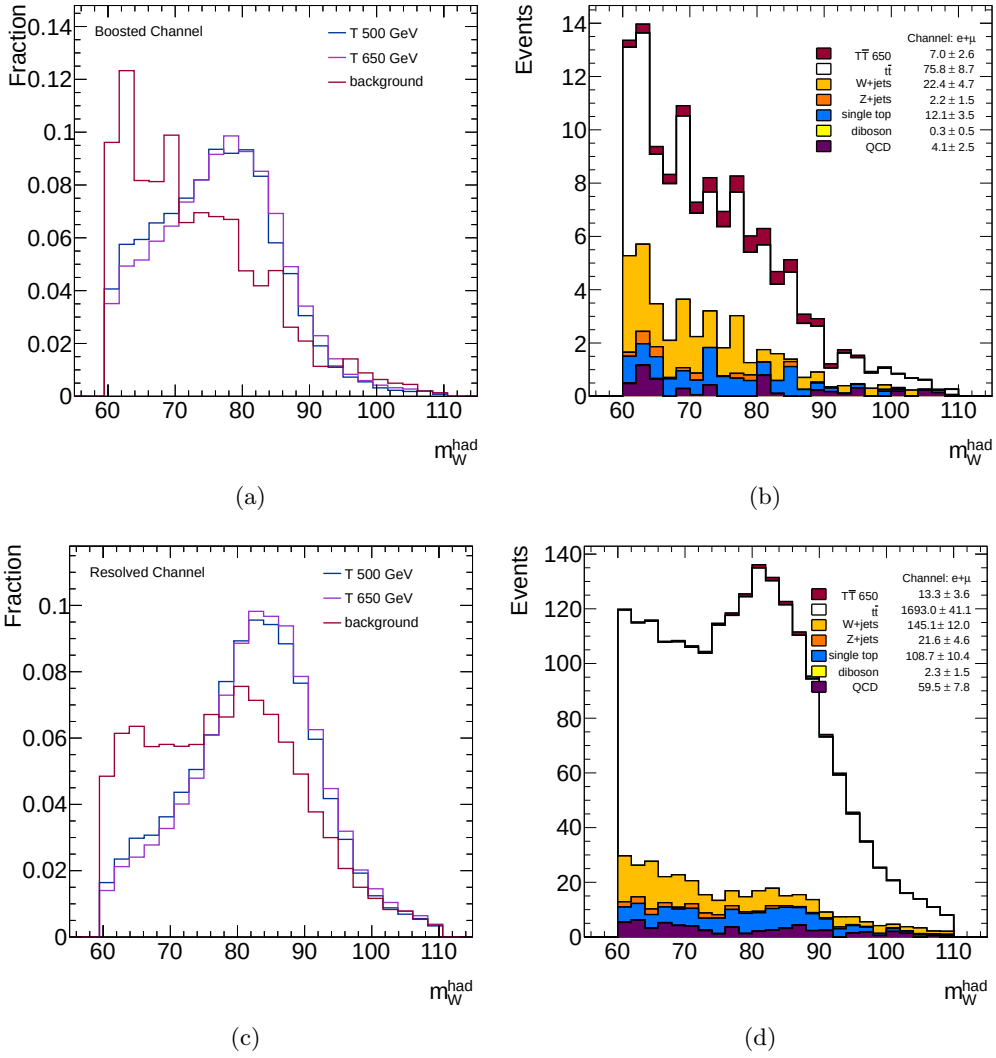


Figure 9.1: Reconstructed W_{had} mass in the BOOSTED ((a), (b)) and the RESOLVED ((c), (d)) channel. The shape difference between the T signal with the mass of 500 GeV and 650 GeV, and the SM background are shown by the distributions normalised to unity in (a) and (c). The stacked plot of the expected distributions of the SM background processes superimposed with the T-signal with the mass of 650 GeV are shown in (b) and (d). The expected distributions are normalised to an integrated luminosity of 4.7 fb^{-1} assuming the production cross-section.

9.2.5 Reconstruction of the T mass

To reconstruct T-quarks, the identified b-jets (see Section 9.2.1) need to be combined with the corresponding reconstructed W-bosons. The reconstruction of the hadronically decaying W-boson, W_{had}, is explained in Sections 9.2.2 and 9.2.3. The leptonically decaying W-boson, W_{lep}, is reconstructed as the vector sum of the lepton and neutrino 4-momenta. The lepton is fully defined, whereas only the transverse component of the neutrino momentum is given by $E_{\text{T}}^{\text{miss}}$ while the longitudinal component, p_z^{ν}, is unknown. It can, however, be estimated by requiring the invariant mass of the vector sum of the lepton and neutrino to be equal to the W-boson

mass:

$$P_W^2 = (P_l + P_\nu) = M_W^2. \tag{9.1}$$

That condition is quadratic in p_z^ν which can deliver zero, one or two real solutions. In case no real solution is found, the neutrino is assumed to have the same η as the lepton.

In cases two real solutions are obtained, two configurations are possible. As explained in Section 9.2.1, the assignment of the b-jets is also ambiguous. This allows for four possible T-quark mass solutions. The solution that provides the smallest absolute mass difference between the hadronically and leptonicaly decaying T-quarks, $|M(T_{\text{had}}) - M(T_{\text{lep}}))|$, is considered correct. For the chosen combination, the $M(T_{\text{had}})$ is taken as the reconstructed T-quark mass, M_{reco}, which is the variable on which hypothesis testing is carried out after the final selection, i.e. the final discriminant of the analysis. The M_{reco} distribution normalised to unity and alternatively to an integrated luminosity of 4.7 fb^{-1} assuming the production cross-section is presented in Figure 9.2 for the T-signal and the SM background, for the BOOSTED (Figures 9.2 (a), (b) and (c)) and RESOLVED (Figures 9.2 (d), (e) and (f)) channel separately, as well as combined (Figures 9.2 (g), (h) and (i)).

9.2.6 Discriminating Variables

After implementing the event reconstruction described above, all decay products of the $T\bar{T}$ event are fully defined, allowing for computation of kinematic variables related to them. The following kinematic variables with a good separation power are defined in the reference analysis:

- Total transverse momentum of the hard process $H_{\text{T}} = \sum_{j=1}^{n} p_{\text{T}}^j + p_{\text{T}}^l + E_{\text{T}}^{\text{miss}}$, where p_{T}^j stands for the transverse momentum of the jet j; the sum runs up to $n = 3$ for events with three jets and up to $n = 4$ for events with at least four jets.
- The transverse momenta of the b-jets assigned to the hadronic and leptonic T: $p_{\text{T}}^{b_{\text{had}}}$ and $p_{\text{T}}^{b_{\text{lep}}}$;
- The space angle between the lepton and the neutrino: $\Delta R^{l\nu}$;
- The space angle between W_{had} and each b-jet: $\Delta R^{W_{\text{had}} b_{\text{had}}}$ and $\Delta R^{W_{\text{had}} b_{\text{lep}}}$;
- The space angle between the lepton and each b-jet: $\Delta R^{l b_{\text{had}}}$ and $\Delta R^{l b_{\text{lep}}}$.

These variables are used for the final selection of the reference analysis as well as the training of the NN. The distributions of these eight variables produced from the event sample passing the W_{had}-selection, normalised to unity as well as to the integrated luminosity of 4.7 fb^{-1} assuming the production cross section, are presented in Figures 9.3-9.4 for a signal with a mass of either 500 GeV or 650 GeV, as well as the combined SM backgrounds.

9.2.7 Tag Rate Functions

The final selections in this analysis are extremely tight, leaving only a few expected background events, and also a very poor statistics of simulated events in the final templates. This results

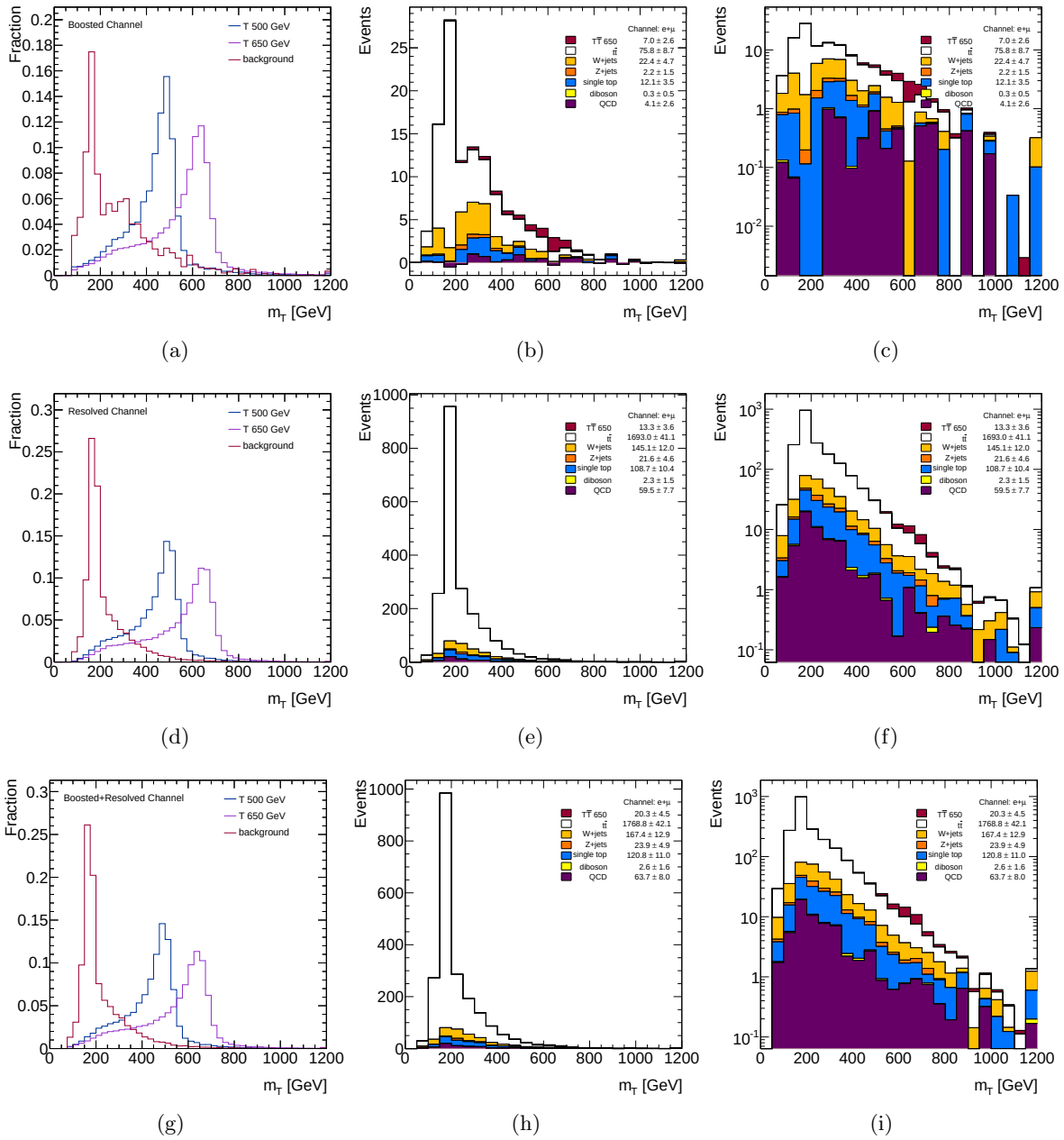


Figure 9.2: Reconstructed T-quark mass for the events passing the W_{had}-selection. The shape difference between the T-signal with the mass of 500 GeV and 650 GeV, and the SM background shown with distributions normalised to the unity are shown in (a), (d) and (g). The stacked plot of the expected distributions of the SM background processes superimposed with the T signal with the mass of 650 GeV normalised to an integrated luminosity of 4.7 fb^{-1} assuming the production cross-section are shown in linear ((b), (e), (h)) and logarithmic ((c), (f), (i)) scales.

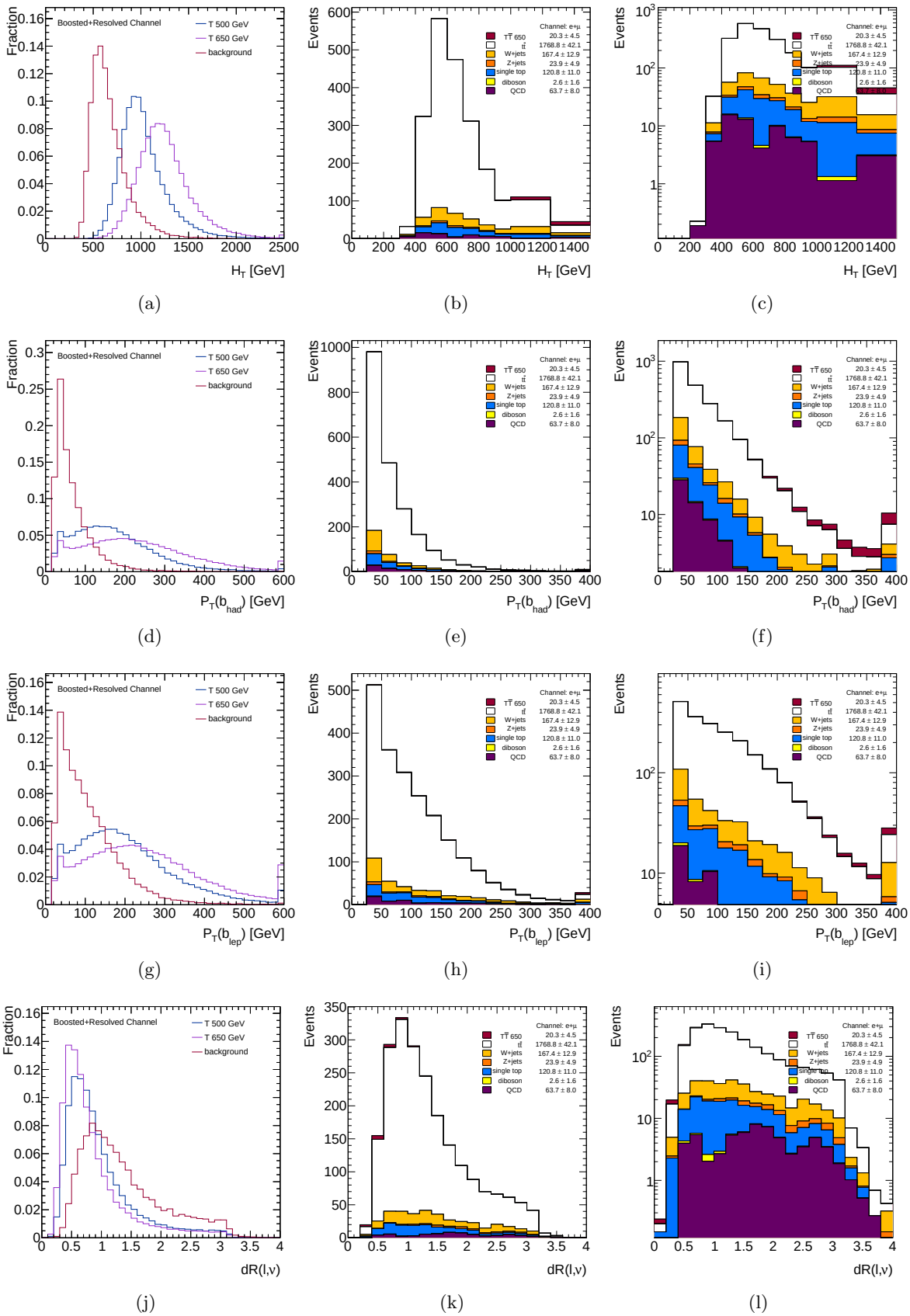


Figure 9.3: Discriminating variables for the signal events passing the W_{had}-selection with T mass of 500 GeV and 650 GeV and the combined MC background. The figures contain events passing either the BOOSTED or the RESOLVED selection. Distributions normalised to unity ((a), (d), (g) and (j)), and the absolute expected distributions normalised to an integrated luminosity of 4.7 fb^{-1} assuming the production cross-section in linear ((b), (e), (h), and (k)) and logarithmic ((c), (f). (i), and (l)) scales.

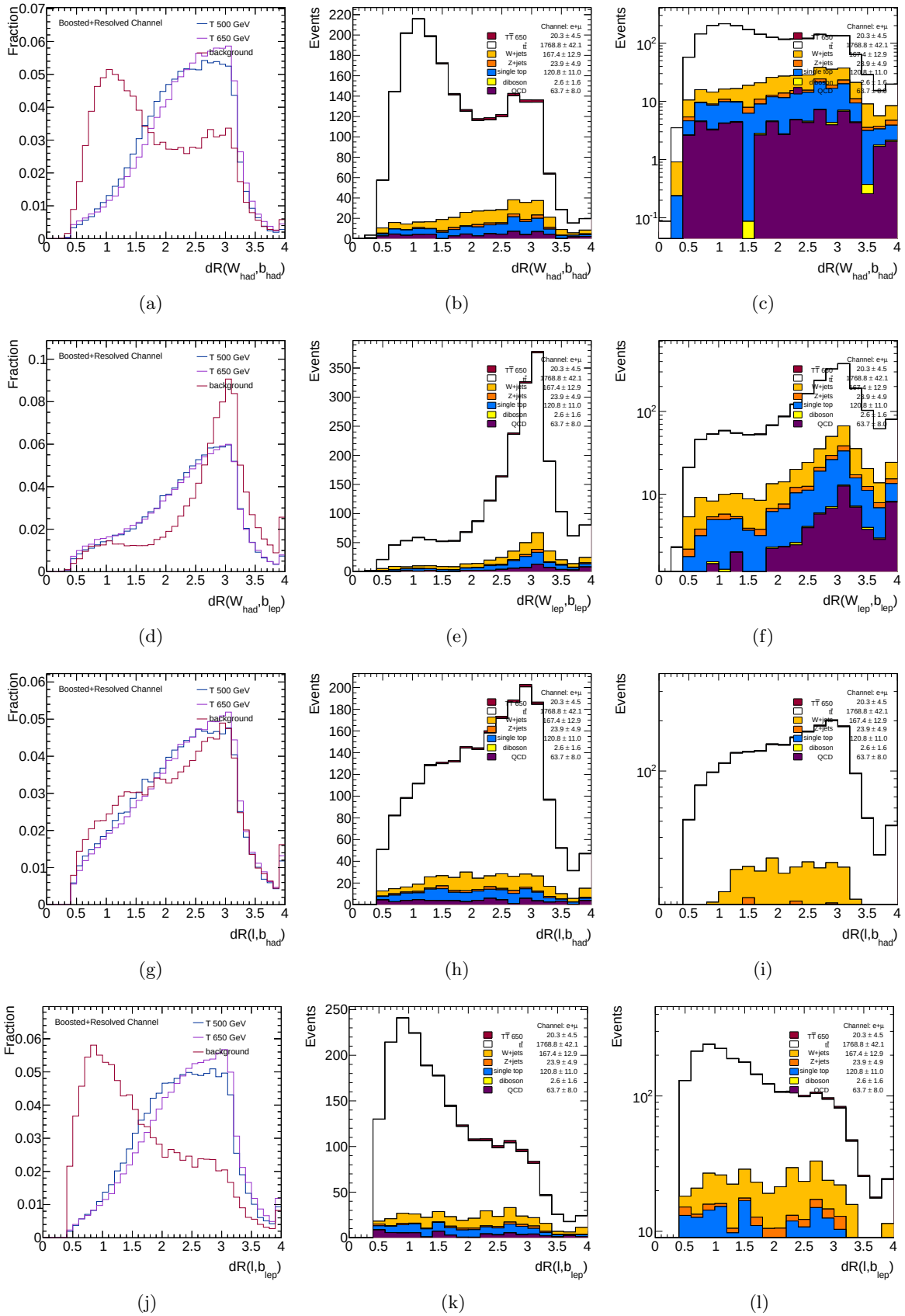


Figure 9.4: Discriminating variables for the signal events passing the W_{had}-selection with T mass of 500 GeV and 650 GeV and the combined MC background. The figures contain events passing either the BOOSTED or the RESOLVED selection. Distributions normalised to unity ((a), (d), (g) and (j)) and the absolute expected distributions normalised to an integrated luminosity of 4.7 fb^{-1} assuming the production cross-section in linear ((b), (e), (h) and (k)) and logarithmic ((c), (f), (i) and (l)) scales.

in unreliable systematic uncertainties on the variable shapes, dominated by large statistical fluctuations. To lessen this problem, a Tag Rate Function (TRF) method had been developed to account for the b-tagging information in simulated samples without rejecting events. Instead of rejecting events which do not contain sufficient number of b-tagged jets, TRF event weights have been applied, which correspond to the probability that a given event has a desired number of b-jets. The starting point of the method are the per-jet tag probabilities which depend on p_{T}, η and true flavour of the jet in question, and are derived from the simulation for a given jet clustering algorithm (jet author), b-tagging algorithm and working point. The TRF weights are then calculated using basic combinatorics. In an event with N jets with individual tag probabilities p_j, the probability of having at least one b-tagged jet $P_{\geqslant 1}$ is given by:

$$P_{\geqslant 1} = 1 - P_{=0}, \tag{9.2}$$

where $P_{=0}$ stands for the probability that no jet in the event is b-tagged:

$$P_{=0} = \prod_{j=1}^{N} (1 - p_j)\,. \tag{9.3}$$

While the TRF weights are applied to the simulated samples, the usual b-tagging requirement is imposed to data and QCD multi-jet background (more details in Section 7.2).

9.3 Multivariate Analysis

The multi-variate analysis employing the NN method described in Section 6.2 has been performed on the samples passing the W_{had}-selection using the discriminating variables explained in Section 9.2.6. This analysis employs the MLP method implemented in the TMVA package version 4.1.2 [173]. The samples and the kinematic variables used for the training as well as the MLP settings and their optimisation are presented in this section.

9.3.1 Input Samples

In this analysis, the NN was trained with two categories of simulated samples: signal and background. The chiral $T\bar{T}$ samples have been used as signal. Trainings have been optimised for two mass points, $m_T = 500$ GeV and $m_T = 650$ GeV. In both cases privately produced samples with high statistics ($\sim$ 1M events) have been used instead of very low statistics official samples (of 75k events) to allow for a higher precision of the machine learning procedure[4]. The samples have been produced following the same generation, simulation, reconstruction and final ntuple production prescriptions as used by the ATLAS MC group. The NN is trained against the combined MC background consisting of the $t\bar{t}$, W/Z+jets, single-top and di-boson samples. The QCD multi-jet background events have not been used in the training, as they often have

[4]The more random events are used for the training in a single training cycle, the larger is their diversity, and the better do they represent the full sample. By evaluating the error function (Equation 6.11) on a more representative sub-sample during the learning procedure, the influence of some non-representative features of a small portion of events onto the training is reduced, which accelerates the convergence of the training, and also diminishes the over-training risk.

Sample	Initial MC Stat.	W_{had}-Selection MC Stat.	W_{had}-Selection Event Yield
$T\bar{T}$(500 GeV)	800k	72k	92.5
$T\bar{T}$(650 GeV)	1M	122k	20.2
MC background	104M	90k	2436.5

Table 9.1: Summary of the sample statistics for signal and background samples. The first column shows the initial number of MC events per sample wile the second and the third show the number of unweighted and expected events in the samples after the W_{had} selection.

very high event weights compared to those of the simulated events, due to which the QCD events would be given a higher importance in the training compared to the other backgrounds. This would tune the NN training to more efficiently reduce the QCD multi-jet background on the cost of the rejection power of the dominant $t\bar{t}$ background, eventually leading to a poorer overall background suppression.

The NN trainings are carried out using events from both the e+jets and μ+jets channels that pass the W_{had}-selection to benefit from the maximal statistical power of the training sample. All samples are normalised to the integrated luminosity of 4.7 fb^{-1} assuming the theoretical production cross sections of the underlying process. A TRF weight (Section 5.7) corresponding to the probability that the event contains at least one b-tagged jet is applied to each event in addition. Since the TMVA implementation of the MLP method does not have an option of treating events with negative weights properly, events with unphysical negative weights, in particular $\sim$ 10% of $t\bar{t}$ events (produced by MC@NLO) have been ignored in the training. Events with negative weights, often occurring in NLO simulations, can introduce unphysical phase space into a training and thus lead to an unfavourable bias. The effects of this additional event selection onto the selection performance of the training are discussed in Section 9.3.7.2.

9.3.2 Sample Splitting

In general, the efficiency of any selection is better when it is applied to the event sample on which it is optimised than to any independent sample of events of the same mixture of the physics processes. The difference in these two efficiencies, i.e. the optimisation bias, is more strongly pronounced in case of a multi-variate selection than in case of a selection based on kinematic requirements, since a multi-variate discriminator is more capable of capturing fine particularities of the sample it is trained on, as explained in the discussion of the over-training phenomenon in Section 6.2.2. Therefore, the MC event sample available after the W_{had}-selection needs to be split in two parts, a training (A) and a testing (B) sample, with exactly the same kinematic properties. The training procedure is then performed on the sample A while the NN performance is evaluated on the sample B, which is finally used in the analysis[5]. Another training with the same settings is performed on the sample B and evaluated on the sample A, as shown in Figure 9.5. Evaluations of both trainings are used in the analysis.

[5]Strictly speaking, the available event sample would need to be split in three parts where one would be used for the training, one for the validation and one for the analysis. However, the differentiation of the latter two is not necessary since by the exclusion of the training events from the analysis the optimisation bias is already removed.

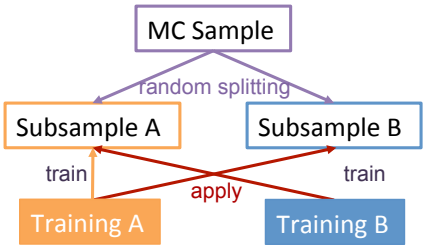


Figure 9.5: Illustration of the sample splitting. The original MC sample is randomly split into subsamples A and B. The training A (B) is trained on the subsample A (B) and evaluated on the subsample B (A).

The samples A and B do not necessarily need to be of equal size. Yet, both of them need to be as large as possible: the training sample in order to provide a sufficient information to the learning process and so ensure a good predictive power of the discriminator, but also the testing sample in order to minimise the statistical uncertainties of the discriminant in the analysis. Thus, a compromise has been met by splitting the available sample in two parts of equal statistics.

The splitting has been performed quasi-randomly based on the ϕ angle of the lepton, ϕ^l. This variable has been chosen as its value is not affected by any systematic uncertainty employed in this analysis (for a description of the systematic uncertainties used see Section 9.6). The events are categorised according to the remainder of division (modulus) by 2 of the integer content of the absolute value of ϕ^l, multiplied by a large (integer) number in order to retrieve the decimal places and so increase the randomness of the variable. In particular, the events are ordered as follows:

- sample A: $\text{Int}(|\ \phi^l\ | \cdot 125434413) \bmod 2 = 0$,
- sample B: $\text{Int}(|\ \phi^l\ | \cdot 125434413) \bmod 2 = 1$.

The multiplicative factor has been chosen iteratively in a way to ensure a good agreement between the two sub-samples, both in the size and the kinematic distributions used in the training.

The statistical error obtained by splitting a sample of N elements in two equal parts is given by $\sigma_{stat} = \sqrt{N/2}$, which in this case equals 246.9 for the signal sample and 184.4 for the combined MC background. As shown in Table 9.2, the number of events in the sub-samples A and B is within $1\sigma_{stat}$ for both signal and background. For completeness, the splitting is also quoted in terms of expected event yields for the signal and background as well as observed yields for data. The level of agreement in the kinematic distributions used for the NN training between these two sub-samples is discussed in Appendix C.2.

9.3.3 Input Variables

The choice of the input variables is a crucial step in building an efficient NN training. For a successful learning process, a NN needs to be supplied with sufficient information through the input variables with good discrimination powers. An insufficient set of variables leads to a poor predictive power of the NN outcome, while on the other hand a superfluous number of uninformative variables amplifies the dimensionality of the optimisation problem, and thus the computational burden as well as the training statistics requirement. Furthermore, a simultaneous usage of highly correlated variables introduces redundancies which could increase the

	$T\bar{T}$(650 GeV)	MC background	Data
MC Stat. A	60831	33954	–
MC Stat. B	61073	34022	–
Absolute Diff. [events]	242	68	–
Absolute Diff. [σ_{stat}]	0.98	0.37	–
Relative Diff. [%]	0.4	0.2	–
Event Yield A	10.1352	1191.8	1205
Event Yield B	10.1044	1187.93	1220
Absolute Diff. [events]	0.0308	3.87	15
Absolute Diff. [σ_{stat}]	1.02	0.03	1.03
Relative Diff. [%]	0.3	0.3	1.24

Table 9.2: MC statistics and expected event yields of the subsamples A and B and the differences between them for the signal and background. The absolute differences are expressed in terms of the number of simulated (expected) events for the upper (lower) half of the table, as well as in terms of the number of standard deviations σ_{stat} defined as $\sigma_{stat} = \sqrt{N/2}$. Relative differences are expressed in percents. Observed event yields in the samples A and B are shown in the last column.

number of local minima of the error function in the space of the synapse weights, and could thus enhance the risk of the back-propagation procedure converging to a false minimum.

A number of algorithms for the variable choice optimisation have been developed and are in some training interfaces incorporated into the training procedure. Since the goal of this analysis is to explore the potential of the sensitivity improvement compared to an equivalent cut-based analysis, the set of the input variables has not been optimised. Instead, in order to make a direct comparison between the cut-based and multi-variate selection, the variables with which have been used for the final selection of the cut-based analysis (Section 9.2.6) are adopted as the NN input variables. The discrimination power of these variables is expressed in terms of the *separation* $\langle S^2 \rangle$:

$$\langle S^2 \rangle = \sum_{i=1}^{n} \frac{(s_i - b_i)^2}{s_i + b_i}, \tag{9.4}$$

with the signal and background rates s_i and b_i in the bin i and the sum running over all bins of the variable in question, normalised to unity. Separation for the NN input variables between the combined simulated background and the signal with the T-quark mass of 500 and 650 GeV is presented in Table 9.3. The separation gain of the signal with the mass of 650 GeV relative to that with the mass of 500 GeV is pointed out in the last row of this table. The transformation of the higher T-quark mass into the energy of its decay products is visible through the improvement of the separation of the p_{T}-like variables H_{T}, $p_{\mathrm{T}}^{b_{\mathrm{had}}}$ and $p_{\mathrm{T}}^{b_{\mathrm{lep}}}$ as well as the $\Delta R^{l\nu}$ variable, which shows the boost enhancement of the W-boson when increasing the T-quark mass from 500 to 650 GeV. For completeness, the discriminant M_{reco} is included in the table, according to which it has the second best separation power after the W_{had}-selection, slightly weaker than the H_{T} variable.

Although the choice of variables was not optimised, the correlations among them have been studied, to assure the absence of any negative effects of redundant information. The correlations have been evaluated in terms of the Pearson's (linear) correlation coefficients. For variables X

	H_{T}	M_{reco}	$p_{\mathrm{T}}^{b_{\mathrm{had}}}$	$p_{\mathrm{T}}^{b_{\mathrm{lep}}}$	$\Delta R^{l\nu}$	$\Delta R^{lb_{\mathrm{lep}}}$	$\Delta R^{W_{\mathrm{had}}b_{\mathrm{had}}}$	$\Delta R^{W_{\mathrm{had}}b_{\mathrm{lep}}}$	$\Delta R^{lb_{\mathrm{had}}}$
$\langle S^2\rangle_{500}$	0.527	0.572	0.358	0.179	0.155	0.196	0.120	0.062	0.006
$\langle S^2\rangle_{650}$	0.707	0.600	0.475	0.295	0.280	0.207	0.131	0.056	0.008
$\Delta\langle S^2\rangle/\langle S^2\rangle_{500}$	34.2%	4.7%	32.7%	64.8%	80.6%	5.6%	9.2%	−9.7%	33.3%

Table 9.3: Separation power of the NN training variables and discriminant M_{reco} between the combined MC background and the T signal with $m_T = 500$ GeV and $m_T = 650$ GeV respectively.

and Y with mean values $\bar{x}$ and $\bar{y}$ respectively, the correlation coefficients are given by:

$$R_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}, \tag{9.5}$$

with i running over all events.

Figure 9.6 shows the linear correlations between the input variables used in this analysis. The correlation matrices represent the linear correlations for two signal samples used, i.e. the T samples with masses of 500 GeV and 650 GeV and the combined backgrounds. The only pair of rather correlated variables are $\Delta R^{W_{\mathrm{had}}b_{\mathrm{had}}}$ and $\Delta R^{lb_{\mathrm{lep}}}$. This correlation is present in both signal samples as well as in the background, but is more pronounced for the signal. It reflects the fact that events with the decay products close in space on one side of the decay tend to have the decay products close to each other also on the other side of the decay. In other words, the events in question tend to be rather symmetric, i.e. with similar boosts of the initial quark and anti-quark. The correlation is more pronounced in the signal than in the background, suggesting that the signal events are more symmetric. The effect of this correlation on the discrimination power of the NN discriminator has been studied and the conclusion is presented in Section 9.3.7.5.

9.3.4 Training Configuration

The training optimisation is carried out iteratively by varying numerous NN parameters and training settings, always using the sub-sample A defined in Section 9.3.2 as the training sample. Every training is evaluated in terms of its validity and discrimination power. The training validation is elaborated in Section 9.3.5, while the evaluation of the performance is explained in Section 9.3.6. In this section, the configuration of the best performing NN training is presented. A comparison of the performance of the best training and a set of trainings with different parameters is elaborated in Section 9.3.7.

The NN is trained with the high statistics signal sample with the T-quark mass of 650 GeV using 8 discriminating variables presented in Section 9.3.3 as input. The network architecture consists of a single hidden layer with 28 nodes ($N + 20$ with N being the number of the input variables). Taking into account one bias node, each in the input and hidden layer, the number of synapses in the network adds up to 281. The neurone activation function used is tanh (Equation 6.10). The best performing learning procedure is back-propagation. These settings are summarised in Table 9.4.

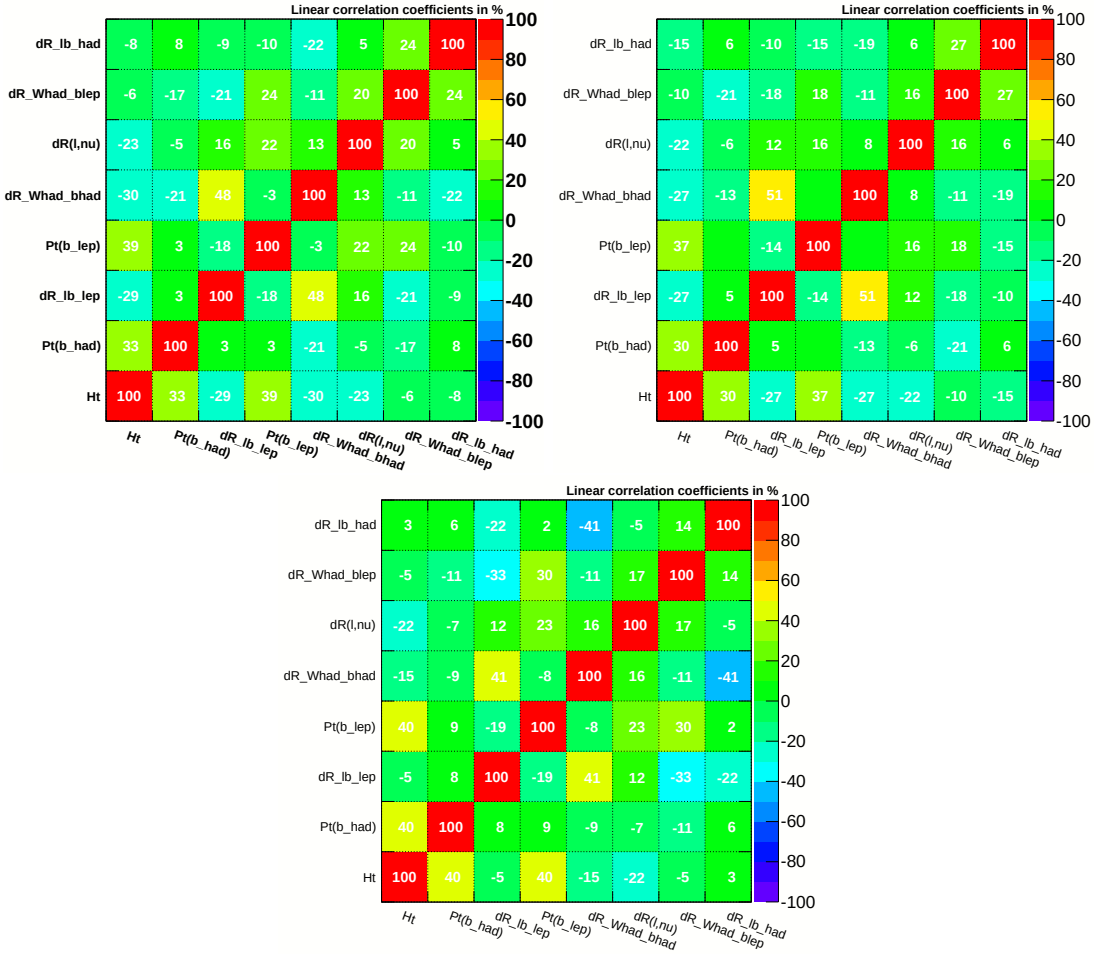


Figure 9.6: Linear correlations between the input variables for signal with $m_T = 500$ GeV (top left), signal with $m_T = 650$ GeV (top right) and combined MC background (bottom).

Setting	Chosen option
T mass	650 GeV
Number of input variables	8
Number of hidden layers	1
Number of nodes in the hidden layer	28
Number of synapses	281
Neuron activation function	tanh
Estimator function	cross-entropy
Learning procedure	back-propagation
Number of training cycles	700

Table 9.4: Optimal NN settings.

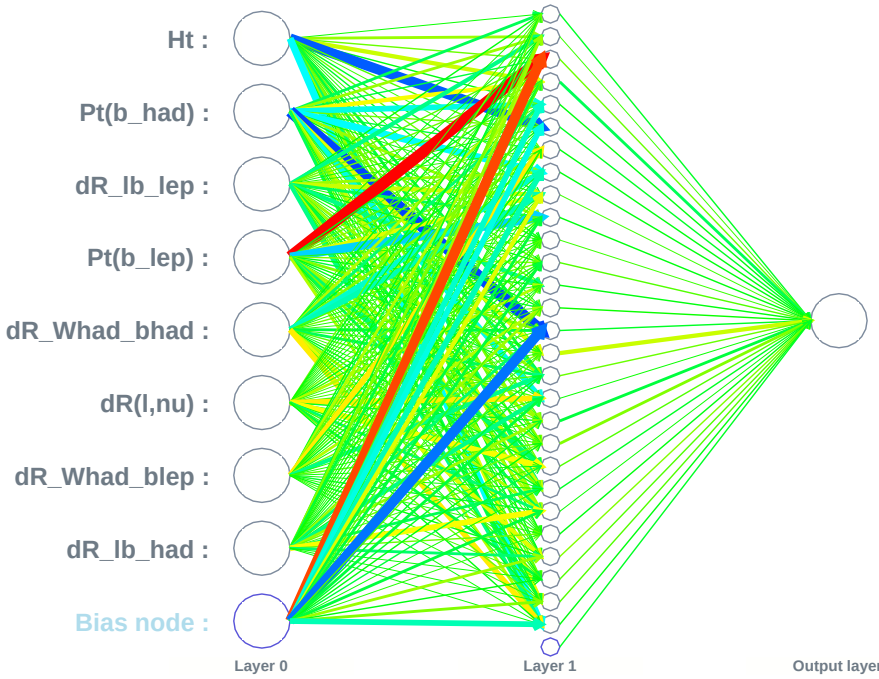


Figure 9.7: Final network architecture. Eight input variables are arranged in the input layer and accompanied by an independent bias node. The hidden layer shows 28 hidden nodes. The output node is the only node of the output layer. The synapses between the input ant the hidden layer, as well as those between the hidden and the output layer nodes are painted as colourful lines. The width and colour scheme depict the magnitude and the sign of the synapse weight respectively. More details are given in the main text.

The resulting network architecture is presented in Figure 9.7. The nodes are represented as circles and the synapses as straight lines. The input nodes obtain the normalised values of the input variables, and the weight of the output node represents the MLP output, as explained in Section 6.2. The synapse weights are represented by their width and colour. The width of a line is proportional to the absolute value of corresponding weight. The colours ranging from green to red represent the positive synapse weights in the increasing order of their weights. The negative synapse weights carry the colours in the range from blue-green to blue in decreasing order.

Variable	Importance
$p_{\mathrm{T}}^{b_{\mathrm{had}}}$	245.5
$p_{\mathrm{T}}^{b_{\mathrm{lep}}}$	169.4
H_{T}	78.95
$\Delta R^{l\nu}$	24.75
$\Delta R^{lb_{\mathrm{lep}}}$	12.77
$\Delta R^{lb_{\mathrm{had}}}$	8.71
$\Delta R^{W_{\mathrm{had}}b_{\mathrm{had}}}$	5.94
$\Delta R^{W_{\mathrm{had}}b_{\mathrm{lep}}}$	1.18

Table 9.5: Variable ranking according to the importance.

The strength of the contribution the variable i has provided to the network is expressed in terms of the importance I_i given by:

$$I_i = \bar{x_i}^2 \sum_{j=1}^{n_s} (w_{ij}^0)^2, \tag{9.6}$$

where the sum runs over all outgoing synapses of the input neurone corresponding to the variable i. The variable importance for the optimal training is presented in Table 9.5. This table shows that the momentum-related variables make a stronger impact to the training than the angular distributions. Transverse momenta of the b-jets have the highest importance, followed by the H_{T} variable. Among the angular variables, $\Delta R^{l\nu}$ has the strongest impact.

In Section 9.3.2 the necessity of training a multi-variate discriminator on an independent set of events to that used in the analysis has been explained. However, proceeding in the explained way, i.e. using only approximately 1/2 of the available statistics in the analysis leads to an enhancement of the statistical uncertainties of the discriminant templates. Nevertheless, this issue can be bypassed by a simple training rotation, as performed in this analysis. Instead of using (say) sample A for training and sample B for testing and analysis, two NN trainings with identical settings are performed: one on the sub-sample A (training A) and the other one on the sub-sample B (training B). If the performance of both training measured on the corresponding testing sample is comparable, both trainings can be used on complementary subsets: training A on the sample B and training B on the sample A. In this way it is ensured that no training is ever applied to events on which it has been trained, therefore avoiding an optimisation bias. The same sample splitting and training evaluation mechanism is applied to the data events as well. The validation of the method is presented in the next section.

9.3.5 Training Validation

The error function for both training and testing samples is monitored in every training cycle. Figure 9.8 shows the error function evolution during the training procedure for training and testing samples for both training A and training B. For both samples and trainings, the error function shows a few large fluctuations at the beginning of the training, followed by a slowly decreasing trend until the convergence is reached after $\sim$ 500 cycles. The absence of any raise of the error function for the testing samples indicates that there are no signs of over-training.

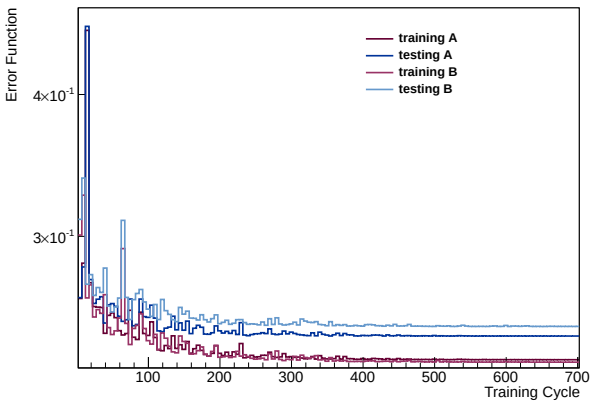


Figure 9.8: The estimator (error function) value evolution through the training cycles is shown for trainings A (darker colour lines) and B (lighter colour lines). For both trainings the error function is evaluated on the corresponding training sample (red) and testing sample (blue).

The level of predictability of a training can be studied by comparing the NN response distribution MLP evaluated on the training sample used and the corresponding testing sample. Any irregularity of the training procedure, such as an insufficiently randomness of the sample splitting, over-training or under-training, can result in differences in the MLP variable between the training and the testing sample. A comparison of the MLP distribution between the training and testing samples is shown in Figure 9.9 for the trainings A (left) and B (right). The distributions are normalised to an integrated luminosity of 4.7 fb^{-1} assuming the production cross section of the underlying processes. The figure demonstrates that the categorisation efficiency is compatible between the

training and testing sample within 1 or 2 statistical σ for most of the bins for both trainings.

However, the two trainings differ slightly in their prediction power. Both trainings describe the signal better than the background, especially in the proximity of the signal region (i.e. in the high tail of the MLP distribution), which is due to the higher training statistics of the signal sample. In particular, in the last bin of the distribution (which contains the signal region, as discussed later on), both trainings see a very good agreement between the training and testing samples for the signal, while they see more background events in the testing sample. In other words, more background events are categorised as signal-like in the testing sample than in the training sample. This suggests that the trainings have difficulties with separating background events that look similar to signal events from the signal, which can be explained by the scarceness of such events, resulting in an insufficient information for a more detailed machine learning.

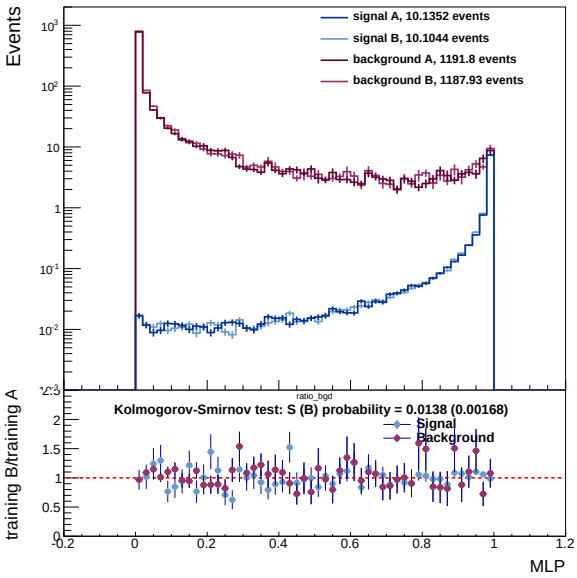


Figure 9.10: The comparison of the NN response MLP evaluated on the corresponding testing samples for the trainings A and B.

In Figure 9.10 the MLP distributions of the testing sample only (i.e. of the events used in the analysis) for the trainings A and B are compared. The distributions agree well between the the two trainings for both signal and background, especially in the last bin which is essential for the analysis.

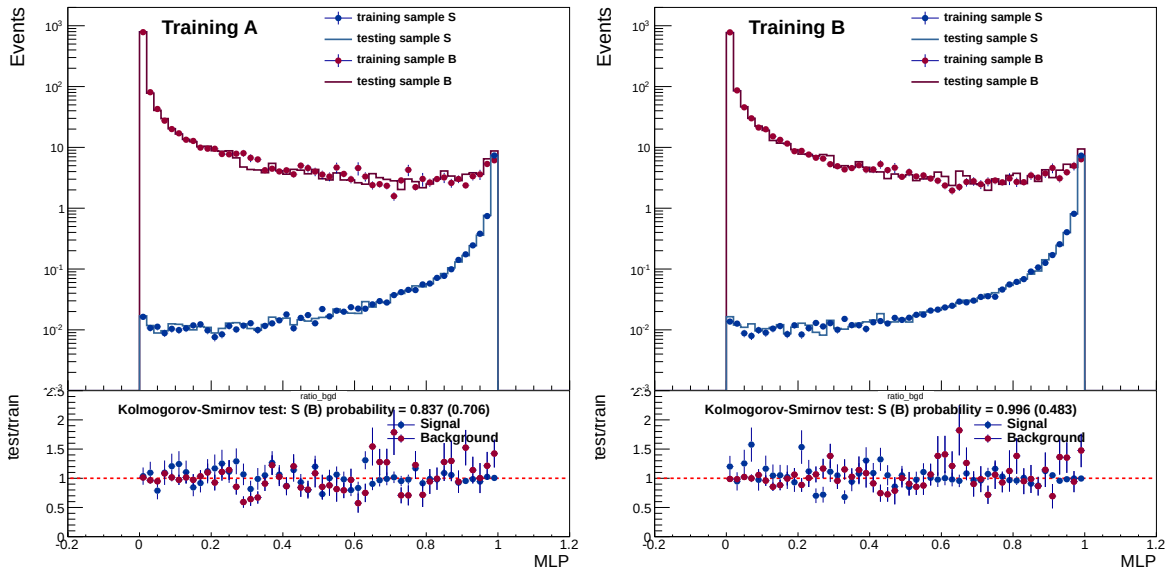


Figure 9.9: The NN response MLP for the trainings A (left) and B (right) shown for training and testing sample and for signal and background separately.

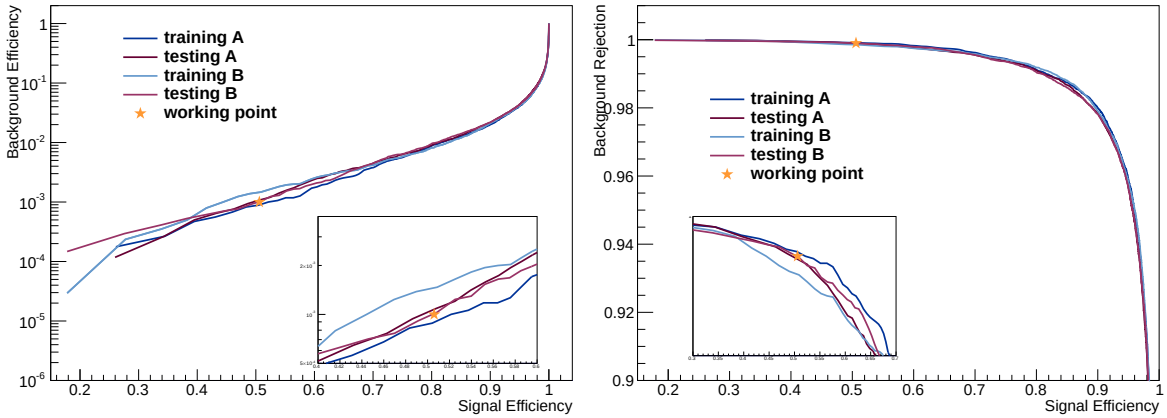


Figure 9.11: Background efficiency (left) and rejection (right) as function of the signal efficiency measured with the MC events for the training (blue lines) and the testing (red lines) samples of the trainings A and B. The working point is shown by an orange star. The magnification of the figures in the region close to the working point is shown in the boxes at the bottom of the plots.

9.3.6 Training Performace

The feature of interest of a selection is its power of discriminating signal from background, which can be expressed in terms of various quantities, such as the background efficiency or rejection as a function of the signal efficiency, signal over background ratio S/B, where S (B) stands for the number of expected signal (background) events, significance $S/\sqrt{S+B}$, etc. The choice of the quantity used to evaluate a selection depend on what the selection is optimised for (e.g. high signal efficiency or purity). The training feature of interest for this analysis is the NN response curve, i.e. the curve showing the background acceptance as a function of the signal acceptance, in terms of the number of expected events. In this section, the performance of the trainings A and B is quantified in terms of the mentioned quantities.

Figure 9.11 shows the background efficiency (left) and rejection (right) as a function of the signal efficiency measured with simulated events. The figures are shown for the trainings A and B. The real efficiencies are those measured in the testing samples while the efficiencies measured on the corresponding training samples are shown for a comparison. The working point defined later on in Section 9.5.2 is shown by an orange star for completeness.

The difference between the training (blue line) and the testing (red line) efficiency of a given training illustrates the optimisation bias, which is rather small for the trainings A and B shown in the plots. Although around the working point the training sample efficiency for the training A seems to be slightly better than the corresponding testing sample efficiency, due to the high statistical uncertainties in high tail of the MLP distribution, across the full range the efficiencies measured on the training samples are slightly better than those derived from the testing samples, as expected.

The S/B and the significance $S/\sqrt{S+B}$ as a function of the MLP threshold are shown in the left plot of Figure 9.12. The S/B curve raises continuously up to the MLP cut at 0.999 where it reaches a value of 4.73 after which it starts shrinking. The maximal significance of 2.84 is reached at the MLP cut at 0.9945. The working point (orange vertical line in the plot) is defined by an MLP threshold slightly tighter than that providing the maximal significance and corresponds to $S/B = 2.98$ and $S/\sqrt{S+B} = 2.78$.

The NN response MLP represents a nonlinear combination of all input variables. However, studying the linear correlations between the MLP variable and the input variables shown in Figure 9.13 provides a clue of which region of which variable corresponds to signal and which to background. As the low region of the MLP variable corresponds to background and the high tail to signal, for the variables highly correlated to the MLP variable, such as the H_{T} variable, the high region corresponds to signal and the low region to background. On the other hand, for the variables anti-correlated to the NN response, such as $\Delta R^{l\nu}$, the low values of the variable contribute more significantly to signal and the high values to background. For the completeness, the M_{reco} discriminant is included in the correlation matrices. The high linear correlation between M_{reco} and the $\Delta R^{W_{\mathrm{had}}b_{\mathrm{had}}}$ for the signal clearly illustrates the correlation between the boost of the W-boson and the mass of the object it originates from. For the two dimensional distributions of the input variables as a function of the NN output and M_{reco} see Appendix C.6.

9.3.7 Training Optimisation

Once the training with the optimal response curve had been established iteratively (see Section 9.3.4), the optimality of that training configuration was tested by performing a set of trainings, which differ from the optimal configuration in a single setting. For each training setting, several options have been tested, keeping all other setting as in the optimal training. All trainings have been performed on the event sample A defined in Section 9.3.2. Accordingly, their performance is evaluated on the sample B and compared to that of the best training of the type A. The difference of the performance has been evaluated in terms of the NN response function defined in Section 9.3.6 close to the signal region. Figure 9.14 shows the response functions for different options of a single setting while the other parameters remain unchanged. The curve corresponding to the optimal training is drawn as a black line in all plots. The orange and blue stars show the working points for the final NN selection (see Section 9.5.2) and the reference TIGHT selection (see Section 9.5.1) respectively.

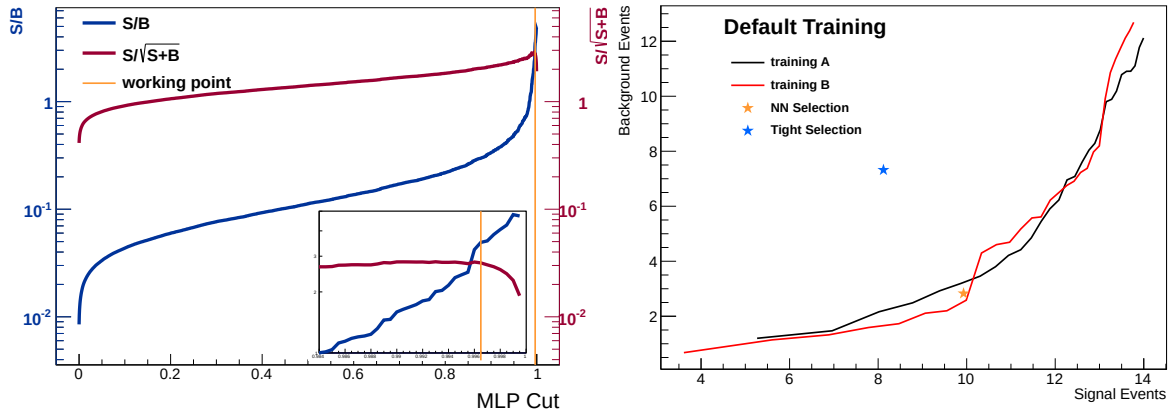


Figure 9.12: Left: Signal to background ratio and the expected significance as a function of the MLP cut. The working point is shown by an orange line. The magnification of the region close to the working point is shown in the box at the bottom right of the plot. Right: the NN response curve for the trainings A and B. The working point determined on a combination of the two trainings is presented by an orange star. The working point of the TIGHT selection is shown by a blue star.

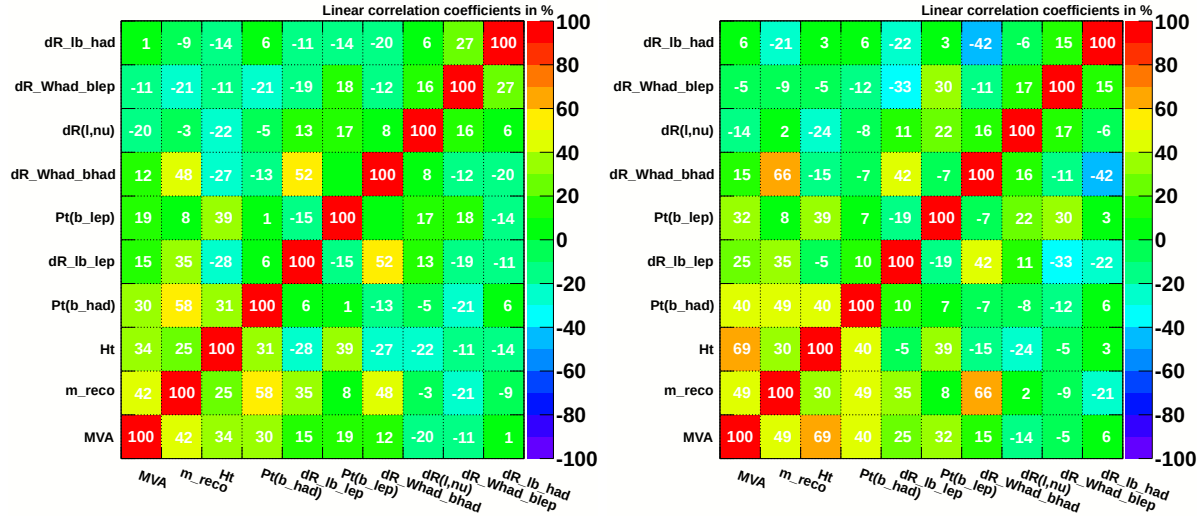


Figure 9.13: Linear correlations between the NN response (labeled as "MVA" in the plot), the M_{reco} discriminant and input variables for signal with $m_T = 650$ GeV (left) and combined MC background (right).

9.3.7.1 NN Architecture

Given the limited training samples statistics, a simple network architecture with a single hidden layer has been chosen, to avoid generating a too large number of degrees of freedom to be trained. However, the number of the nodes in the hidden layer has been optimised for. Figure 9.14(a) shows that among the tested configurations, the best performance is achieved with $N+20$ nodes in the hidden layer, where N stands for the number of input variables.

9.3.7.2 Input Pre-processing

As discussed in Section 9.3.1, $t\bar{t}$ events with negative weights ($\sim 10\%$ of the simulated events) are excluded from the training to avoid their potentially negative influence on the training. To test the effect of the events with negative weights, a training has been performed with them included. According the the Figure 9.14(b), this effect is negligible.

9.3.7.3 Training Options

After the renormalisation of the input variables, the first step in the signal propagation through the NN is the neurone activation function. Four functions have been tested, one linear and three non-linear (see Equation 6.10). As clearly shown in Figure 9.14 (c), the non-linear activation functions result in a much better response function than the linear activation function, because they allow the learning procedure to profit from non-linear correlations among the input variables. The neurone function which delivers the best NN response curve is tanh. Cross-entropy (CE) is found to perform better than the mean square estimator (MSE), as shown in Figure 9.14 (d). Back-propagation, labeled as "MLP" in Figure 9.14 (e), turns out to be the best performing learning method. The other two learning methods, the BFGS algorithm without (MLPBFGS) and with (MLPBNN) bayesian regulators, take 5 and 10 times longer for the training and require $\sim 50\%$ more training cycles due to a slower convergence. The plot labeled

as "Number of Training Cycles" shows that the trainings trained in 600, 700 and 800 cycles perform equally well. However, since the training B shows a slightly slower convergence in the testing sample, which can be seen in Figure 9.9, trainings are performed with 700 training cycles.

9.3.7.4 Signal Mass Point

Trainings with T-quark masses of 500 GeV and 650 GeV have been compared. According to Figure 9.14 (d), the training with $m_T = 650$ GeV shows a stronger separation power.

9.3.7.5 Number of Variables

The set of input variables has not been optimised for this analysis, as it was adopted from [206]. However, it has been observed that the variables $\Delta R^{W_{\mathrm{had}} b_{\mathrm{had}}}$ and $\Delta R^{lb_{\mathrm{lep}}}$ are rather highly correlated in the signal samples, as explained in Section 9.3.3. The potential negative effects of this correlation have been studied by comparing the training with all eight variables with those without one of these two correlated variables. Figure 9.14 (f) shows that the training indeed profits from both of these variables, so both are used. Although the variable $\Delta R^{lb_{\mathrm{lep}}}$ in addition shows a very poor separation, it is included in the construction of the NN output to enable a fair comparison of the reference selection to a multi-variate selection with the same set of variables.

9.4 Expected and Observed Event Yields

In this section, the observed event yields as well as those expected for the integrated luminosity of 4.7 fb^{-1} after the pre-selection and the W_{had}-selection are presented. Before looking at data in the region where the T-signal is expected, the agreement between the observed and expected event yields as well as the goodness of the reconstruction of the variables used in the final selection, have been studied in the regions enriched in background and depleted in signal. Several so-called control regions are defined by placing requirements on the representative variables, as described in Section 9.4.1. Variables with a good separation between the signal and the backgrounds, such as M_{reco} and H_{T}, represent a particularly suitable basis of the control region definitions, as they allow for eliminating the majority of the signal while cutting off a relatively small part of the background phase space. It is beneficial to use several control regions in order to allow for the examination of each variable in its full range. For example, the high tail of the M_{reco} distribution (in which the presence of the signal is expected), which is cut off (blinded) to define a signal region, can be looked at in a sample from which most of the signal events are removed by imposing a requirement on some other variable, such as H_{T}. The results of the comparison between data and simulation in the control regions after the pre-selection and the W_{had}-selection are shown in Sections 9.4.2 and 9.4.3, while those with no blinding requirements are shown in Section 9.4.4. The expected and observed distribution of the NN response is shown in all regions, while the corresponding distributions of the input variables can be found in Appendices quoted in each section. In these figures, the the SM backgrounds are divided into a $t\bar{t}$ and a non-$t\bar{t}$ component, since this division is used also in the hypothesis testing, as explained further below in Section 9.7.1.

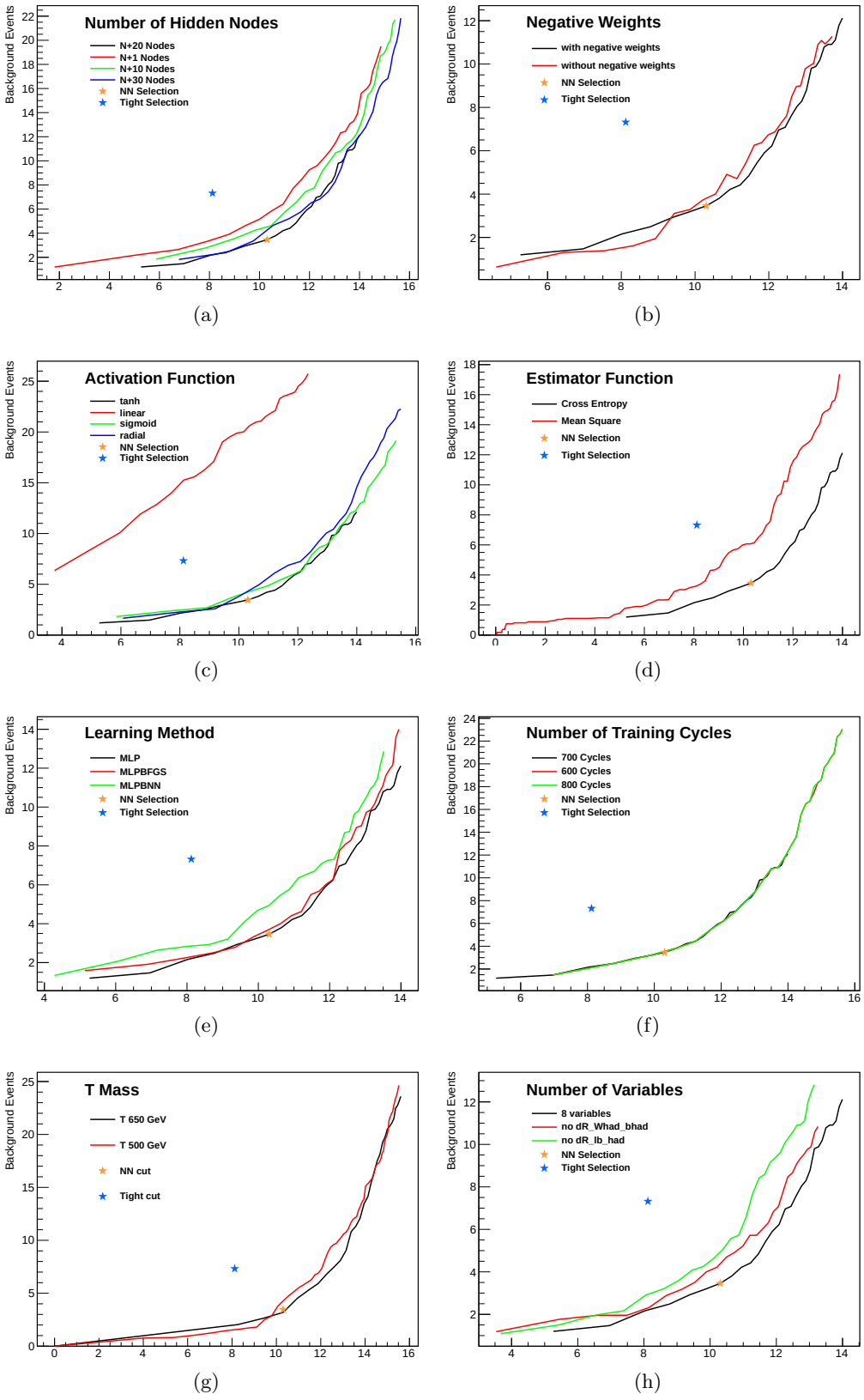


Figure 9.14: Optimisation of the MLP training parameters: number of the nodes in the hidden layer (a), employing or not events with the negative event weights in the training (b), neurone activation function (c), estimator function (d), learning method (e), number of training cycles (f), T mass point (g), and number of variables (f). The default setting is drawn in black in all plots. The working point or the Tight and NN-based final selections defined in Section 9.5 are shown as blue and orange star respectively.

9.4.1 Control Regions

Three control regions have been defined by demanding the following requirements:

- $\geqslant 4$ jets, $= 0\,b$-tags
  Given the b-tag veto, this region is dominated by the W+jets component while containing very few expected signal events.
- $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $M_{\text{reco}} < 350$ GeV
  By requiring a low reconstructed mass, the region is depleted in signal while allowing a closer examination of the high tail of the H_{T} distribution, This region is dominated by the $t\bar{t}$ background.
- $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $H_{\text{T}} < 700$ GeV
  This region provides an insight into the high tails of the M_{reco} distribution while ensuring a low signal contamination. The dominant component in this region is the $t\bar{t}$ background.
- $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $MLP < 0.8$
  This region shows the level of the agreement between the expectation and the observation when cutting on the NN response MLP.

Except for the last control region which is examined only after the W_{had}-selection (as the NN training is performed on the sample passing the W_{had}-selection), all other control regions are examined after two selection stages: after the pre-selection (Section 9.4.2) and after the W_{had}-selection (Section 9.4.3).

9.4.2 Event Yields in the Control Regions after Pre-selection

The expected and observed event yields in the first three control regions defined in Section 9.4.1 applied to the pre-selection are summarised in Tables 9.6–9.8. The expected number of events is shown for the signal with the mass of 500 GeV and 650 GeV, as well as for all SM backgrounds separately and together. The event yields are presented for the e+jets and μ+jets channels separately and combined. The plots of the NN input variables corresponding to these regions are shown in Appendices C.4.1–C.4.3.

The 0 b-tags control region sees overall $\sim 0.3\%$ fewer data events than expected. The low M_{reco} and low H_{T} control regions see a data excess of $\sim 8\%$. However, the plots in the mentioned figures show that most of these fluctuations are within 1σ systematic error bands, except for a small number of individual bins with slightly larger fluctuations.

9.4.3 Event Yields in the Control Regions after W_{had}-selection

The expected and observed event yields in the control regions defined in Section 9.4.1 applied to the W_{had}-selection are given in Tables 9.9–9.12. The expected number of events is shown for the signal with the mass of 500 GeV and 650 GeV, as well as for all SM backgrounds separately and together. The event yields are presented for the e+jets and μ+jets channels separately and combined. The comparison between the expected and observed MLP distribution in the

Sample	e	μ	$e+\mu$
$T\bar{T}(500\ \text{GeV})$	10.8±0.2	11.5±0.2	22.3±0.22
$T\bar{T}(650\ \text{GeV})$	1.8±0.02	1.8±0.02	3.6±0.03
$t\bar{t}$	2718.0±6.0	4166.0±7.4	6884.0±9.5
W+jets	9814.1±69.8	18597.5±100.4	28411.5±122.3
Z+jets	2024.8±23.0	2136.1±28.9	4161.0±36.9
single top	289.3±3.1	451.2±3.7	740.5±4.8
di-boson	151.4±2.1	233.4±2.6	384.9±3.4
Multi-jet	2137.2±34.9	2171.0±21.5	4308.2±41.0
Total prediction	17134.8±81.7	27755.2±107.0	44890±134.7
Data	16151	27615	43766

Table 9.6: Event yields with statistical uncertainties in the $\geqslant 4$ jets, $= 0\,b$-tags control region after pre-selection. The total prediction includes all SM backgrounds.

Sample	e	μ	$e+\mu$
$T\bar{T}(500\ \text{GeV})$	74.3±1.0	80.4±1.1	154.8±1.5
$T\bar{T}(650\ \text{GeV})$	10.5±0.2	11.0±0.16	21.5±0.2
$t\bar{t}$	15634.4±26.5	24268.3±33.1	39902.7±42.4
W+jets	1908.2±24.1	3545.2±32.1	5453.3±40.1
Z+jets	342.2±6.1	357.6±7.1	699.9±9.4
single top	1088.8±8.9	1715.6±11.1	2804.4±14.2
di-boson	35.1±0.7	53.8±0.9	88.9±1.1
Multi-jet	957.4±28.2	1392.9±15.7	2350.3±32.3
Total prediction	19966.1±46.9	31333.4±50.5	51299.4±68.9
Data	21159	34321	55480

Table 9.7: Event yields with statistical uncertainties in the $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $M_{\text{reco}} < 350$ GeV control region after pre-selection. The total prediction includes all SM backgrounds.

first three control regions is shown in Figure 9.15 for the e+jets and the μ+jets channels. The corresponding plots of the NN input variables are contained in Appendices C.4.5-C.4.8.

All signal regions after the W_{had}-selection see a data excess of $\sim 10\%$. In the 0 b-tags control region there is 9.6% more data, but the MLP variable has an agreement within 1σ systematic between the data and the expectation. The low M_{reco} region has a data excess of 12.1%. The agreement between the data and the expectation in this region is within 1σ except for two bins with fluctuations to $\sim 1.5\sigma$ and $\sim 3\sigma$. In the last bin of the MLP distribution, which contains the signal region, there are slightly fewer data events than expected in the electron channel. In the muon channel, however, the last bin sees a data excess of $\sim 3\sigma$, while 3 other bins see a data excess of $\sim 2\sigma$. In the low H_{T} control region an overall data excess of 10.1% is observed. The agreement between the data and the expectation is within 1σ for the bins which contain more than one expected event, while the fluctuations become larger in the bins with fewer than one event expected. Only $\sim 2\%$ of events in this region exceeds an MLP value of 0.2, which clearly shows that most of events with $H_{\text{T}} < 700$ GeV are classified as background by the NN discriminator. The low MLP signal region has a 12.3% excess of observed events, and an agreement between the expectation and observation mostly within 1σ systematic.

Sample	e	μ	$e+\mu$
$T\bar{T}$(500 GeV)	11.7±0.4	14.5±0.5	26.1±0.6
$T\bar{T}$(650 GeV)	0.42±0.03	0.58±0.04	1.0±0.05
$t\bar{t}$	14771.3±25.7	23232.6±32.3	38003.8±41.3
W+jets	1686.6±22.7	3241.5±30.7	4928.1±38.1
Z+jets	301.7±5.7	328.4±6.8	630.0±8.9
single top	1009.6±8.6	1627.8±10.8	2637.3±13.7
di-boson	33.0±0.7	51.5±0.8	84.6±1.1
Multi-jet	900.9±27.3	1361.9±15.5	2262.7±31.4
Total prediction	18703±45.0	29843.7±48.9	48546.7±66.5
Data	19781	32643	52424

Table 9.8: Event yields with statistical uncertainties in the $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $H_{\mathrm{T}} < 700$ GeV control region after pre-selection. The total prediction includes all SM backgrounds.

Sample	e	μ	$e+\mu$
$T\bar{T}$(500 GeV)	3.8±0.08	4.2±0.09	8.0±0.12
$T\bar{T}$(650 GeV)	0.78±0.02	0.80±0.01	1.6±0.02
$t\bar{t}$	120.4±1.3	159.8±1.5	280.2±1.9
W+jets	281.2±11.4	413.0±13.9	694.2±18.0
Z+jets	54.5±3.6	49.9±4.2	104.4±5.6
single top	10.7±0.58	14.6±0.69	25.3±0.91
di-boson	4.7±0.37	5.4±0.39	10.1±0.54
Multi-jet	59.1±5.9	39.1±2.9	98.3±6.6
Total prediction	530.7±13.4	681.8±14.9	1212.6±20.1
Data	561	769	1330

Table 9.9: Event yields with statistical uncertainties in the $\geqslant 4$ jets, $= 0\,b$-tags control region after W_{had}-selection. The total prediction includes all SM backgrounds.

Sample	e	μ	$e+\mu$
$T\bar{T}$(500 GeV)	9.9±0.38	10.9±0.40	20.8±0.55
$T\bar{T}$(650 GeV)	1.5±0.05	1.5±0.06	3.0±0.08
$t\bar{t}$	680.9±5.6	909.9±6.5	1590.8±8.5
W+jets	47.5±3.8	74.4±4.8	122.0±6.1
Z+jets	9.6±1.1	9.2±1.3	18.8±1.7
single top	40.3±1.9	49.2±2.1	89.5±2.8
di-boson	0.80±0.10	1.04±0.12	1.84±0.15
Multi-jet	26.14±5.23	25.04±2.25	51.18±5.69
Total prediction	805.3±8.84	1068.8±867	1874.1±12.4
Data	894	1207	2101

Table 9.10: Event yields with statistical uncertainties in the $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $M_{\mathrm{reco}} < 350$ GeV control region after W_{had}-selection. The total prediction includes all SM backgrounds.

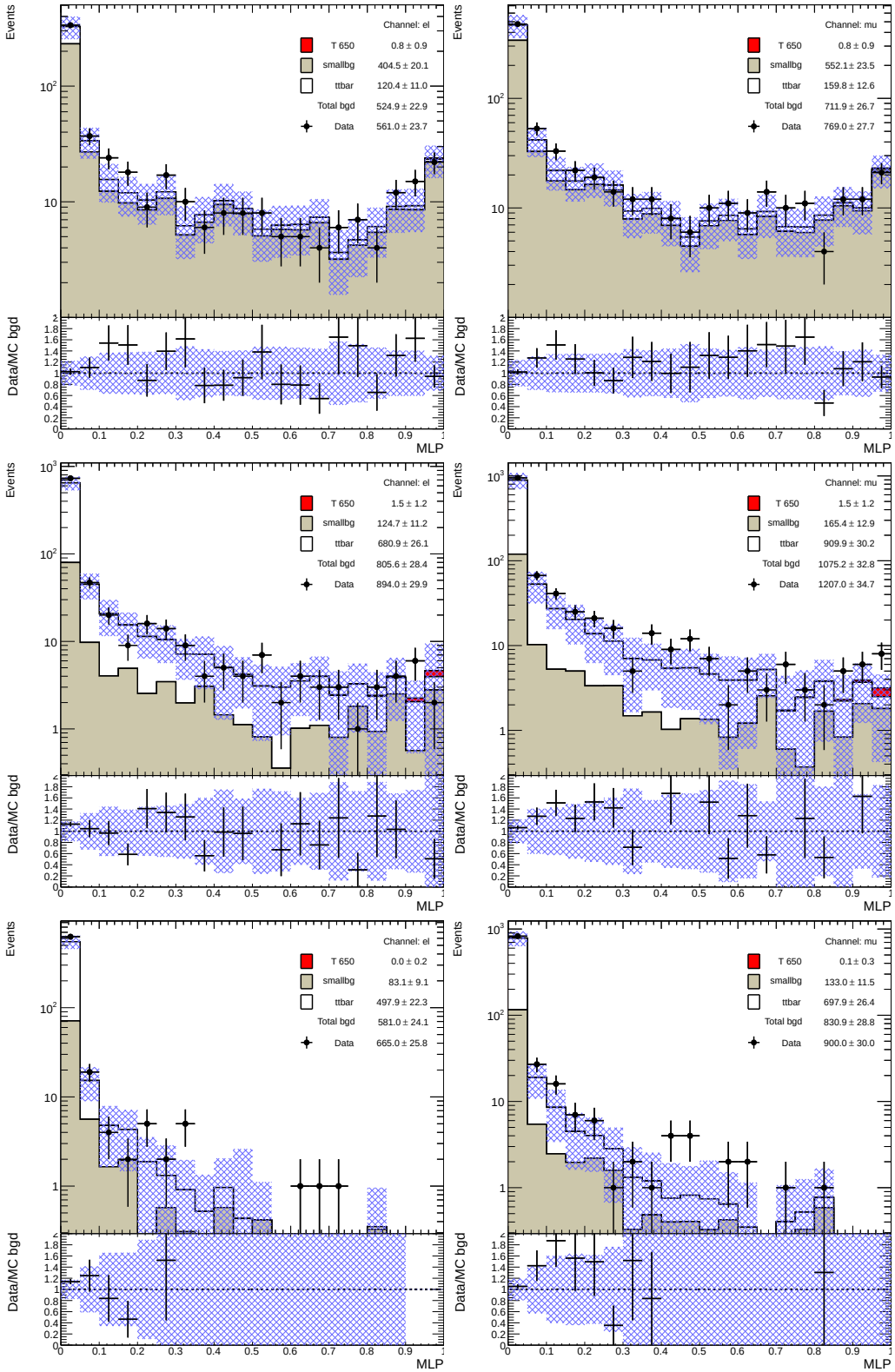


Figure 9.15: MLP distribution in the $\geqslant$ 4 jets, $= 0\,b$-tags (top row), $\geqslant$ 4 jets, $\geqslant 1\,b$-tags, $M_{\text{reco}} < 350$ GeV (middle row), and $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $H_{\text{T}} < 700$ GeV (bottom row) control regions after W_{had}-selection, in the e+jets (left) and μ+jets (right) channels. The blue shaded area represents the 1σ systematic error bands on the expected event yields.

Sample	e	μ	$e+\mu$
$T\bar{T}$(500 GeV)	2.0±0.17	2.8±0.20	4.7±0.26
$T\bar{T}$(650 GeV)	0.05±0.01	0.08±0.01	0.13±0.02
$t\bar{t}$	497.8±4.7	697.9±5.6	1195.7±7.3
W+jets	31.7±3.0	55.9±4.2	87.6±5.2
Z+jets	5.7±0.85	7.3±1.1	13.0±1.4
single top	29.3±1.6	41.0±1.9	70.3±2.5
di-boson	0.73±0.10	0.89±0.11	1.6±0.15
Multi-jet	15.4±4.5	22.8±2.1	38.2±4.95
Total prediction	580.6±7.4	825.9±7.6	1406.5±10.6
Data	665	900	1565

Table 9.11: Event yields with statistical uncertainties in the $\geqslant 4$ jets, $\geqslant 1$ b-tags, $H_{\mathrm{T}} < 700$ GeV control region after W_{had}-selection. The total prediction includes all SM backgrounds.

Sample	e	μ	$e+\mu$
$T\bar{T}$(500 GeV)	10.68±0.39	13.14±0.44	23.82±0.59
$T\bar{T}$(650 GeV)	0.66±0.04	0.80±0.04	1.5±0.06
$t\bar{t}$	729.55±5.83	985.44±6.78	1714.99±8.94
W+jets	49.78±3.91	92.44±5.58	142.23±6.81
Z+jets	10.30±1.17	10.74±1.32	21.03±1.76
single top	46.39±2.04	60.54±2.33	106.93±3.10
diboson	1.04±0.12	1.21±0.12	2.25±0.17
Multi-jet	27.72±5.49	29.47±2.43	57.19±6.00
Total prediction	864.78±9.22	1179.84±9.49	2044.62±13.23
Data	965	1331	2296

Table 9.12: Event yields with statistical uncertainties in the $\geqslant 4$ jets, $\geqslant 1$ b-tags, $NN < 0.8$ control region after W_{had}-selection. The total prediction includes all SM backgrounds.

9.4.4 Event Yields in the Full Region after W-selection

The event yields for the e+jets and μ+jets channels separately as well as combined after the W_{had}-selection are given in Table 9.13. After W_{had}-selection, e+jets channel sees $\sim 11\%$ more data than expected, while in the μ+jets channel data overshoots the expectation by $\sim 15\%$.

The expected and observed NN response distribution is presented in Figure 9.16 for the electron and muon channels separately. In both channels, the per-bin observed event yields are mostly within the 1σ systematic error bands, except for several bins. The bins with the fluctuations exceeding 1σ systematic uncertainty are distributed randomly across the MLP spectrum in the electron channel, while in the muon channel several bins in the low region (covering $0.05 < MLP < 0.3$), one bin around $MLP = 0.5$, as well as the last bin of the distribution see $> 1\sigma$ upward fluctuations. In both channels, the last bin exceeds the expectations, whereby in the electron channel this excess is within 1σ systematic, while in the muon channel it reaches $\sim 2\sigma$. The plots showing a comparison between the expectation and observation in the NN input

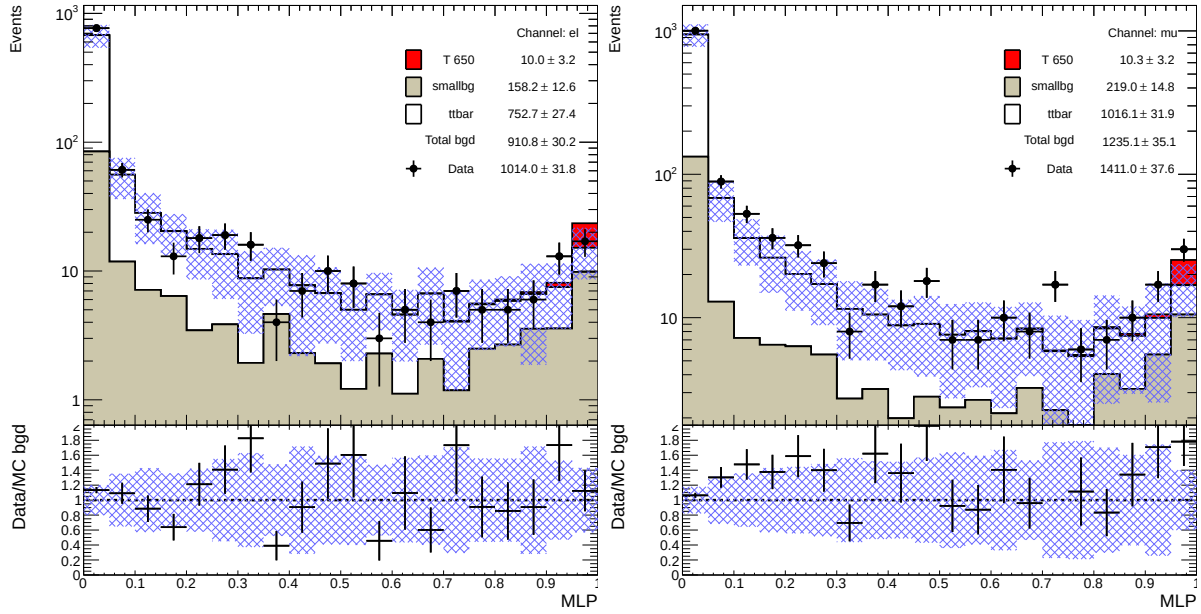


Figure 9.16: NN output distribution for W_{had}-selection in the e+jets (left) and μ+jets (right) channels. The blue shaded area represents the 1σ systematic error bands on the expected event yields.

variables at different stages of the selection are presented in Appendix C.4.9).

9.5 Final Selection

The sample passing the consecutively applied pre-selection (Section 7.2) and the W_{had}-selection (Section 9.2.4) is further refined by imposing either of the two final selections defined in this section. Both final selections are based on the discriminating variables defined in Section 9.2.6. In the so-called cut-based final selection, very stringent kinematically motivated requirements are imposed on the discriminating variables, as explained in Section 9.5.1. In the NN-based

Sample	e	μ	$e+\mu$
$T\bar{T}(500\ \text{GeV})$	45.16±0.81	47.77±0.83	92.93±1.16
$T\bar{T}(650\ \text{GeV})$	9.99±0.15	10.30±0.16	20.29±0.22
$t\bar{t}$	752.70±5.94	1016.13±6.89	1768.82±9.10
W+jets	59.29±4.19	98.09±5.46	157.38±6.88
Z+jets	12.03±1.25	11.82±1.38	23.85±1.86
single top	53.14±2.18	67.70±2.48	120.84±3.30
di-boson	1.29±0.14	1.31±0.13	2.60±0.19
Multi-jet	32.10±5.70	31.57±2.51	63.67±6.23
Total prediction	910.55±9.57	1226.61±9.58	2137.16±13.54
Data	1014	1411	2425

Table 9.13: W_{had}-selection event yields with statistical uncertainties for an integrated luminosity of 4.7 fb^{-1}.

Sample	e	μ	$e+\mu$
$T\bar{T}$(500 GeV)	14.23±0.46	14.20±0.45	
$T\bar{T}$(650 GeV)	4.03±0.10	4.08±0.10	
$t\bar{t}$	1.64±0.34	2.25±0.38	3.89±0.51
W+jets	1.02±0.44	0.751±0.395	1.77±0.59
Z+jets	0.258±0.229	0.007±0.007	0.265±0.229
single top	0.654±0.236	0.700±0.407	1.35±0.47
di-boson	0.037±0.029	0±0	0.037±0.029
Multi-jet	1.04±0.59	0.255±0.224	1.30±0.63
Total prediction	4.65±0.88	3.96±0.72	8.62±1.13
Data	6	7	13

Table 9.14: Tight selection event yields with statistical uncertainties for an integrated luminosity of 4.7fb^{-1}.

election described in Section 9.5.2, a single requirement is imposed to the NN discriminator MLP obtained by training a NN with the same variables. Both final selections yield a very clean sample with very few SM background events. All data events accepted by either final selection are presented in in Appendix C.7.

9.5.1 Cut-based Final Selection

The cut based final selection called Tight selection in the following is performed by imposing the following requirements:

- $H_{\mathrm{T}} \geqslant 750$ GeV,
- the p_{T} of the harder b-jet is required to be $\geqslant 160$ GeV and the one of the softer jet is required to be $\geqslant 60$ GeV,
- $\Delta R^{l\nu} \leqslant 1.4$.
- $\Delta R^{W_{\mathrm{had}}b_{\mathrm{had}}}$ and $\Delta R^{W_{\mathrm{had}}b_{\mathrm{lep}}} \geqslant 1.4$,
- $\Delta R^{lb_{\mathrm{had}}}$ and $\Delta R^{lb_{\mathrm{lep}}} \geqslant 1.4$.

These requirements adopted from [206], where they were established based on generator level studies. The event yields for the e+jets and μ+jets channels separately as well as combined after the Tight selection are given in Table 9.14. The expected and observed distributions of the discriminating variables after the Tight selection are presented in Appendix C.4.10.

9.5.2 NN-based Final Selection

As the events deemed more likely to be signal are assigned the values of the NN discriminator MLP close to unity, the NN-based signal region is defined by imposing a lower bound on the MLP variable. The exact value of MLP at which the bound is placed has been optimised in a way to reach the highest possible expected limit on the T-quark mass. As discussed in Section

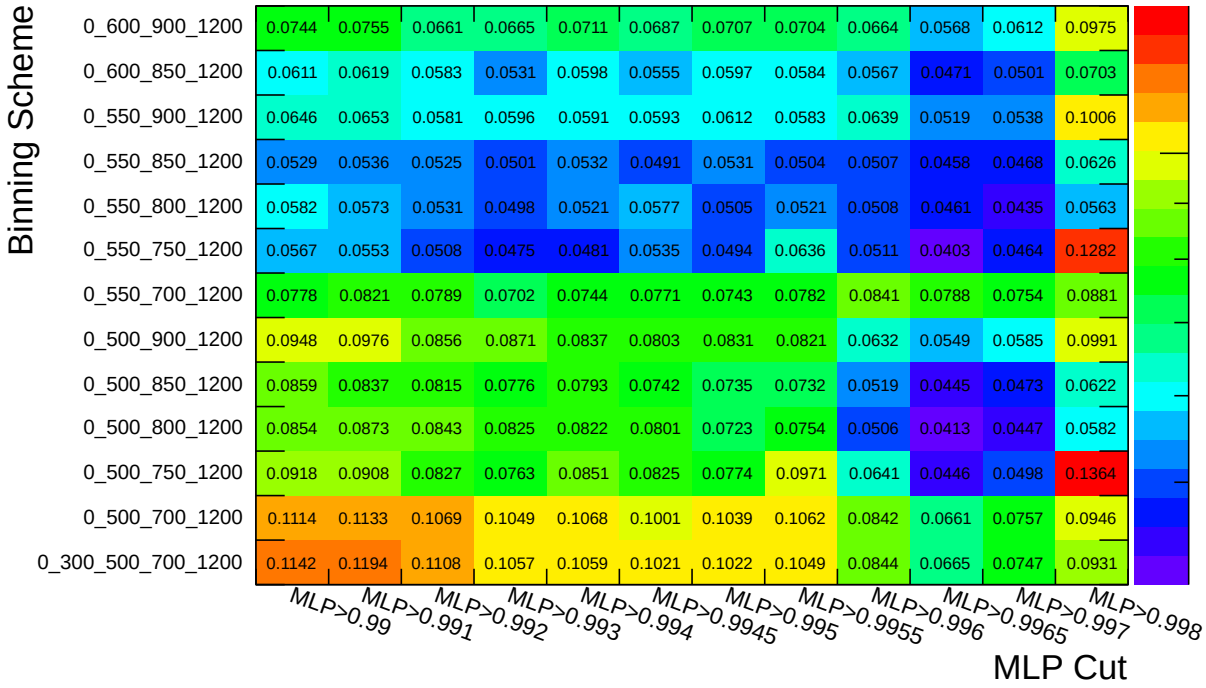


Figure 9.17: Expected CLs for the T with the mass of 700 GeV as a function of the cut on the NN output and the binning scheme (with quoted values given in GeV).

6.3, the core of the hypothesis testing is a binned log likelihood function, which relies on the bin contents of the discriminant distributions for the signal and background. In other words, the likelihood function is dependent on both the MLP threshold and the binning of the M_{reco} discriminant. Therefore, these two settings are optimised simultaneously.

The expected limit has been produced using the method described in Section 6.3 for a set of binning schemes and a range of MLP thresholds. A coarse optimisation is performed by producing the median expected CLs value for the T-signal with the mass of 700 GeV (the mass point closest to the expected limit achievable with this selection) tested against the background only hypothesis as a function of the MLP requirement and the binning scheme, as shown in Figure 9.17. For a number of best configurations, that is those with the lowest expected CLs values, the expected limit at 95% confidence level (C.L.) is computed by a linear interpolation between two CLs values: one below the limit (i.e. smaller than 0.05) and one above the limit (i.e. greater than 0.05). According to Figure 9.18, the best limit is achieved with the selection requirement $NN > 0.9965$ and the binning scheme $(0, 500, 800, 1200)$ GeV. The event yields for the e+jets and μ+jets channels separately as well as combined after the Tight selection are given in Table 9.15. The distributions of the discriminating variables of expected and observed events passing the NN>0.9965 selection are presented in Appendix C.4.11.

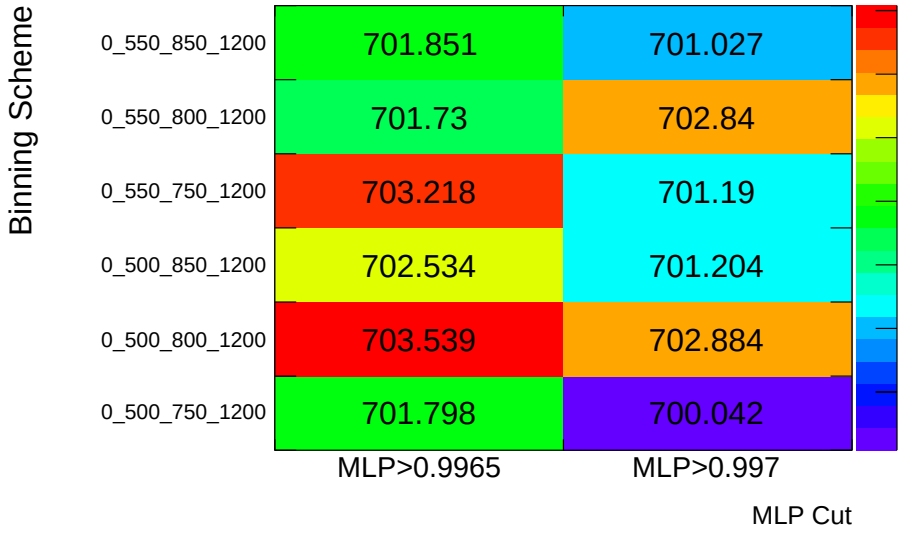


Figure 9.18: Expected limit on the T mass as a function of the cut on the NN output and the binning scheme (with quoted values given in GeV). The plot shows that in the considered range the expected limit is rather stable, i.e. changes only by very little for the tested final MLP cuts and binning schemes.

Sample	e	μ	$e+\mu$
$T\bar{T}$(500 GeV)	10.31±0.39	9.66±0.37	19.97±0.54
$T\bar{T}$(650 GeV)	5.11±0.11	5.01±0.11	10.12±0.15
$t\bar{t}$	0.380±0.163	0.549±0.167	0.929±0.233
W+jets	0.286±0.211	0.227±0.140	0.513±0.253
Z+jets	0.021±0.015	0±0	0.021±0.015
single top	0.363±0.171	0.671±0.420	1.03±0.45
di-boson	0.001±0.001	0±0	0.001±0.001
Multi-jet	0.096±0.383	0.505±0.286	0.601±0.478
Total prediction	1.148±0.497	1.95±0.55	3.10±0.74
Data	3	8	11

Table 9.15: $MLP > 0.9965$ selection event yields with statistical uncertainties for integrated luminosity of 4.7fb^{-1}.

9.6 Systematic Uncertainties

Various systematic uncertainties are assigned to the expected distributions to account for different uncertainties arising from the event simulation and detector modelling. The derivation and treatment of the sources of systematic uncertainties used in this analysis are explained in Sections 9.6.1 – 9.6.4. The resulting systematic uncertainties on the Tight and NN final selections are summarised in 9.6.5.

9.6.1 Luminosity

The systematic uncertainty on the integrated luminosity is estimated to be 3.9% using the Van der Meer scans [207]. This uncertainty affects the overall yield and is applied to all simulated components.

9.6.2 Object Reconstruction Uncertainties

In this section, the uncertainties affecting the reconstruction of the physics objects used in the analysis are explained. Those associated with the reconstruction and identification of leptons are described in Section 9.6.2.1, while those affecting jets are explained in Section 9.6.2.2.

9.6.2.1 Lepton Uncertainties

Lepton trigger, reconstruction and identification The differences in lepton trigger, reconstruction and identifications efficiency between simulation and data are corrected for in simulated samples by applying the scale factors derived from the T&P method [208] using $Z \to \ell\ell$ ($l = e$, μ) data. These scale factors are assigned uncertainties. For each source of uncertainty the quadratic sums of the statistical and systematic uncertainties on the corresponding scale factor is taken as the overall systematic uncertainty. Both leptons are assigned a per lepton uncertainty of $1 - 1.5\%$ [142]. These uncertainty sources affect only the overall yields.

Lepton momentum scale and resolution The agreement between data and simulation in momentum scale and resolution of leptons is studied using the di-lepton mass distribution in $Z \to \ell\ell$ and the J/ψ data. For electrons, E/p studies have been performed on $W \to e\nu$ data in addition. Small discrepancies are corrected for by corresponding correction factors. The electron energy scale is exceptionally applied to data in all regions and to MC only in the crack region. All other corrections are applied to MC only. The uncertainties on the these correction factors are found to be at the sub-percent level and thus neglected in this analysis.

9.6.2.2 Jet Uncertainties

Jet reconstruction efficiency The jet reconstruction efficiency agrees between data and simulation within 0.4% for jets of momenta below 35 GeV and within uncertainty for the jets of higher momenta. Given that most of the jets used in this analysis are well above 35 GeV, the systematic uncertainty assigned to efficiency is very small and is neglected in this analysis.

Jet vertex fraction efficiency In this analysis jets are required to satisfy $|JVF| > 0.75$. The efficiency of this cut is measured in $Z \to \ell\ell + 1$-jet data, using specific selections to separate the events containing hard-scatter jets from those containing mainly pile-up jets. This efficiency is a function of the p_{T} of the jet and varies $89-99\%$ for hard-scatter jets and $2-11\%$ for pile-up jets. For both types of events, efficiency and inefficiency scale factors are derived to correct for the data/MC discrepancies. The per-event correction weight is computed as the product of all per-jet scale factors and applied to MC. The effect of this uncertainty on the normalisation of signal and backgrounds is estimated to be $2.5-3.5\%$.

Jet energy scale The jet energy scale (JES) and the uncertainty assigned to it have been derived combining information from test-beam data, LHC collision data and simulation. The JES uncertainty is composed of 16 individual uncorrelated sources which is supplied by the so-called `MultijetJESUncertaintyProvider` tool [209]. In this analysis the corresponding 16 sources of uncertainty are added in quadrature to a single per-jet JES uncertainty scale factor. All jet-related kinematic variables as well as the missing transverse energy are recomputed accordingly. The JES uncertainty ranges $1-1.5\%$ for jets with the p_{T} ranging from ~ 50 GeV to ~ 1 TeV and reaches up to almost 4% for the jets with the p_{T} outside of that range. This uncertainty source affects both the shape and the normalisation of the kinematic distributions including the discriminant.

Jet energy resolution The jet energy resolution (JER) is measured in di-jet events with two different in-situ methods: di-jet balance method [210] and bisector method [211], which agree within uncertainties as explained in [212]. Both methods rely on the measurement of the fractional p_{T} resolution, $\sigma(p_{\mathrm{T}})/p_{\mathrm{T}}$, which at a fixed rapidity equals to the fractional jet energy resolution $\sigma(E)/E$. The p_{T} and rapidity dependent fractional p_{T} resolution $\sigma(p_{\mathrm{T}}, \eta)$ is provided by the `JERUncertaintyProvider` tool [213]. A systematic variation is constructed by smearing the jet p_{T} spectrum. The p_{T} of every jet is scaled by a factor $1 + r$, where r is a random number drawn from a Gaussian distribution with a mean of 0 and a width equal to the corresponding $\sigma(p_{\mathrm{T}}, \eta)$. This variation is by definition one-sided, corresponding to an up-variation. The corresponding down-variation is constructed by symmetrising the shift of the up-variation with

respect to the nominal distribution of the discriminant. Typical JER uncertainties range from below 10% for jets with $p_{\mathrm{T}} > 200$ GeV to $\sim 20\%$ for jets with $p_{\mathrm{T}} < 40$ GeV (Figure 5.9).

Jet mass scale and resolution The jet mass variable is used in the W_{had}-selection in the BOOSTED channel. Uncertainties on the jet mass are estimated using an in-situ comparison of the calorimeter jets to their associated tracks in several MC samples using various generators, hadronic shower models and amounts of dead material [214]. The estimation of these uncertainties used in this analysis is that which was available at the time when analysis was optimised. This estimation was derived for the anti-k_t jets with $R = 1.0$ using an older version of the reconstruction software than that used in this analysis, without the refined jet calibration. Compared with a later estimation, performed with the jet definition and reconstruction software used in this analysis, the uncertainties used in this analysis are twice larger, and thus very conservative. The per-jet systematic uncertainty on the jet mass scale (JMS) is taken to be 4.5% for jets with $p_{\mathrm{T}} < 400$ GeV and 6.0% otherwise. An additional uncertainty of 1% is added in quadrature, to account for mis-modelling of the pile-up which is not fully covered by the JMS uncertainty estimation used. The jet mass resolution (JMR) the uncertainty is taken to be 20%, regardless of the jet p_{T}. It is symmetrised in the same manner as described for JER. The JMS and JMR are treated as uncorrelated with JES and JER.

Flavour Tagging Uncertainties In this analysis, flavour tagging relies on the MV1 tagger at the operating point of 70% efficiency for b-jets. This efficiency, as well as the efficiencies for other flavours, is measured from data and is corrected in MC by dedicated p_{T} dependent scale factors. These scale factors are assigned systematic uncertainties in the range of $6\% - 20\%$ for b-jets, $12\% - 22\%$ for c-jets and $\sim 6\%$ for light jets. The scale factors and the corresponding uncertainties are supplied by `BTaggingCalibrationDataInterface` [215]. For b-jets with $p_{\mathrm{T}} >$ 200 GeV, scale factors are taken as those from the last p_{T} bin (140-200 GeV) available from the accurate measurement. The following additional errors are added for the extrapolation: 0.11 for jets with $p_{\mathrm{T}} \in$ (200-300 GeV) and 0.12 for jets above that. Similarly, for c-jets the uncertainties of 0.16 (200-300 GeV) and 0.17 (300-500 GeV) are added in quadrature to the corresponding SF values from the bin (90-140 GeV).

9.6.3 Normalisation Uncertainties

In this section, the estimation of the largest among the systematic uncertainties on the normalisation of the background components is presented.

9.6.3.1 MC Cross-section Uncertainties

The simulated samples are assigned a systematic uncertainty on their theoretical production cross sections. The uncertainties on the signal samples are given in Table 7.2 for each mass point separately. The $t\bar{t}$ production cross section uncertainty is estimated at approximate NNLO to be $+9.9\% - 10.7\%$ using HATHOR tool [192]. The single top and di-boson cross sections have been assigned $+4.7\%/-3.7\%$ [199] and $\pm 5\%$ [216] uncertainty respectively. Z+jets cross sections suffer rather large uncertainties given that they are obtained from the LO calculation

using ALPGEN. The Z+jets normalisation uncertainty is estimated as ±4% uncertainty on the inclusive cross section added in quadrature to 24% for each additional jet.

9.6.3.2 W+jets Normalisation Uncertainty

The overall W+jets event yield is affected by an $\sim$ 13% systematic uncertainty originating from various sources assigned to the data-driven method used to estimate it. However, for this analysis which mostly deals with rather hard jets, a more conservative uncertainty of ±48% has been chosen following the Top group recommendation for simulation-based normalisation uncertainty [142].

9.6.3.3 Multijets Normalisation Uncertainty

The systematic uncertainty on the QCD multi-jet normalisation described in Section 7.5.1 arises from variations in the QCD estimations using different control regions and calibrations of the method. It is estimated to be ±50% on the sample passing the $t\bar{t}$ pre-selection in the electron channel [142]. In this analysis, the same systematic uncertainty is conservatively used in the muon channel as well (in which the uncertainty does not exceed ±40%). In addition to this uncertainty, the QCD normalisation is affected by a statistical uncertainty of 70% due the the very poor statistics of the loose and tight data samples after the final selections.

9.6.4 Modelling Uncertainties

9.6.4.1 $t\bar{t}$ Modelling

Evaluation of the $t\bar{t}$ modelling uncertainties is studied on the specialised simulated samples explained in this section. These samples are processed with the fast detector simulation AFII and summarised in Table 7.4.

Parton Shower The nominal $t\bar{t}$ sample is generated with MC@NLO, the event generator including full NLO QCD corrections for partonic hard sub-processes. According to the studies concerning $t\bar{t}$ events, kinematic distributions agree reasonably between LO and NLO modelling except for the p_{T} of the $t\bar{t}$ system for which the initial and final state radiation (ISR and FSR) generated at LO are softer than if generated at NLO. A systematic uncertainty on this particular issue is estimated from the dedicated samples generated with AcerMC and interfaced to Pythia with varied parameters. Half the relative difference between the maximal and minimal amount of ISR and FSR is taken as a variation which is then propagated to the nominal sample in a symmetric way.

Fragmentation The magnitude of the variation due to the inaccuracies of the fragmentation models has been studied on samples generated with the same parton level generator, Powheg, interfaced to different two hadronisation models, Herwig and Pythia and processed in AFII. The relative difference between the two is symmetrised and propagated to the nominal sample.

Event generation To estimate the systematic effect due to inaccuracies in parton level modelling, kinematic distributions produced by the MC@NLO are compared to the corresponding distributions from another NLO generator, POWHEG, interfaced to the same hadronisation model, HERWIG. The comparison is done with the samples processed through the AFII simulation and their relative difference is symmetrised and propagated to the nominal sample processed by the full simulation.

9.6.5 Systematic Effects on the Final Selection

The effects of the individual systematic uncertainties on the discriminant for each component of the samples after the final cut-based and NN-based selections are summarised in the Tables 9.16 and 9.17. Since after the final selections all backgrounds other than $t\bar{t}$ deliver rather poor statistics, they are merged into a single sample here called small backgrounds. For each component of the sample Table 9.18 shows which source of uncertainty affects the normalisation only and which also the shape of the template in question.

Source of uncertainty	T 500 GeV	T 650 GeV	$t\bar{t}$	smallbg
Luminosity	+3.9/-3.9 %	+3.9/-3.9 %	+3.9/-3.9 %	+3.9/-3.9 %
e identification SF	+1.1/-1.1 %	+1.1/-1.1 %	+0.9/-0.9 %	+0.4/-0.4 %
e reconstruction SF	+0.4/-0.4 %	+0.4/-0.4 %	+0.4/-0.4 %	+0.2/-0.2 %
e trigger SF	+0.3/-0.3 %	+0.3/-0.3 %	+0.2/-0.2 %	+0.1/-0.1 %
μ identification SF	+0.4/-0.4 %	+0.4/-0.4 %	+0.4/-0.4 %	+0.1/-0.1 %
μ reconstruction SF	+0.2/-0.2 %	+0.2/-0.2 %	+0.2/-0.2 %	+0.1/-0.1 %
μ trigger SF	+0.7/-0.7 %	+0.7/-0.7 %	+0.7/-0.7 %	+0.2/-0.2 %
JER	+2.2/-2.2 %	+1.3/-1.3 %	+13.2/-12.1 %	+3.2/-3.2 %
JES	+5.6/-7.0 %	+2.2/-3.2 %	+28.3/-23.5 %	+9.9/-10. % 2
JMR	+3.2/-3.2 %	+5.3/-5.3 %	+3.5/-3.1 %	+5.5/-5.5 %
JMS	+0.9/-0.8 %	+0.8/-1.7 %	+4.3/-3.4 %	+0.0/+0.0 %
JVF	+1.6/-1.7 %	+1.8/-1.8 %	+2.0/-2.0 %	+0.6/-0.6 %
b-tagging SF	+5.4/-8.2 %	+5.6/-8.7 %	+6.9/-7.8 %	+2.3/-2.9 %
c-tagging SF	+0.3/-0.3 %	+0.3/-0.3 %	+1.0/-1.1 %	+0.0/-0.0 %
mistag SF	+0.1/-0.1 %	+0.1/-0.1 %	+0.3/-0.3 %	+0.3/-0.3 %
MC generation	–	–	+83.2/-80.5 %	–
Fragmentation	–	–	+0.7/-0.7 %	–
Parton shower	–	–	+10.4/-10.4 %	–
$t\bar{t}$ cross section	–	–	+9.9/-10.7 %	–
QCD cross section statistical	–	–	–	+12.2/-12.2 %
QCD cross section systematic	–	–	–	+13.5/-13.5 %
$W + b\bar{b}/c\bar{c}$ normalisation	–	–	–	+7.6/-8.0 %
$W + b\bar{b}/c\bar{c}$ scale	–	–	–	+4.6/-4.8 %
$W + c$+jets scale	–	–	–	+0.6/-0.5 %
W+light jets normalisation	–	–	–	+5.1/-5.1 %
W+jets charge asymmetry	–	–	–	+6.8/-5.4 %
Z+jets cross section	–	–	–	+53.0/-5.4 %
Di-boson cross section	–	–	–	+0.0/-0.0 %
Single top cross section	–	–	–	+1.3/-1.0 %
Total	+9.8/-12.3 %	+9.3/-11.8 %	+90.6/-86.7 %	+58.8/-25.9 %

Table 9.16: Relative variation of the normalisation of the reconstructed T-quark mass due to individual and total uncertainties after the Tight selection.

Source of uncertainty	T 500 GeV	T 650 GeV	$t\bar{t}$	smallbg
Luminosity	+3.9/-3.9 %	+3.9/-3.9 %	+3.9/-3.9 %	+3.9/-3.9 %
e identification SF	+1.1/-1.1 %	+1.1/-1.1 %	+0.9/-0.9 %	+0.4/-0.4 %
e reconstruction SF	+0.5/-0.5 %	+0.5/-0.5 %	+0.4/-0.4 %	+0.1/-0.1 %
e trigger SF	+0.3/-0.3 %	+0.3/-0.3 %	+0.2/-0.2 %	+0.1/-0.1 %
μ identification SF	+0.4/-0.4 %	+0.4/-0.4 %	+0.4/-0.4 %	+0.2/-0.2 %
μ reconstruction SF	+0.2/-0.2 %	+0.2/-0.2 %	+0.2/-0.2 %	+0.1/-0.1 %
μ trigger SF	+0.7/-0.7 %	+0.7/-0.7 %	+0.7/-0.7 %	+0.3/-0.3 %
JER	+0.8/-0.8 %	+1.2/-1.2 %	+11.1/-11.1 %	+12.4/-12.4 %
JES	+14.4/-13.3 %	+4.8/-6.8 %	+29.9/-26.3 %	+26.8/-14.3 %
JMR	+3.3/-3.3 %	+5.4/-5.4 %	+2.9/-2.9 %	+11.5/-11.5 %
JMS	+1.1/-1.0 %	+1.3/-1.4 %	+4.7/-5.7 %	+0.0/+0.0 %
JVF	+1.7/-1.8 %	+1.8/-1.8 %	+3.2/-2.7 %	+0.8/-0.9 %
b-tagging SF	+5.6/-8.9 %	+5.6/-8.9 %	+8.4/-9.8 %	+3.0/-4.1 %
c-tagging SF	+0.3/-0.3 %	+0.3/-0.3 %	+1.1/-1.1 %	+0.1/-0.1 %
mistag SF	+0.1/-0.1 %	+0.1/-0.1 %	+0.3/-0.3 %	+0.4/-0.4 %
MC generation	–	–	+71.0/-65.2 %	–
Fragmentation	–	–	+13.4/+4.3 %	–
Parton shower	–	–	+1.5/-1.5 %	–
$t\bar{t}$ cross section	–	–	+9.9/-10.7 %	–
QCD cross section statistical	–	–	–	+22.0/-22.0 %
QCD cross section systematic	–	–	–	+13.9/-13.9 %
$W + b\bar{b}/c\bar{c}$+jets normalisation	–	–	–	+4.6/-4.9 %
$W + b\bar{b}/c\bar{c}$+jets scale	–	–	–	+2.1/-2.2 %
$W + c$+jets scale	–	–	–	+0.2/-0.2 %
W+light jets normalisation	–	–	–	+3.2/-3.1 %
W+jets charge asymmetry	–	–	–	+2.9/-2.5 %
Z+jets cross section	–	–	–	+7.3/-0.9 %
Di-boson cross section	–	–	–	+0.0/-0.0 %
Single top cross section	–	–	–	+2.2/-1.8 %
Total	+16.5/-17.0 %	+10.3/-13.4 %	+80.5/-73.2 %	+42.5/-35.4 %

Table 9.17: Relative variation of the normalisation of the reconstructed T-quark mass due to individual and total uncertainties after the NN selection.

In both final selections the $t\bar{t}$ component suffers the highest relative systematic uncertainties. The overall uncertainty on the $t\bar{t}$ component in the Tight selection is $^{+90.6}_{-86.7}$%, which makes $^{+3.53}_{-3.37}$% expected events and 75% of the total background uncertainty. It is dominated by the event generator uncertainty (labeled as MC generator in the tables) of $^{+83.2}_{-80.5}$% followed by the jet energy scale (JES) uncertainty of $^{+28.3}_{-23.5}$%. For the NN selection, the overall $t\bar{t}$ uncertainty is $^{+80.5}_{-73.2}$% which is equivalent to $^{+0.78}_{-0.65}$ expected events and makes almost 57% of the overall background uncertainty. The dominant up variation of +71% results from the MC modelling uncertainty.

9.7 Results

Using the statistical method explained in Section 6.3.1, data samples passing the Tight and the NN final selections are tested against the background only (b) and the signal plus background ($s + b$) hypotheses. The templates used for hypothesis testing are discussed in Section 9.7.1.

Source of uncertainty	T 500/650 GeV		$t\bar{t}$		smallbg	
	Norm	Shape	Norm	Shape	Norm	Shape
Luminosity	●	○	●	○	●	○
e identification SF	●	○	●	○	●	○
e reconstruction SF	●	○	●	○	●	○
e trigger SF	●	○	●	○	●	○
μ identification SF	●	○	●	○	●	○
μ reconstruction SF	●	○	●	○	●	○
μ trigger SF	●	○	●	○	●	○
JER	●	●	●	●	●	●
JES	●	●	●	●	●	●
JMR	●	●	●	●	●	●
JMS	●	●	●	●	●	●
JVF	●	●	●	●	●	●
b-tagging SF	●	●	●	●	●	●
c-tagging SF	●	●	●	●	●	●
mistag SF	●	●	●	●	●	●
MC generation	–	–	●	●/○	–	–
Fragmentation	–	–	●	●/○	–	–
Parton shower	–	–	●	●	–	–
$t\bar{t}$ cross section	–	–	●	○	–	–
QCD cross section statistical	–	–	–	–	●	○
QCD cross section systematic	–	–	–	–	●	○
$W + b\bar{b}/c\bar{c}$+jets normalisation	–	–	–	–	●	○
$W + b\bar{b}/c\bar{c}$+jets scale	–	–	–	–	●	○
$W + c$+jets scale	–	–	–	–	●	○
W+light jets normalisation	–	–	–	–	●	○
W+jets charge asymmetry	–	–	–	–	●	○
Z+jets cross section	–	–	–	–	●	○
Diboson cross section	–	–	–	–	●	○
Single top cross section	–	–	–	–	●	○

Table 9.18: Summary of the component-wise treatment of systematic uncertainties as normalisation and or shape uncertainties in both final selections. A full circle (●) stands for YES and an empty one (○) for NO. The event generator (MCGEN) and fragmentation (Frag) are considered to be shape uncertainties in the Tight selection and normalisation only in the NN selection.

The corresponding test statistics resulting from the pseudo experiments based on the presented templates are shown in Section 9.7.2. The discovery potential and the exclusion limits are presented in Sections 9.7.3 and 9.7.4 respectively. The analysis infrastructure is validated by comparing the outcomes of the cut-based analysis implemented here to those of the analysis presented in [206]. The sensitivity improvement acquired from the multi-variate selection is evaluated by comparing the exclusion limits achieved with these two final selections, as discussed in 9.7.4.

9.7.1 Templates

The M_{reco} distributions composed of the events passing the final selections are used as the hypothesis testing templates. This variable is chosen to be the final discriminant of the analysis because of its high sensitivity to the presence of resonant production of Wb pairs. The processes producing real Wb pairs, such as $T\bar{T}$ signal and $t\bar{t}$ background, are expected to have an enhancement in the M_{reco} distribution corresponding to the mass of the particle from which Wb pairs originate, i.e. T- and t-quark respectively. Thus, the M_{reco} distribution has a very good separation power. For an illustration, the separation of the M_{reco} distribution as well as the NN input variables after the Tight and NN final selections is presented in Table 9.19. The table shows that the M_{reco} has a better separation than the other variables presented. It further shows that the separation of the M_{reco} variable is better after the NN final selection than after the Tight selection, which illustrates the advantage of the NN selection.

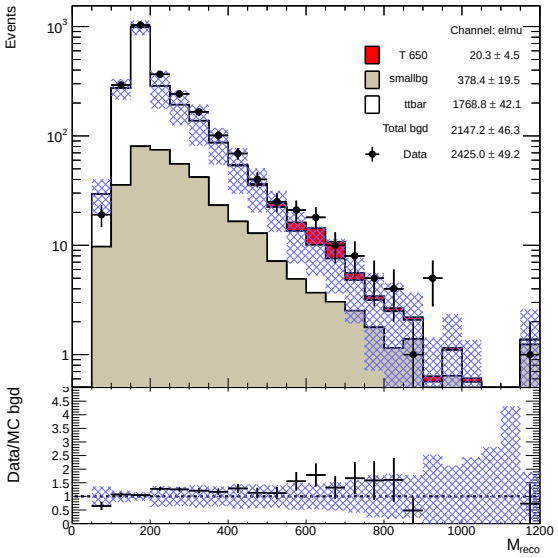


Figure 9.19: Data/MC comparison of the reconstructed T-quark mass distribution M_{reco} for the events passing the W_{had}-selection in the electron and muon channel combined.

As both final selections are extremely tight, resulting in very poor statistics of the selected samples, the templates for the electron and muon channels are merged and treated as a single template in the hypothesis testing procedure. For the same reason, the non-$t\bar{t}$ backgrounds, which deliver rather poor event yields and suffer large statistical uncertainties (Table 9.15), are merged into a single sample, called here small backgrounds (*small bkg*), while the dominant

	M_{reco}	$p_{\text{T}}^{b_{\text{lep}}}$	H_{T}	$\Delta R^{lb_{\text{lep}}}$	$\Delta R^{W_{\text{had}}b_{\text{had}}}$	$\Delta R^{W_{\text{had}}b_{\text{lep}}}$	$\Delta R^{l\nu}$	$\Delta R^{lb_{\text{lep}}}$	$p_{\text{T}}^{b_{\text{had}}}$
$\langle S^2\rangle_{Tight}$	0.355	0.108	0.338	0.123	0.112	0.123	0.112	0.143	0.237
$\langle S^2\rangle_{NN}$	0.395	0.384	0.378	0.356	0.338	0.292	0.239	0.224	0.137

Table 9.19: Separation power of the discriminant M_{reco} and the NN training variables between the combined simulated background and the T-signal with $m_T = 650$ GeV after the Tight and NN final selections.

$t\bar{t}$ background is treated separately. The M_{reco} distribution composed of events passing the W_{had}-selection for both electron and muon channels is shown in Figure 9.19.

The binning of the reconstructed T-quark mass distribution after both final selections has been optimised to provide the lowest possible expected CL_s value for the T-signal hypothesis with the mass point closest to the limit achievable with that selection (i.e. 650 GeV for the Tight, and 700 GeV for the NN selection). The last and the first bin of the M_{reco} distribution have been fixed in the way to avoid negative event yields of the $t\bar{t}$ background in these bins. For the range between the first and the last bin, the number of bins and the position of the bin edges have been varied until an optimal configuration has been achieved.

The CL_s values obtained after the Tight selection using the T-signal with a mass of 650 GeV for each tested binning scheme is shown in Figure 9.20. According to this figure, the best performing binning scheme of the M_{reco} distribution after the Tight selection is $(0, 500, 700, 1200)$ GeV. For the NN selection, the binning has been optimised simultaneously with the selection requirement on the NN response as explained in Section 9.5.2. The binning scheme $(0, 500, 800, 1200)$ GeV has been found to provides the best expected T-quark mass limit.

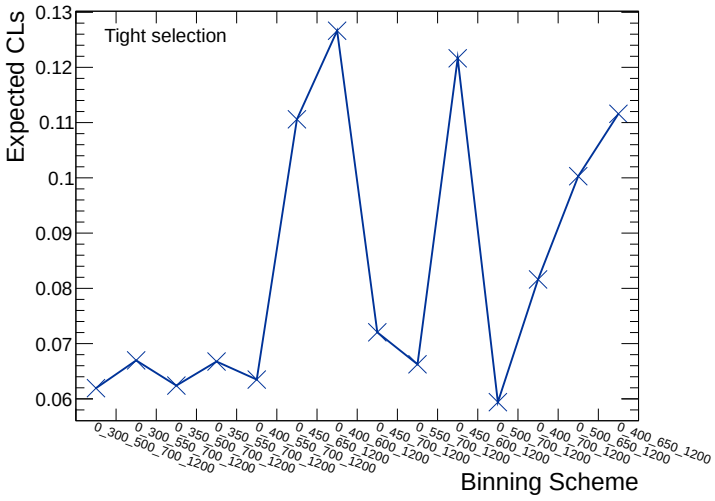


Figure 9.20: Expected CLs value for the T signal with the mass of 650 GeV as a function of the binning scheme for the Tight selection.

The M_{reco} templates used for the Tight and the NN selections showing total systematic uncertainty bands are shown in Figure 9.21. Comparing the two templates for the T-quark mass of 650 GeV (Figures 9.21 (e) and (f)), several differences become obvious. The NN selection yields a $\sim$ 25% higher signal acceptance, while achieving a background suppression which is stronger by an overall factor of $\sim$ 2.8. The NN selection is especially more efficient in removing the $t\bar{t}$ component from the low mass region. The NN selection accepts even $\sim$ 50% of the signal and only $\sim$ 0.15% of the backgrounds passing the W_{had}-selection, while the Tight selection accepts $\sim$ 40% of signal and $\sim$ 0.4% of backgrounds. However, while the Tight selection sees an overall excess of data events of factor $\sim$ 1.5 compared to the expectation, the NN selection sees even $\sim$ 3.55 times more data events than expected. Nonetheless, the data is compatible with the presence of the signal with the mass of 650 GeV only in the first bin of the discriminant. In the second bin data content is too low, and in the third too high compared to the $s + b$ hypothesis, which increases the sensitivity of the analysis. This also demonstrates the advantage of using binned discriminant rather than a single bin counting experiment. The corresponding event yields of the individual components including statistical and systematic uncertainties are presented in Table 9.20. The effects of individual systematic uncertainties on each component separately are summarised previously in Tables 9.16 and 9.17.

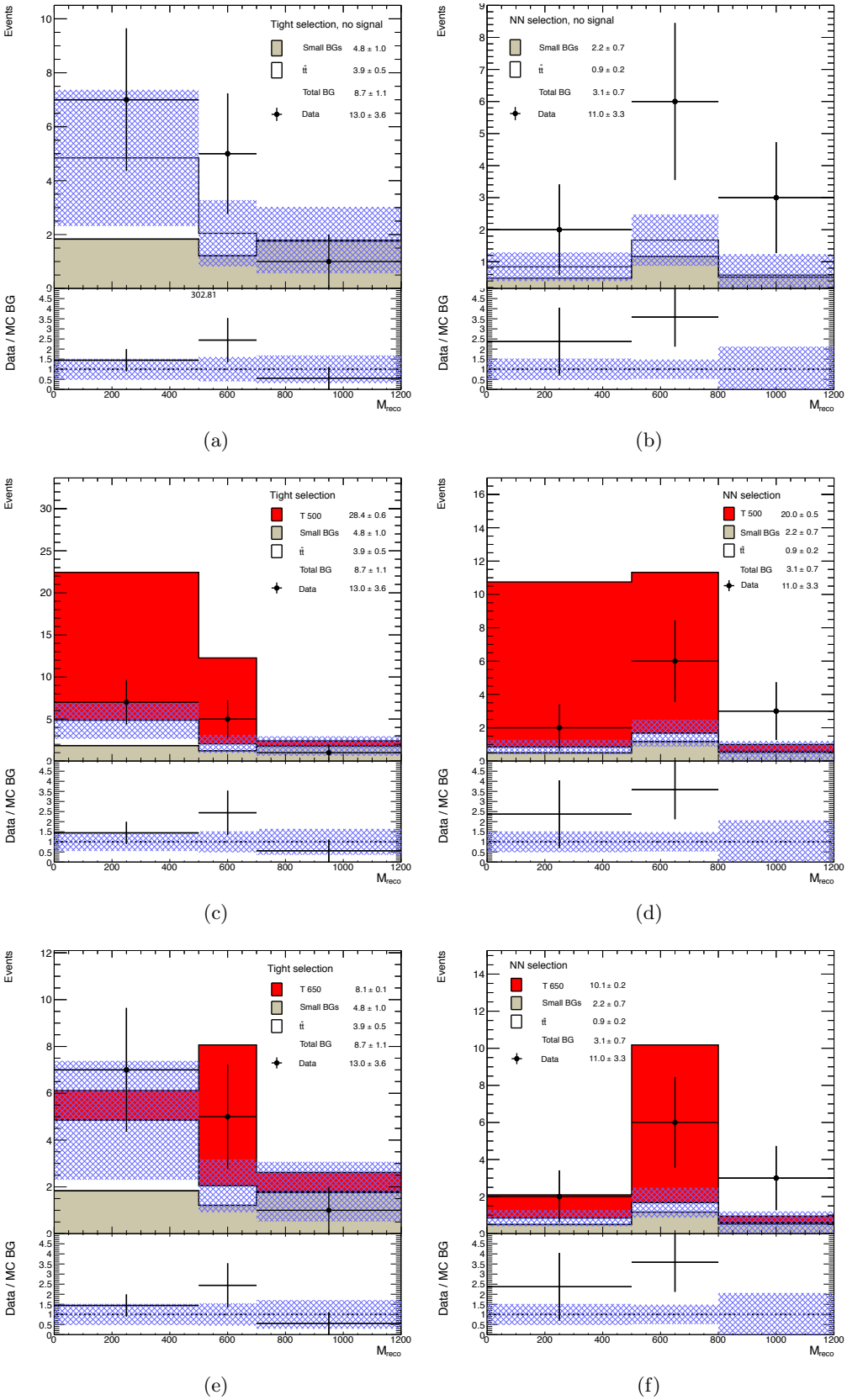


Figure 9.21: M_{reco} template for the Tight selection ((a), (c), and (e)) and the NN selection ((b), (d), and (f)). Figures (a) and (b) show the SM expectation divided into stacked $t\bar{t}$ (white) and small bkg (grey) components, as well as data (black points). Figures (c) and (d) include the T-quark signal with $m_T = 500$ GeV, while Figures (e and (f) include the signal with ($m_T = 650$ GeV), superimposed to the backgrounds.

	Tight	NN
T 500 GeV	$28.43 \pm 0.64\ ^{+2.78}_{-3.50}$	$19.97 \pm 0.54\ ^{+3.29}_{-3.40}$
T 650 GeV	$8.12 \pm 0.14\ ^{+0.75}_{-0.95}$	$10.12 \pm 0.15\ ^{+1.05}_{-1.35}$
$t\bar{t}$	$3.89 \pm 0.51\ ^{+3.53}_{-3.37}$	$0.93 \pm 0.23\ ^{+0.78}_{-0.65}$
Small bkg.	$4.79 \pm 1.03\ ^{+2.82}_{-1.24}$	$2.17 \pm 0.71\ ^{+0.92}_{-0.77}$
Total bkg.	$8.69 \pm 1.15\ ^{+4.67}_{-3.77}$	$3.10 \pm 0.74\ ^{+1.37}_{-1.12}$
Data	13	11

Table 9.20: Tight and NN selection event yields broken into individual components with statistical and systematic uncertainties respectively. Background event yields are quoted for $t\bar{t}$ and *small bkg* separately, as well as combined (Total bkg).

9.7.2 Test Statistic

The expected and observed CL_s values are determined from the LogLikelihood Ratio (LLR) distributions generated for the b-only and the $s+b$ hypotheses as explained in Section 6.3.1, with all systematic uncertainties included. In Figure 9.22 the resulting LLR_b (blue curve) and LLR_{s+b} (red curve) distributions are shown generated with 1M pseudo experiments (PEs) each. To illustrate the effect of the systematic uncertainties, the LLR distributions assuming the T-signal with the mass of 650 GeV for the samples passing the Tight and NN selection, are shown once produced without (Figures 9.22 (a) and (a)), and once with systematic uncertainties (Figures 9.22 (c) and (d)). The figures show that allowing for the variations of the template bin content in the PEs within the systematic uncertainties results in a broadening of the LLR distributions, which reduces the separation between the $s+b$ and b hypothesis test statistics, and thus weakens the sensitivity of the analysis.

The CL_b and CL_s values are computed following Equations 6.17 and 6.18. The LLR_{obs} from these equations stands for the LLR value obtained by using data instead of pseudo data. LLR_{obs} is shown in the plots in Figure 9.22 by a black vertical line. For calculating the expected CL_b and CL_s the LLR_{obs} is substituted by the median of the LLR_b distribution which is indicated by a vertical blue line in the plots. For the NN selection the LLR_b and LLR_{s+b} distributions are more separated than in case of the Tight selection which results from the difference in the separation power of the corresponding M_{reco} templates. The broader distributions for the NN selection result from the somewhat broader systematic error bands compared to the Tight selection.

In Figure 9.23 the median expected LLR_b and LLR_{s+b} values with the corresponding 1 and 2 standard deviation (s.d) widths are plotted as a function of the T-quark mass for the Tight (left) and the NN selection (right). The LLR_b and LLR_{s+b} overlap more for the Tight selection, suggesting a better separation power of the NN selection throughout the considered mass range. The observed LLR_{obs} as a function of the T-quark mass is shown by black solid lines in both plots. For the Tight selection, LLR_{obs} is compatible with the median expectation of the b-hypothesis within 1 σ, although slightly lower, i.e. pulling towards the $s+b$ hypothesis. For the NN selection, the LLR_{obs} follows the shape of the b-hypothesis median expectation within $1.1-1.3$ σ, and becomes more compatible with the $s+b$ hypothesis above the T-quark mass of 660 GeV. However, around this region the 2 s.d. bands of both selections largely overlap, and the 1 s.d. bands start overlapping as well, so that no clear statement can be made. In addition, the solid compatibility of the observation with the b-hypothesis across the considered

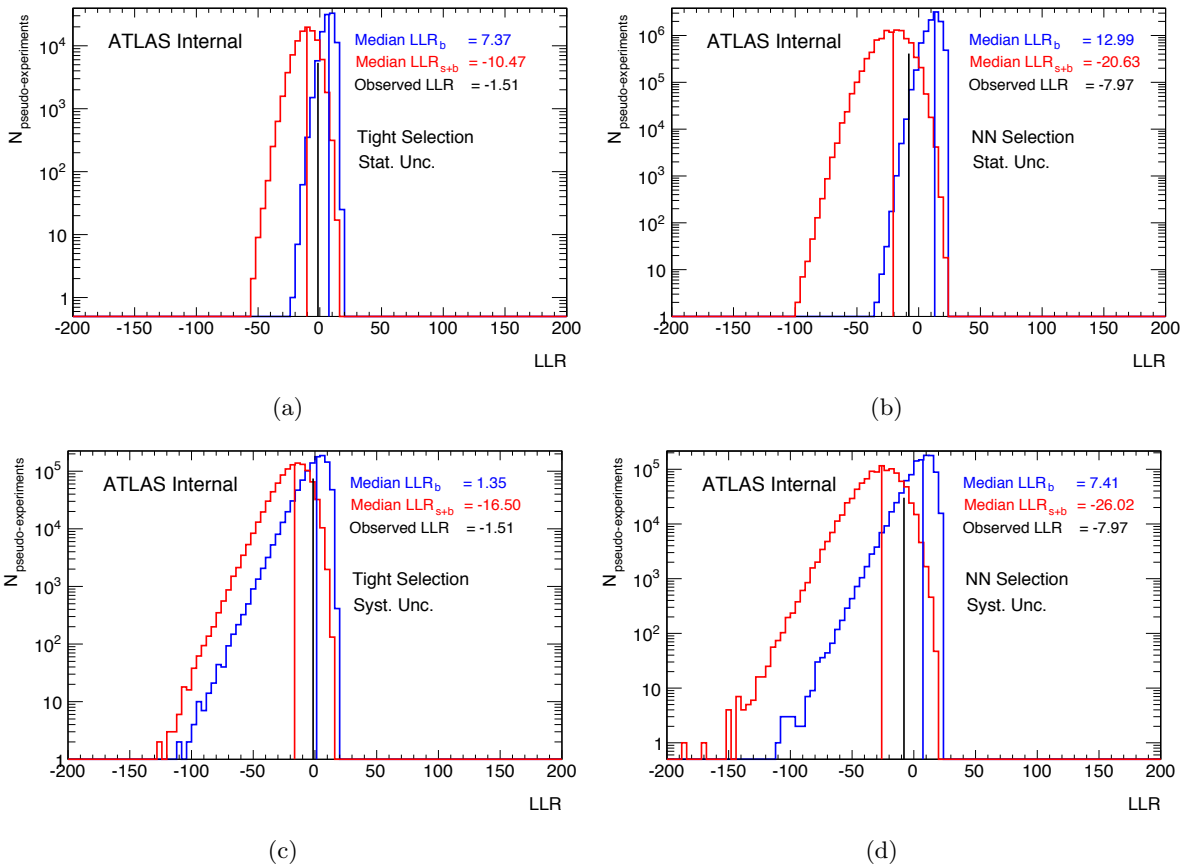


Figure 9.22: LLR distributions of the $s+b$ (red) and b (blue) hypotheses obtained with 1M PEs each for $m_T = 650$ GeV for the Tight selection (Figures (a) and (c)) and NN selections (Figures (b) and (d)), based on statistical uncertainties only (Figures (a) and (b)), or including the systematic uncertainties (Figures (c) and (d)).

mass range suggests that the observation is rather a fluctuation of the SM background than an indication to the presence of the signal.

The LLR distributions are the broader the more systematic uncertainties are allowed for. Accordingly, the CL_s values increase when the LLR distributions broaden. Following this argument, the treatment of the systematic uncertainties can be validated by ensuring that the expected CL_s value increases steadily as the various systematic uncertainties are added one by one. Figure 9.24 shows the median expected CL_s as a function of the T-quark mass for the successive inclusion of the systematic uncertainties listed in Section 9.6 to the hypothesis testing based on the Tight (Figures 9.24 (a) and (c)) and NN (Figures 9.24 (b) and (d)) final selection templates. For both selections, the full mass range is shown (upper row), as well as the range close to the exclusion limit is magnified (lower row). For T-quark masses lower than 600 GeV some irregularities are present due to large statistical fluctuations. When LLR_b and LLR_{s+b} distributions are well separated, as in case of low T-quark masses, especially if no or few systematic uncertainties are allowed for, very few pseudo events assuming the $s+b$ hypothesis are generated with LLR_{s+b} higher than the median of LLR_b, which results in a very high statistical uncertainties on the CL_{s+b}. Nevertheless the expected CL_s as a function of the T mass behaves as expected for T-quark masses above 600 GeV and in particular around the limit. From the comparison of the limit without systematic uncertainties (the black line in the

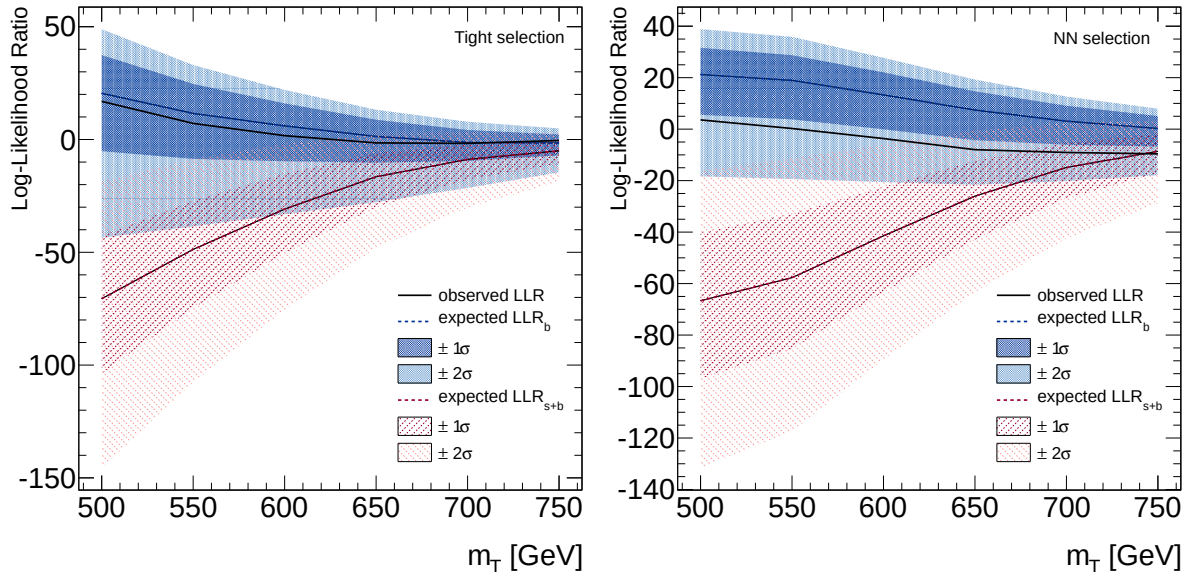


Figure 9.23: The observed and the median expected LLR as well as the 1 and 2 s. d. width for the $s+b$ (red) and b (blue) hypotheses as a function of the T-quark mass for the Tight selection (left) and the NN selection (right).

plots in Figure 9.24) and with all systematic uncertainties applied (the light grey line laying on the top of all the others in the plots), it turns out that the systematic uncertainties deteriorate the mass limit for ~ 40 GeV in both final selections.

9.7.3 Sensitivity Test

Based on the LLR distributions for all tested T-quark masses, the discovery potential of both selections has been studied using the CL_b method. The expected and observed signal significance (p-value) of the b-only hypothesis (1-CL_b) is shown in Figure 9.25 for both final selections. The observed sensitivity is well within 2σ for the both Tight and NN selection. The median expected sensitivity is between 2 σ and 3 σ below, and within 2 σ above 630 GeV for the Tight selection. The expected uncertainty of 2 σ on 1-CL_b is still far above the 5σ discovery threshold ($\sim 6 \cdot 10^{-7}$) in the full mass range considered. The median expected signal significance is above 3σ for the T masses below 600 GeV, between 2 and 3σ in the range $600-700$ GeV and within 2σ above 700 GeV for the NN selection. The 2 σ uncertainty band reaches almost down to 5σ for the T-quark mass of 500 GeV and stays slightly above 3σ up to almost 750 GeV. This shows that the NN selection has a better expected sensitivity than the Tight selection. For both selections, however, the observed sensitivity is within the statistical fluctuation.

This low observed significance despite of a data excess of a factor of ~ 3.55 over the expectation is due to the shape difference between data and the $s+b$ hypothesis expectation. In the case where the shape difference is neglected, sensitivity is higher, but still not significant, as shown in Appendix C.10.

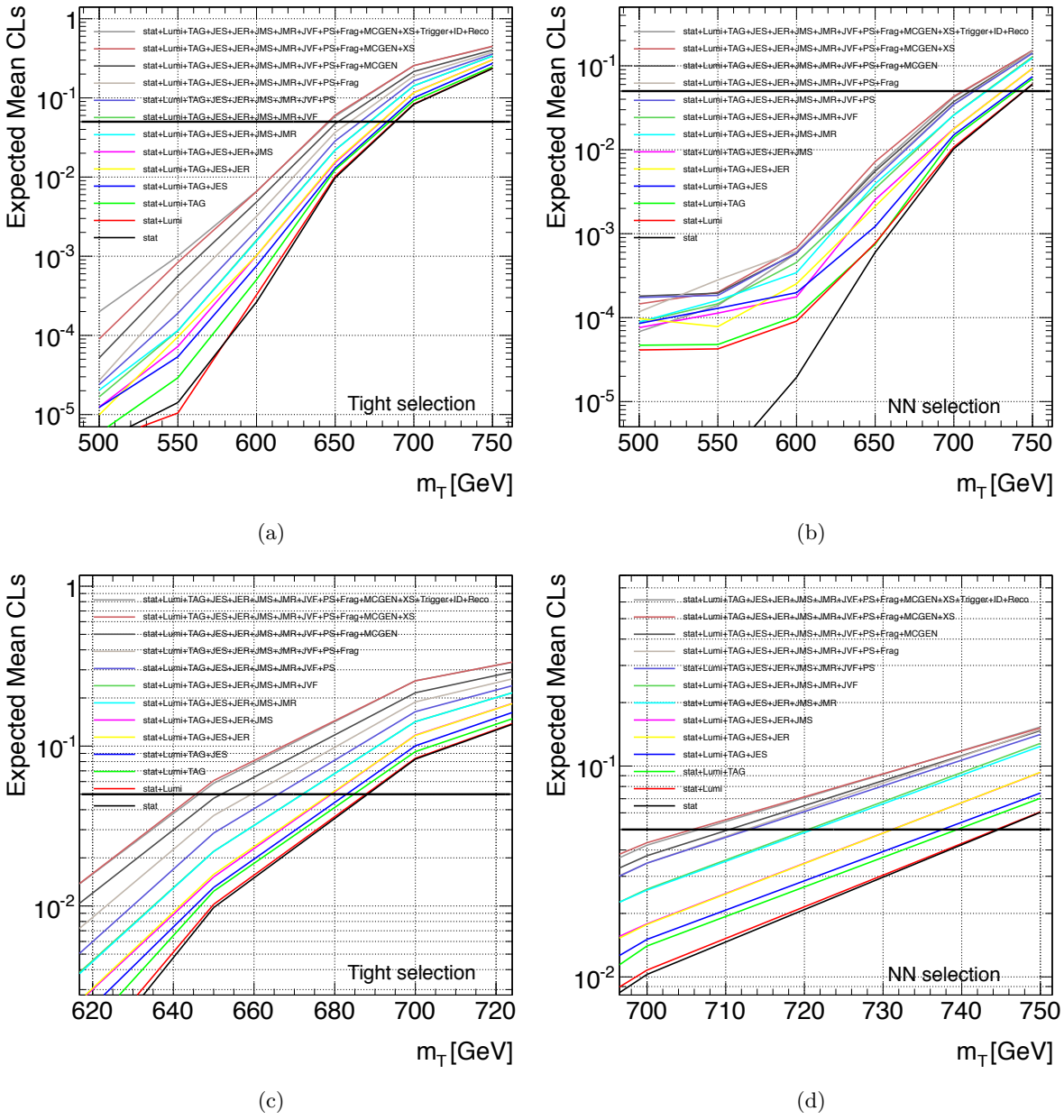


Figure 9.24: Propagation of the the expected CL_s as a function of the T-quark mass, as systematic uncertainty are applied one by one for the Tight selection (Figure (a)) and the NN selection (Figure (b)). The plots in the bottom row show the corresponding magnification focused on the mass range close to the expected limit.

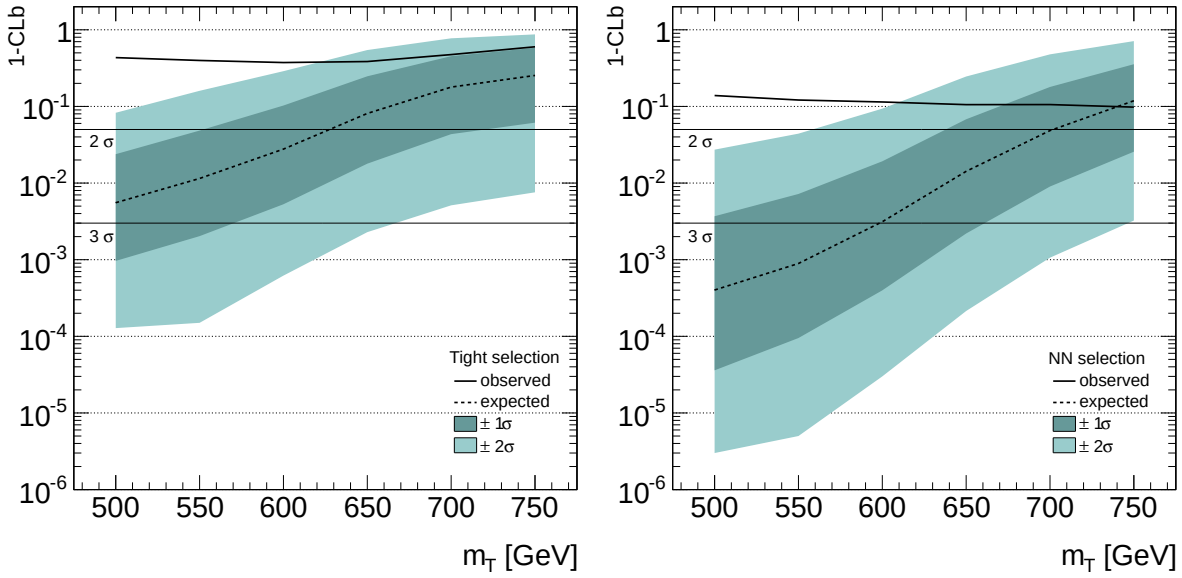


Figure 9.25: Expected and observed CL_b values as a function of the T-quark mass assuming the theoretical cross section for the signal ($R = 1$) obtained from the LLR test statistic generated from 1M pseudo-experiments based on the M_{reco} discriminant for the Tight selection (left) and the NN selection (right).

9.7.4 Final Limits and Discussion

The compatibility of the data with the vector-like T-quark is measured in terms of the CL_s value defined in Section 6.3.1 and further explained in Section 9.7.2. Low CL_s values correspond to a low compatibility. The signal mass for which $CL_s = 0.05$, assuming its theoretical production cross section, is considered excluded at 95% C.L. In Figure 9.26, the observed and the median expected CL_s values are plotted as a function of the T-quark mass. The observed (expected) lower bound on the T-quark mass is extracted by a linear interpolation between the closest mass points providing the CL_s values below and above the 95% C.L.

For each T-quark mass point, the excluded production cross section is found by scanning a range of scale factors (SF) and finding that which provides a $CL_s = 0.05$. These scans are performed using 100k PEs for the b and $s + b$ hypotheses each, and are in addition used to find the 1 and 2 σ error bands on the expected excluded cross section. The expected and observed excluded cross sections as a function of the T mass obtained using the Tight and NN final selections, are shown in Figure 9.27. The 1 σ and 2 σ systematic error bands are assigned to the expectation. As a comparison, the expected and observed cross section lines with 1 σ systematic error bands from [206] are shown in the same plot. The measured and expected excluded cross sections with 1 and 2 σ bands are summarised in Tables 9.21–9.22 for both final selections.

The 95% C.L. lower bounds on the T mass using the analysis infrastructure as close as possible to that published in [206] are found to be $m_T > 618.6\,(640.3)$ GeV observed (expected). The corresponding limits in the published analysis are $m_T > 656\,(638)$ GeV observed (expected).

The differences in the observed results between this analysis and [206] result from the differences in the expected and observed event yields. The expected background yield in this analysis is lower by $\sim 24\%$ mostly due to the use of updated multi-jet background estimates which provides

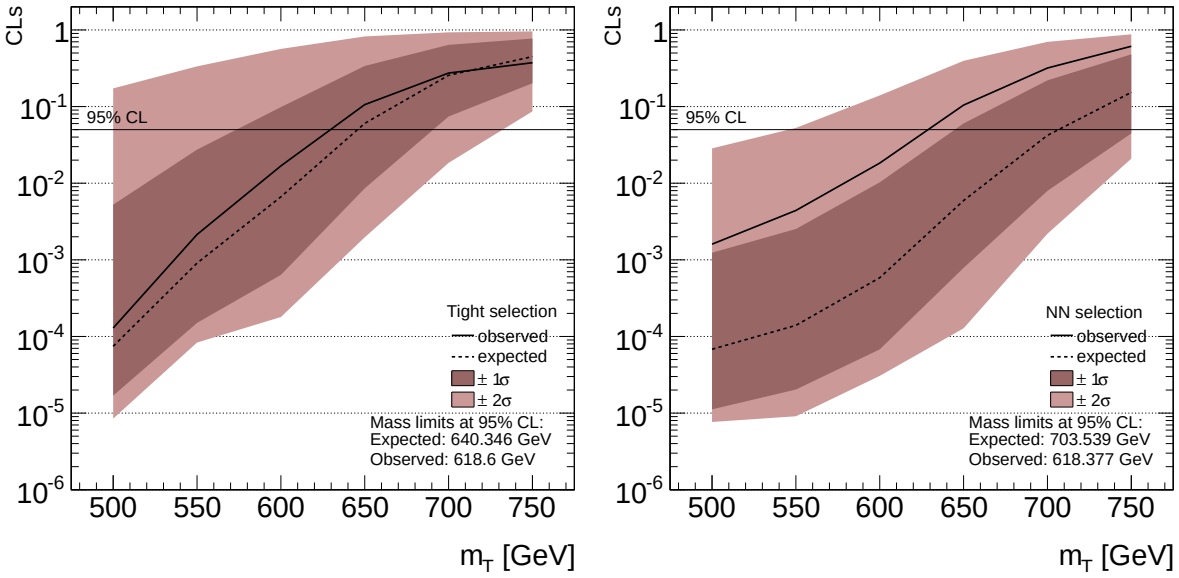


Figure 9.26: Expected and observed CL_s values as a function of the T-quark mass assuming the theoretical cross section for the signal ($R = 1$) obtained from the LLR test statistic generated from 1M pseudo-experiments based on the M_{reco} discriminant for the Tight selection (left) and the NN selection (right).

T mass (GeV)	500	550	600	650	700	750
Observed (pb)	0.181	0.123	0.088	0.074	0.062	0.049
Expected (pb)	0.122	0.081	0.057	0.052	0.051	0.047
Expected + 1 s.d	0.076	0.051	0.037	0.034	0.034	0.031
Expected − 1 s.d	0.201	0.135	0.097	0.087	0.084	0.075
Expected + 2 s.d	0.056	0.038	0.029	0.027	0.025	0.022
Expected − 2 s.d	0.339	0.226	0.162	0.145	0.138	0.121

Table 9.21: Observed and expected cross sections 95% C.L. limits vs T mass for the Tight selection.

T mass (GeV)	500	550	600	650	700	750
Observed (pb)	0.188	0.115	0.077	0.061	0.052	0.050
Expected (pb)	0.112	0.063	0.042	0.033	0.028	0.025
Expected + 1 s.d	0.073	0.041	0.028	0.022	0.019	0.016
Expected − 1 s.d	0.181	0.108	0.069	0.054	0.045	0.041
Expected + 2 s.d	0.056	0.033	0.022	0.016	0.015	0.013
Expected − 2 s.d	0.300	0.175	0.115	0.090	0.075	0.070

Table 9.22: Observed and expected cross sections 95% C.L. limits vs T mass for the NN selection.

~ 2.9 times lower number of events than in [206] (Table 9.16). At the same time, the observed number of events is higher in this analysis due to the difference in the b-jet identification. While in this analysis the b-jets are identified after the pre-selection, in the original analysis it is done before the $p_{\mathrm{T}} > 25$ GeV requirement for jets. The downward variation in the expectation and upward in the observation result in a slightly improved expected and a significantly weakened observed T-quark mass limit with respect to the result published in [206].

The observed and expected limits obtained with the NN selection and binning are 618.4 GeV and 703.5 GeV, respectively. The improvement of the expected sensitivity of ~ 60 GeV corresponding to a lower expected cross section by a factor ~ 1.56 at the T mass of 650 GeV arises from a higher sample purity resulting from a powerful NN-based signal and background identification. For the T-quark sample with the mass of 650 GeV in the Tight selection, a signal to background ration of 0.93 is expected, while the NN selection yields 3.26 times more signal than background. The deterioration of the observed exclusion limit is however twofold. When compared to the result from [206], the same argument regarding the differences in the b-jet identification discussed above applies here. Compared to the Tight selection result of this analysis, the NN selection sees a much larger data excess. Figure 9.16 shows that in the electron channel that excess is a large upward fluctuation of the data in the last bin of the NN response (MLP distribution). This excess can be understood as a consequence of a general feature of the electron channel, since already after the W_{had}-selection an overall data excess of $\sim 11\%$ is seen, while at the same time all NN input variables have a good agreement between data and simulation (Appendix C.4.9). In the muon channel, however, the poor modelling of the high tail of the H_{T} distribution causes the data overshoot in the tail of the NN response, as shown in Figure 9.16, which results in a sizable data excess in the signal region. The excess is however only slightly above 1σ of the expected exclusion cross section, as can be seen in the bottom plot of Figure 9.27.

9.8 Conclusion

A search has been presented for the pair-production of a heavy quark T decaying exclusively into a W-boson and a b-quark based on the vector-like T as a benchmark model. The search is focused on the decay mode in which one W-boson decays into a lepton and a neutrino and the other W-boson decays into a pair of light quarks, which results in final states with one isolated high momentum electron or muon, at least three jets, and significant missing transverse energy. The search is carried out with 4.7 fb^{-1} of pp collision data at $\sqrt{s} = 7$ TeV collected by the ATLAS detector in 2011. The analysis employs two approaches for the final selection: a so-called cut-based approach, based on selection requirements on a set of kinematic variables, and a so-called NN-based approach, based on the multi-variate NN technique introduced in Section 6.2. In this analysis it has been demonstrated that an analysis based on an NN discriminator achieves a better separation between signal and background and consequently a better expected sensitivity compared with an equivalent cut-based selection. The SM background to this search consists of $t\bar{t}$, single top, QCD multi-jet, and di-boson (WW, WZ, ZZ) production, as well as the W/Z-boson production in association with jets. The multi-jet background was estimated using the data-driven matrix method. Other backgrounds were estimated from the simulation and normalised to the corresponding theoretical cross section calculated at NNLO. The selection requirement on the NN discriminator and the binning of the final discriminant, the reconstructed T-quark mass M_{reco}, were optimised simultaneously to achieve the best expected

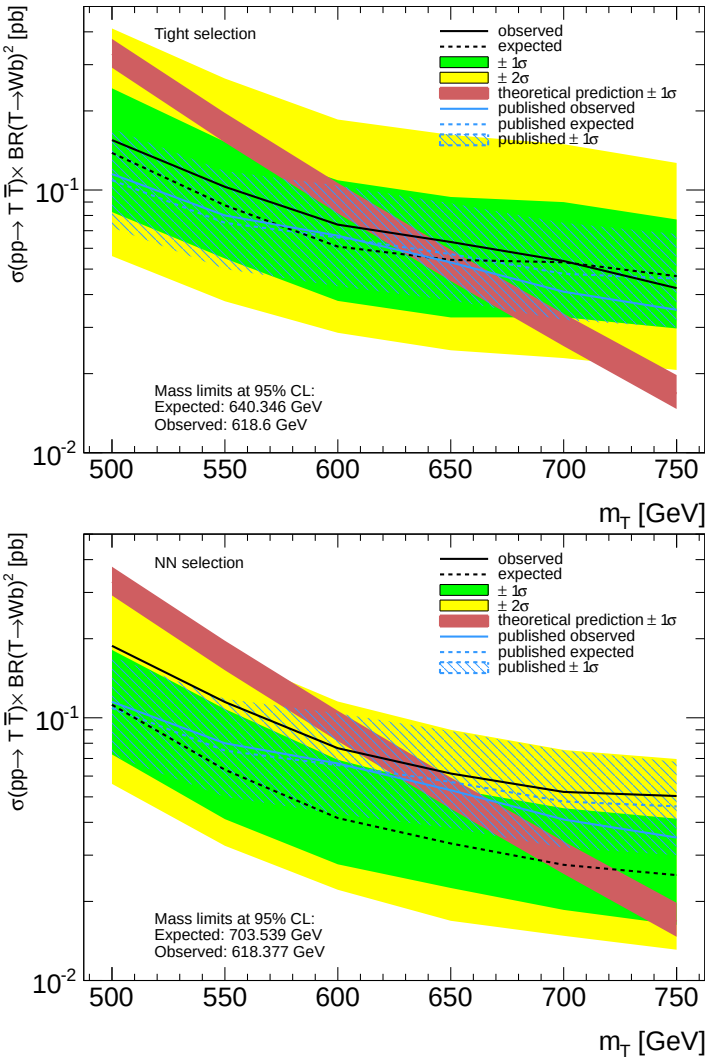


Figure 9.27: Expected (full black line) and observed (dashed black line) excluded cross section as a function of the T mass with the expected 1 (green band) and 2 σ (yellow band) systematic error bands. The theoretical prediction with theoretical 1 σ uncertainty is shown in red. The experimental result from [206] with systematic 1 σ uncertainty is shown in blue. Top: Tight selection; bottom: NN selection.

limit on the $T\bar{T}$ production cross section. The statistical interpretation has been performed using the CLs method explained in Section 6.3.1 based on the log-likelihood ratio distribution obtained from pseudo-experiments generated assuming the b and $s+b$ hypotheses. In the pseudo-experiment generation, systematic uncertainties regarding object reconstruction, flavour tagging, background normalisations and $t\bar{t}$ modelling were taken into account by allowing for Gaussian variations of the generated per-bin event yields according to the corresponding priors, as explained in Section 6.3.2. As a result, the existence of a T-quark lighter than 618.6 GeV (640.3 GeV) observed (expected) is excluded at 95% C.L using the cut-based selection, and correspondingly 618.4 GeV (703.5 GeV) observed (expected) using the NN-based selection. Using the NN techniques, an expected sensitivity improvement of ~ 60 GeV has been achieved compared to the equivalent cut-based method.

The results of this analysis are also applicable to models with a vector-like Y-quark or a singlet state vector-like T-quark decaying with a branching ratio of 100 % to a Wb pair (see Equations 2.96 and 2.96). In addition to this charged current decay, however, the vector-like T can undergo flavour changing neutral current decays according to $T \to Zt$ and $T \to Ht$ as discussed in Section 2.3.2.2. For a generic limit on the vector-like T production, these decays need to be taken into account with their appropriate branching ratios. In order to maximise the sensitivity, this requires dedicated analyses for each of the possible decays, the results of which can then be combined to obtain the limits for vector-like T production as a function of the decay branching ratios. Preliminary ATLAS results have recently become available based on 14.3 and 20.3 fb^{-1} of the $\sqrt{s} = 8$ TeV ATLAS dataset taken in 2012. Fig. 9.28 shows the expected and observed exclusion limits for the vector-like T-quark in the plane of BR($T \to Ht$) versus BR($T \to Wb$) assuming BR($T \to Wb$)+BR($T \to Ht$)+BR($T \to Ht$)=1, based on the results from dedicated searches for vector-like T in the $T \to Ht$ [217], $T \to W^+b$ [218], and $T \to Zt$ [219] decay modes, as well as the search for anomalous production of events with same-sign di-leptons and b-jets [220]. The combined result rules out the singlet (doublet) state vector-like T lighter than 670 (735) GeV.

9.9 Outlook

Various aspects of the analysis presented in this chapter could be extended or improved. Some possible improvements and extensions suggested by the author are summarised in this section.

For the estimation of the signal and $t\bar{t}$ background contribution, only those events of the corresponding simulated samples are used that correspond to e/μ+jets final states. The contributions from final states including τ leptons, as well as those corresponding to di-lepton or fully hadronic final states are not considered. Although these contribution are expected to be small, their inclusion would result in a more accurate signal and background estimation.

For the purpose of studying the sensitivity improvement potential due to a NN-based selection compared to a cut-based selection which is fully validated, the same set of variables was used as in the analysis published in [206]. However, the set of input variables could be extended to exploit further more delicate kinematic differences between the signal and the backgrounds. In particular, the variables that describe the shape of the event, such as centrality, sphericity and aplanarity, could bring additional information into the machine learning procedure. A network re-optimisation effort, including the incorporation of additional variables and investigating different network architecture, has already been undertaken in two undergraduate projects at the University of British Columbia, using the full training infrastructure provided by this analysis. No significant improvement of the discrimination power has been achieved with respect to that shown in this analysis.

The training used a single hidden layer. Recently, the so-called deep learning techniques involving a higher number of hidden layers have demonstrated an advantage over the shallow networks with few hidden layers [222], and could also be considered in future optimisations.

The discriminator of this analysis was obtained using the TMVA implementation of the MLP method [173]. Other MLP implementations, such as e.g. NeuroBayes [223], or different multivariate techniques such as boosted decision trees [224], should be investigated too.

The training used in this analysis was performed on the nominal samples and optimised for the

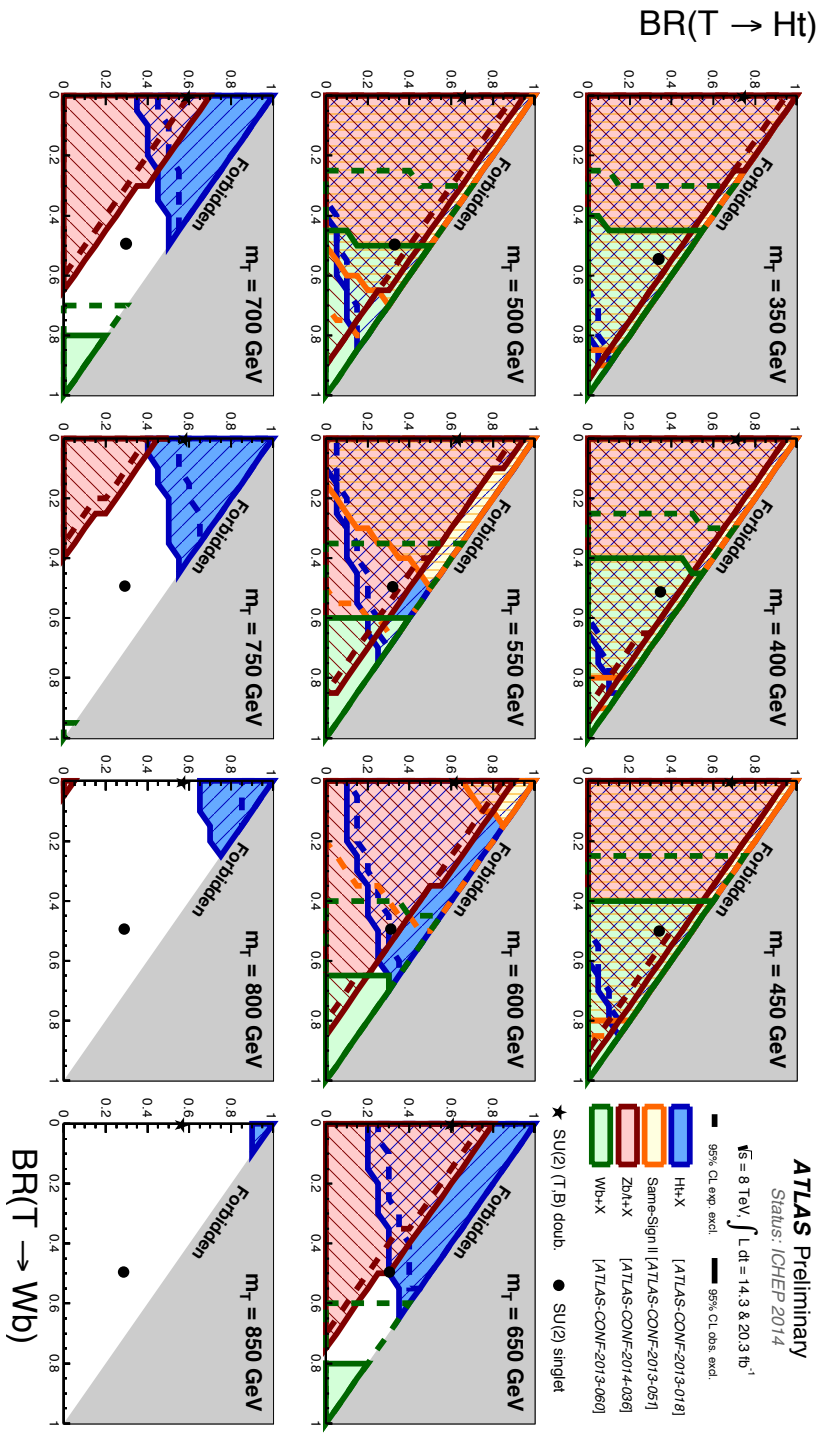


Figure 9.28: Expected and observed exclusion limits on the vector-like T quark in the plane of BR($T \to Ht$) versus BR($T \to Wb$) assuming BR($T \to Wb$)+BR($T \to Ht$)+BR($T \to Ht$)=1. Results are obtained by overlaying the results from [217–220]. Individual plots in the figure are produced using T-quark mass points between 350 and 850 GeV in steps of 50 GeV. The grey area in each plot shows the un-physical region. The contributions of individual analyses are distinguished by the colours, as indicated in the legend. The contours of the observed (expected) limits are drawn as full (dashed) lines. Plot taken from [221].

best signal to background ratio of the sample. A training incorporating the effects of systematic uncertainties, e.g. by including systematic samples generated independently from the nominal samples, could potentially provide a more realistic picture to the training procedure, and should thus be investigated.

The statistical interpretation of this analysis was limited to the signal region only and did not include any constraints on the systematic uncertainties. If the signal region was somewhat looser so that it had a sufficient statistical power to support fitting for the best likelihood by varying the systematic uncertainties, this could be done in the signal region and in a background enriched region simultaneously, which would constrain the individual sources of systematic uncertainties and thus increase the precision of the result. This analysis has furthermore been optimised only for the central value of the expected exclusion cross section regardless the systematic uncertainties on it. Including these uncertainties into the optimisation procedure could perhaps improve the precision of the limit.

This analysis does not distinguish between the T and Y states as the information of the b-jet charge is unavailable. With that information at hand, the analysis could be split into T and Y channels depending on the same-sign or opposite sign combination of Wb pairs, which would also improve the resolution of the mass reconstruction of respective quarks.

Finally, in this analysis only the pair production via the strong interaction was examined. In the currently allowed mass range of vector-like quarks, the rates of certain single production channels are expected to be similar to pair production at the expected LHC centre-of-mass energy in Run II ($\sqrt{s} = 13$ TeV), so that the searches for vector-like quarks produced in this way will become meaningful with the expected LHC data.

As vector-like quarks represent an easily accessible distinguishing feature of hypothetical models (little Higgs models, models with warped extra dimensions etc.) targeted at solving vital open questions of the SM (hierarchy problem), the interest for these exotic quarks persists. The LHC Run II data will provide an ideal opportunity to continue the quest.

Chapter 10

Conclusion

Since the discoveries of the W and Z bosons at CERN in early '80s, spontaneous symmetry breaking (SBB) of the electro-weak (EW) potential was confirmed. The following three decades of the high energy physics research were dominated by the searches for the mechanism that drives the EW SSB, in particular the direct searches for the Higgs boson, predicted by the most widely accepted SSB theory, the Brout-Englert-Higgs (BEH) mechanism. In parallel to these searches, numerous searches for signatures of alternative SSB mechanisms, as well as extensions of the well established Standard Model of particle physics (SM), were undertaken, among which many predicted the existence of new exotic quarks. After a remarkably fruitful era of physics results at Tevatron, including the completion of the SM quark sector with the discovery of the top quark, the start of the LHC operation has set the scene for the exploration of physics phenomena at the new energy frontier. The outstanding success of the CERN experiments was crowned by the discovery of the Higgs boson in the summer of 2012, by which the BEH mechanism was finally confirmed. Before the Higgs boson discovery, the global fit of the existing data was still compatible with the existence of one additional generation of chiral SM quarks, consisting of an up-type t' and a down-type b'. After the discovery, however, models with fourth generation were disfavoured, and the focus of the searches for exotic quarks was shifted towards the models which target the weak points of the SM, such as the hierarchy problem. In particular, new states that are expected to occur at energies already or soon to become accessible with the LHC, such as the vector-like quarks, have become particularly important.

Two searches for heavy quarks decaying into a W-boson and a b-quark have been presented, both performed with LHC pp collision data at $\sqrt{s} = 7$ TeV collected by the ATLAS detector in 2011, in final states with one isolated electron or muon, a significant contribution of missing transverse energy and at least three jets including at least one b-jet. The first analysis is a search for the chiral fourth generation t'-quark carried out with 1.04 fb^{-1} before the Higgs discovery, resulting in observed (expected) lower bounds on the mass of the t'-quark at 404 (394) GeV at 95% C.L. The fourth generation was then found to be hard to conciliate with the newly discovered Higgs boson at the observed rate, and the scope of the searches for heavy quarks was subsequently moved to the vector-like states. The second analysis that was presented is a search for a heavy vector-like quark assumed to decay into a W-boson and a b-quark with a branching ratio of 100%. This analysis was performed on the full 2011 data set corresponding to ~ 4.7 fb^{-1} of data, resulting in observed (expected) lower bounds on the mass of the hypothetical T-quark at 618 (704) GeV at 95% C.L.

In the search for the chiral fourth generation t'-quark, the event reconstruction made use of a sophisticated kinematic likelihood fit, KLFitter, to best match the observed objects to the original t' or $\bar{t}'$ quarks, and to fit the reconstructed t'-quark mass. An extensive study was carried out to establish the fit configuration which provides the optimal resolution of the fitted t'-quark mass. The choice of the reconstructed jets and the t'-quark mass used in the fit, as well as the treatment of b-jets, were tested. In particular, the ability of the fit to recognise which four reconstructed jets originate from the $t'\bar{t}'$ hard scattering if five leading p_{T} jets are provided to the fit, was examined. The study was performed on simulated events of the $t'\bar{t}'$ signal with the assumed t' mass of 400 GeV, and the dominant $t\bar{t}$ background. The performance was quantified in terms of the reconstruction efficiency as a function of the number of objects reconstructed in the detector that were correctly identified in the fit. It was found that the best fit performance is achieved when the four hardest jets of the event are used in the fit. This study provided a powerful final discriminant, the reconstructed heavy quark mass. Statistical interpretation of the analysis was performed on the full kinematic distribution of the reconstructed heavy quark mass, which allowed some of the systematic uncertainties to be constrained. In addition, the correlations between the likelihood function and the relevant observables were studied to identify possible improvements to the algorithm. It has been found that, in order to improve the resolution of the t'-quark mass fitted by considering more than four hardest jets in the event, the fit needs additional kinematic constraints, such as the space angles between the decay products of t'-quarks and W-bosons, in order to reach a better performance with five jets considered.

The search for the heavy vector-like T-quark is based on a simple event reconstruction, identifying all participants of the tree level hard process by making use of the b-tagging probabilities, W-boson mass constraint (to identify the W candidates and calculate the longitudinal momentum of the neutrino) and minimisation of the difference between the reconstructed masses of the T- and $\bar{T}$-quarks. A set of variables was derived from this event reconstruction, mainly inspired by the differences in the boost of the reconstructed W-bosons and T-quarks between the signal and the dominant background processes. Based on these variables, two approaches were developed: a cut-based and a multi-variate analysis. A NN discriminator was optimised by studying the effects of various network architectures, machine learning algorithms, training functions, and a number of other NN parameters. The statistical interpretation was implemented, reproducing the existing cut-based result and producing final limits with the NN-based selection. The requirement on the NN discriminator and the binning of the final discriminant, the reconstructed T-quark mass, were optimised simultaneously, to achieve the best expected exclusion limit. As a result, an overall improvement of ~ 60 GeV on the expected lower bound of the T-quark mass was achieved, compared to the cut-based result. An upper limit on the T-quark mass at 95% C.L. was placed at 618.4 (703.5) GeV observed (expected) using the cut-based final selection, and correspondingly at 618.4 (703.5) GeV observed (expected) using the final selection bases on the NN discriminator. This result fully applies to the vector-like $Y_{-4/3}$-quark. Furthermore, this analysis can be used to set the lower limits on the vector-like $T_{2/3}$-quark in the singlet state, as a function of the relative branching fractions of the possible decay modes ($T \to Wb/Zt/Ht$), providing an optimal performance for the phase space with a high branching fraction of the $T \to Wb$ decay. At the time of completing this thesis, preliminary exclusion limits had been strengthened by analyses performed with 14.3 and 20.3 fb^{-1} of LHC pp collision data at $\sqrt{s} = 8$ TeV, placing the lower bound on the singlet state vector-like T at 670 GeV through the $T \to Wb$ decay, and at the singlet and doublet state at 790 (735) GeV through the $T \to Ht$ ($T \to Zt$) decay.

Appendix A

SCT Byte Stream Errors Study

A.1 Introduction

The basic unit of the SCT readout is a link which reads out the information form one side of an SCT module. The SCT readout is a complex procedure in which errors can occur at various levels. SCT ByteStream (BS) errors consist of different types of data acquisition (DAQ) errors related to timing, synchronization, buffering, link masking and similar issues. Hits recorded by the side of a module that reports certain error types discussed below are considered invalid and are not used in the track reconstruction.

There is thus a certain data loss due to the BS errors. This loss is not taken into account in the Monte Carlo (MC) simulation. The main objective of this study was to quantify the data loss due to the BS errors in order to decide whether or not it needs to be modeled in MC. In addition, this study also aimed to improve the understanding of the BS error behavior. To that end, the geometrical distribution of BS errors as well as their correlation to some relevant quantities have been studied. The quantities used in the study are: instantaneous luminosity; hit occupancy, defined as a ratio of the number of hits in a module to the number of channels in that module; and hits on tracks occupancy, defined as a ratio of the number of hits in a module used for tracking to the number of channels in that module.

The study is based on the information from the SCT monitoring histograms obtained from the Express data stream for the pp collision data at $\sqrt{s} = 7$ TeV collected by the ATLAS detector in 2010 (for the explanation of the ATLAS data streams see Section 3.3.1). Twelve error types accessible in this way and analysed in this study are briefly explained in Section A.2. The SCT monitoring histograms contain only the per run summary information, that is the error rates for each error type in each module side. Information concerning individual events is not available and it is not known whether several errors of different types occurred in the same event, in a given module side. Therefore, it is neither possible to determine the exact number of events with errors, nor to study correlations of the error types. However, it is possible to determine the lower and the upper limit on the fraction of data containing errors, using the approximations described in the following sections. The numbers obtained are also compared to the results from the monitoring histograms of the MinBias data stream and the difference is found to be negligible.

A.2 ByteStream Error Types

Twelve error types stored in the SCT monitoring histograms are briefly described below.

- **TimeOut Error:** Occurs when the module does not return any event before the readout time ends, although a trigger was sent to it. This can happen if the module is dead or un-configurred.
- **BCID error:** Occurs when the bunch crossing ID of the event does not match to that received from the Timing, Trigger and Control (TTC) signal. This implies a synchronization problem and usually occurs when the wrong event is read out.
- **LVL1ID error:** Occurs when the Level 1 ID of the event does not match to that received from the TTC. This implies a synchronization problem and usually occurs when the wrong event is read out.
- **Preamble error:** The event is characterised by a wrong string.
- **Formatter error:** The formatter contains an illegal link.
- **MaskedLink error:** The link was masked off automatically either by the hot-link-masking mechanism or by the stopless removal mechanism. The latter happens when an entire ROD goes BUSY.
- **ABCD Error:** One of the six ABCD chips on a hybrid side has one of the following errors:
  - No data available. Chip has not received the L1 trigger.
  - Buffer overflow.
  - Buffer error. The 128 strips corresponding to the chip with the error will probably not have any hits, but the rest of the module should not be affected. Can be cured by soft reset.
- **Raw Data Error:** Data is received in a wrong format. Can be caused by optical noise. There might be good hits in the link before this error occurs.
- **Rod Clock Error:** Occurs when either the Back of Crate (BOC) clock error bit or the TTC Interface Module (TIM) clock error bit is set. Then BOC does not detect an acceptable signal from the TIM and has to generate its own clock in order to keep the module clocked and thereby avoid thermal issues. This error normally leads to synchronisation issues.
- **Truncated ROD Error:** When the ROD fragment exceeds the specified maximum size it gets truncated.
- **ROB Fragment Error:** An OR of all ROB header errors plus some additional conditions.
- **ByteStream Parse error:** Offline ByteStream decoder reads an invalid word, which is either not a header, trailer or hit, or a non-existing strip number.

Different error types affect data taking in different ways. For example, the so-called *TimeOut* error and *MaskedLink* error affect the data taking link from one side of a module, and flag the data taken by this module side as corrupted until the module is recovered or reconfigured. On the other hand, the so-called *BCId* error and *LVL1Id* error concern individual events making the data recorded by an affected module side invalid for those particular events only. Unless the module is un-configured or has a physical problem, *BCId* error and *LVL1Id* error are cleared by the automatic module resynchronization during the run. The *ABCD* error affects only individual chips of a module side and error flags are set for the strips corresponding to the affected chips only.

A.3 Estimation of the total fraction of data with BS errors

The total fraction of data with the SCT BS errors is given by:

$$\epsilon_{err} = \frac{\sum_{r,ms} \mathcal{L}_r \frac{\epsilon_{ms,r}}{N_{ms}}}{\sum_r \mathcal{L}_r}, \tag{A.1}$$

where $\mathcal{L}_r$ is the integrated luminosity of the run r, $\epsilon_{ms,r}$ the error rate of the module side ms in the run r and N_{ms} the overall number of the SCT module sides (2×4088). The exact quantity $\epsilon_{ms,r}$ cannot be extracted from the information available in the monitoring histograms. It can be estimated under the assumption of either a maximal or a minimal error type correlation. If maximal correlation is assumed, errors of different type occur as often as possible at the same time, i.e. in the same events. Then, the fraction of events $\epsilon^{max}_{ms,r}$ with any type of error is simply given by the largest among the error rates $\epsilon^{t}_{ms,r}$ of all types t:

$$\epsilon^{max}_{ms,r} = \max_t \epsilon^t_{ms,r}.$$

The real value $\epsilon_{ms,r}$ cannot be smaller than this, so this assumption delivers a lower limit for $\epsilon_{ms,r}$.

The opposite extreme case, when the error types are not correlated at all, i.e. errors of different types occur in different events, the overall error rate $\epsilon^{sum}_{ms,r}$ of a module side for the observed run is given by the sum over all error types t:

$$\epsilon^{sum}_{ms,r} = \sum_t \epsilon^t_{ms,r}. \tag{A.2}$$

It should be noted that this value can exceed 1. Thus, we should bare in mind that this value is an approximation and not a real error rate. However, the true value of $\epsilon_{ms,r}$ cannot be larger than this, so $\epsilon^{sum}_{ms,r}$ provides an upper limit for $\epsilon_{ms,r}$.

Considering the quantities defined, one can now define the lower and the upper limit, ϵ^{low}_{err} and ϵ^{up}_{err} respectively, of the fraction of data with errors:

$$\epsilon^{low}_{err} = \frac{\sum_{r,ms} \mathcal{L}_r \frac{\epsilon^{max}_{ms,r}}{N_{ms}}}{\sum_r \mathcal{L}_r}, \tag{A.3}$$

$$\epsilon^{up}_{err} = \frac{\sum_{r,ms} \mathcal{L}_r \frac{\epsilon^{sum}_{ms,r}}{N_{ms}}}{\sum_r \mathcal{L}_r}. \tag{A.4}$$

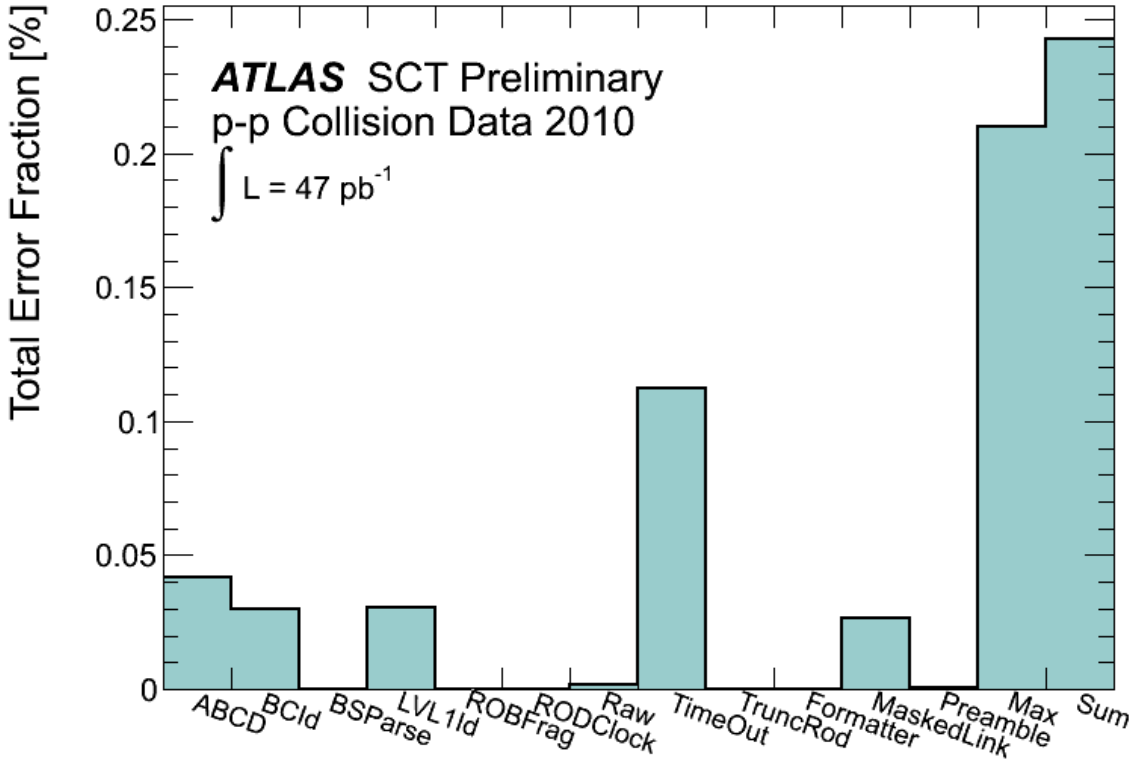


Figure A.1: Fraction of data with BS errors of each type, ϵ^{t}_{err}, as well as the lower and upper limit on the total fraction of data containing any type of BS errors, ϵ^{low}_{err} (corresponding to the bin labeled as "Max") and ϵ^{up}_{err} (corresponding to the bin labeled as "Sum").

A further relevant quantity is the overall fraction of data with errors of a particular type t:

$$\epsilon^{t}_{err} = \frac{\sum_{r,ms} \mathcal{L}_r \frac{\epsilon^{t}_{ms,r}}{N_{ms}}}{\sum_r \mathcal{L}_r}. \tag{A.5}$$

The plot in Figure A.1 shows the values of ϵ^{t}_{err}, where t stands for all error types listed in Section A.2, as well as ϵ^{low}_{err} and ϵ^{up}_{err}, corresponding to the bins labeled as Max and Sum in the plot respectively. All numbers shown are extracted from the SCT monitoring histograms are shown in the plot in the figure A.1.

The last two bins of the plot correspond to ϵ^{low}_{err} and ϵ^{up}_{err} respectively, implying that $0.21\% \leqslant \epsilon_{err} \leqslant 0.25\%$. It turns out that the dominant contribution comes from the *TimeOut* error type. Most of the *TimeOut* errors are a consequence of dis-functional SCT transmitting units (TXs). When a TX stops operating, the module corresponding to it stops responding and thus automatically reports a *TimeOut* errors until the TX is disabled or replaced.

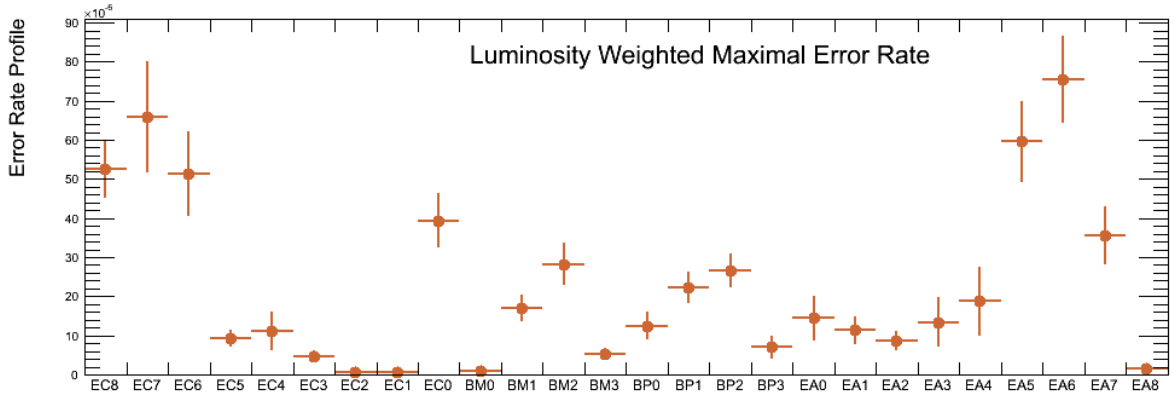


Figure A.2: Fraction of data with errors as a function of the SCT layers. $EC0 - EC08$ stand for the endcap C discs, $EA0 - EA8$ stand for the endcap A discs, $BM0 - BM3$ stand for the negative side barrel half layers and the $BP0 - BP3$ for the positive side barrel half laeyers.

A.4 Geometrical distribution of BS errors

BS errors are caused by various module problems. An immediate indication of the presence of un-configured or masked off modules, or TX deaths, is an increase of the BS error rates. Some of such problems can be solved during the run by recovering the problematic modules manually or in an automatic recovery procedure.

The plot in figure A.2 shows the break-down of the second to last bin (given the name "Max") of the plot in the figure A.1 into the SCT barrel layers and endcap discs. The barrel layers are divided into negative eta ($BM0 - BM3$) and positive eta ($BP0 - BP3$) half layers. The endcap discs are labeled as $EC0 - EC8$ for the endcap C, and correspondingly $EA0 - EA8$ for the endcap A.

The plot shows that overall the positive eta side of the detector has collected slightly more data reporting BS errors. Furthermore it turns out that the fraction of data containing errors is rising in first three barrel layers and then dropping in the last layer. Except for the $EA8$ disc, which has a surprisingly low fraction, and the $EC0$ disc, which has exceptionally high fraction of data with errors, the outer endcap discs have in general more errors than the inner discs. This behavior is a combined effect of the module quality, calibration, position of the modules with respect to the beam etc. For example, it is known that the modules with a higher quality are installed in the inner layers while the modules with a lower quality are used for the outer layers of the SCT.

A.5 Dependence of the BS Error rates on the luminosity

During the 2010 data taking period, the beam parameters changed several times. The center of mass energy and the number of bunches per beam were increased, and the number of protons per bunch rose in several steps, resulting in a dramatic increase of the instantaneous luminosity (see Section 3.1.1 for details). The correlation of the BS error rates with the instantaneous luminosity cannot be studied directly from the monitoring histograms. Nevertheless, one can

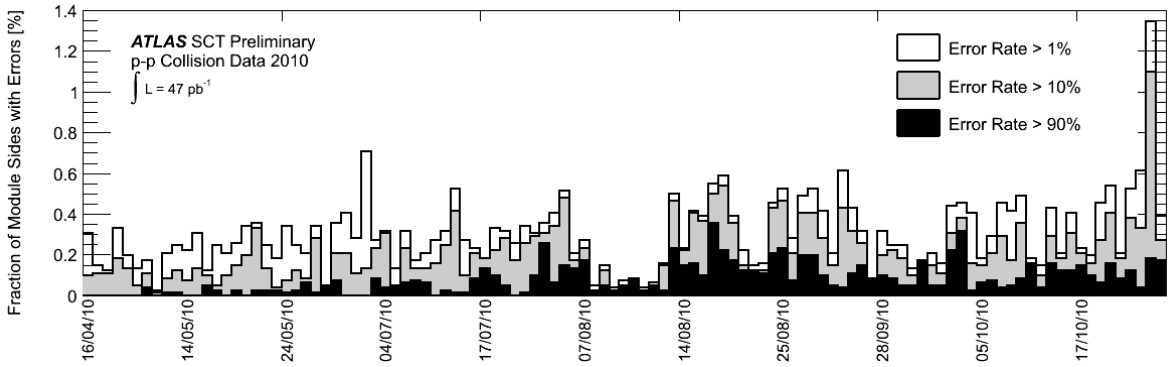


Figure A.3: Number of the SCT module sides reporting any kind of BS errors with the rate over $x\%$, x being 1, 10 and 90. The numbers are reported for the data taking runs from April throughout October 2010.

gain an indirect insight into this correlation by studying the behavior of the errors throughout the runs, bearing in mind that the luminosity was significantly increasing with time.

Figure A.3 shows the number of module sides reporting any kind of BS errors with the rates of least 1%, 10% and 90%. Appart form the run 167844, the second last run of the campaign, which reports exceptionally large number of errors, there is no obvious increase of the number of module sides with errors through the data taking. According to the plot, the run 167844 has significantly more module sides with low error rates than the other runs, but comparably many module sides with high error rates than the other runs.

Furthermore, it is relevant to know whether the error rates increase with the amount of data collected during the run, i.e. with the integrated luminosity of the run. Figure A.4 shows the correlation of the maximal module side error rate summed over all SCT module sides in an observed run to the integrated luminosity of that run. The plot shows that the runs with high integrated luminosity do not have higher average error rates. The point with the highest error rate in the plot corresponds to the run 167844 discussed above.

A.6 Dependence of the BS Error rates on the hit occupancy and the hits on tracks occupancy

Since it is known that the hit occupancy of a wafer is proportional to the instantaneous luminosity, by studying the correlation of the BS Error rates and the hit occupancy one can indirectly learn how the BS error rates behave with respect to the instantaneous luminosity. Studying the correlation of the error rates to the hits on tracks occupancy, however, shows the effect of the space points and track reconstruction in presence of the BS errors.

Figure A.5 shows the average hit occupancy in every barrel half layer and endcap disc. The average hit occupancy of a module side drops with the radial distance from the beam pipe, as expected. The negative and the positive pseudo-rapidity sides of the barrel are furthermore quite symmetric in terms of the average hit occupancy. The geometrical behavior of the BS Error rates from the figure A.2 shows in turn no strong correlation to this pattern.

In fact, it is not possible to make a general statement about the overall error rate dependence

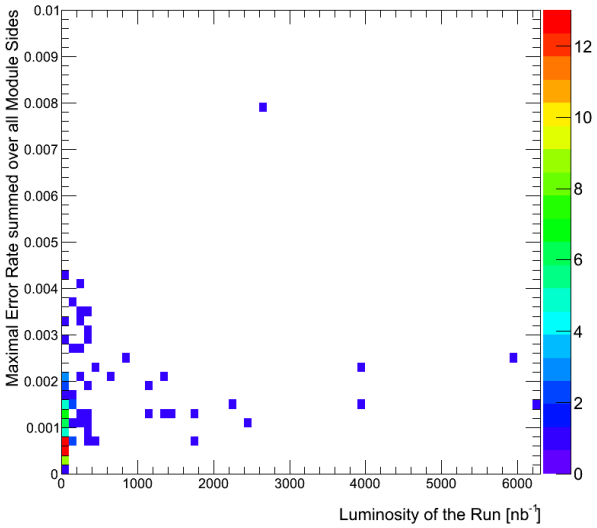


Figure A.4: Maximal module side error rate averaged over total SCT in the observed run

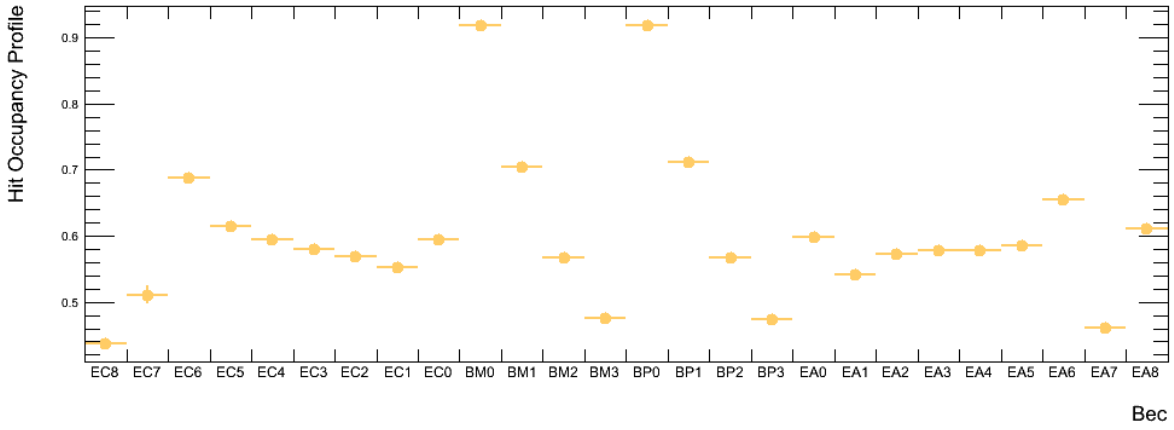


Figure A.5: Average Hit Occupancy of a module side versus SCT layer. The x-axis labels follow the same scheme as in the figure A.2.

on the hit occupancy because different error types depend on the hit occupancy in different manners. In the following, several error types are discussed separately. The overall effect is investigated in terms of $\epsilon^{max}_{ms,r}$ and $\epsilon^{sum}_{ms,r}$.

TimeOut Error Figures A.6 (a) and (b) show the two dimensional histograms of the *TimeOut* error rate for one module side in one run in bins of the hit occupancy, and the hits on tracks occupancy of that module side in the observed run, respectively. The error rate seems to be anti-correlated to both sorts of occupancy. If a module reports a *TimeOut* error in an event, it has not recorded any hit in that event. The more *TimeOut* errors occur for a module during a run, the fewer hits are recorded in that run. The number of hits used for tracks is further reduced by the presence of the other BS errors or simply by the choice of the tracking algorithms, which makes the hits on tracks occupancy concentrated around lower values compared to the

hit occupancy. The *MaskedLink* error type behaves in a similar manner as the *TimeOut* error type with respect to the occupancies.

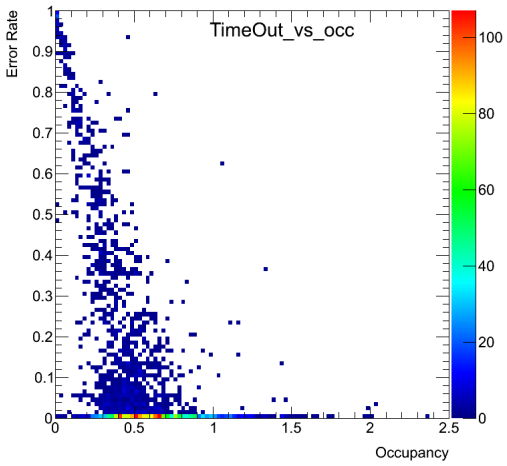
BCId Error *BCId* error is an indication of a synchronization problem. In the presence of *BCId* errors the module side still records the hits, but these hits are not used for track reconstruction. The occupancy distribution in Figure A.6 (c) shows no dependence on the hit occupancy. On the other hand, Figure A.6 (d) shows that the *BCId* error rate is somewhat anti-correlated to the hits on tracks occupancy. This can be explained by the fact that the more events report the *BCId* error for a given module, the fewer hits recorded by it are usable for the track reconstruction.

ABCD Error This error type is slightly correlated to the occupancy. Since in case of this error only individual chips do not work properly, hits are still recorded by the rest of the module. It looks like the *ABCD* error rate is more likely to happen if there are more hits in the module.

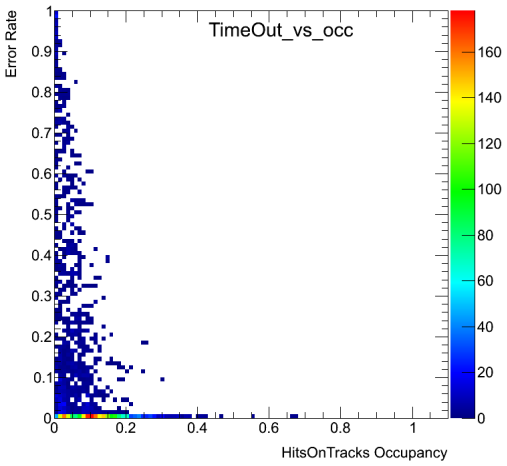
Max and Sum Errors The overall dependence of the error rates on the occupancy can be expressed in terms of the scatter plot of the $\epsilon^{max}_{ms,r}$ and $\epsilon^{sum}_{ms,r}$ as a function of the occupancy. Figure A.7 shows that $\epsilon^{max}_{ms,r}$ and $\epsilon^{sum}_{ms,r}$ are slightly anti-correlated to the occupancies, since they are dominated by the *TimeOut* errors, as it follows from the figure A.1.

A.7 Conclusion

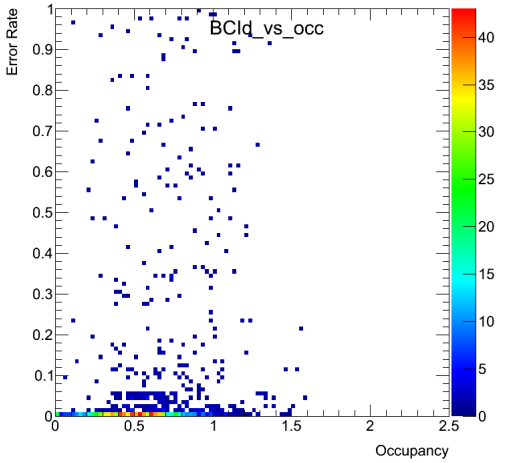
Studies of the SCT monitoring histograms from the Express stream show that the total fraction of 7 TeV *pp* collision data taken by the ATLAS detector in 2010 reporting any kind of the SCT BS errors is at least 0.21% but does not exceed 0.25%. This fraction is sufficiently small to be neglected in the MC simulation. It is furthermore found that the BS error rates do not grow significantly with the increasing instantaneous or integrated luminosity. The dependence of the BS errors on the hit occupancy and hits on tracks occupancies has been studied and it has been found that different error types depend on occupancies in different manners.



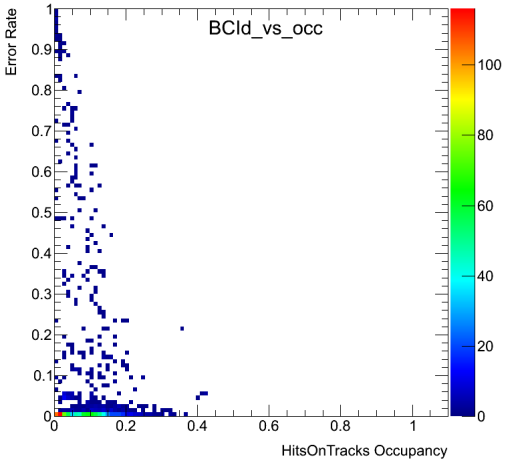
(a)TimeOut error rate versus Hit Occupancy



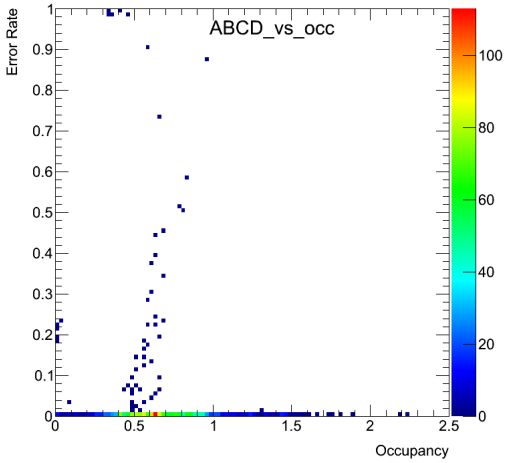
(b)TimeOut error rate versus Hits on Tracks Occupancy



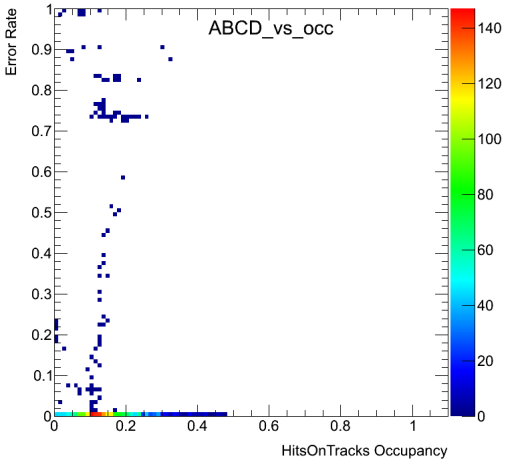
(c)BCId error rate versus Hit Occupancy



(d)BCId error rate versus Hits on Tracks Occupancy

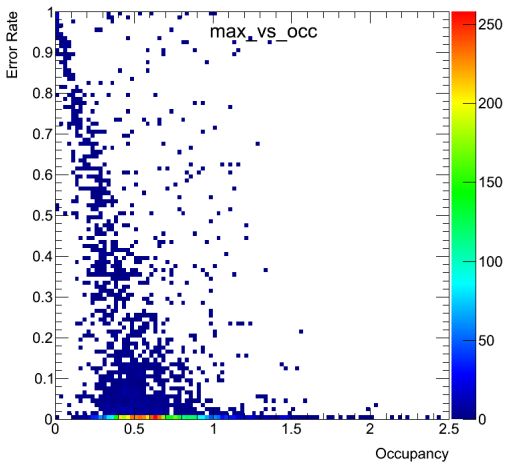


(e)ABCD error rate versus Hit Occupancy

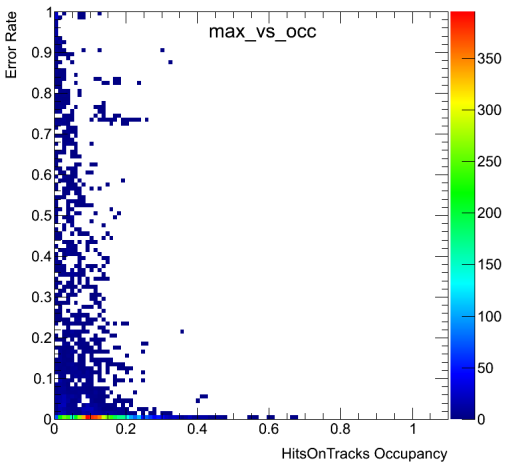


(f)ABCD error rate versus Hits on Tracks Occupancy

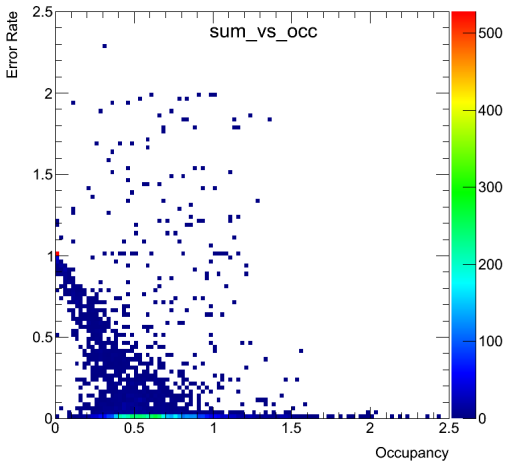
Figure A.6: Error Rates versus Occupancies



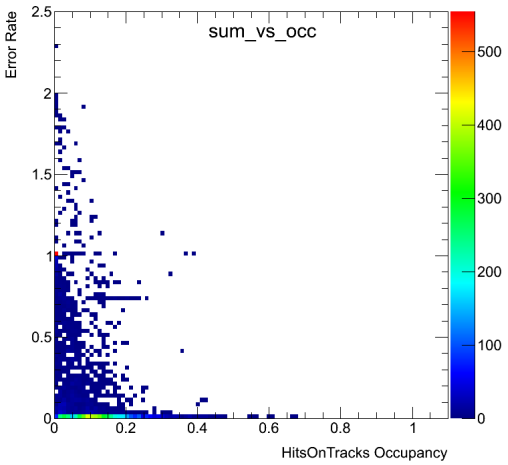
(a)Max error rate versus Hit Occupancy



(b)Max error rate versus Hits on Tracks Occupancy



(c)Sum error rate versus Hit Occupancy



(d)Sum error rate versus Hits on Tracks Occupancy

Figure A.7: Error Rates versus Occupancies

Appendix B

KLFitter Studies - Matching and Identification Efficiency

This chapter contains supplementary plots to Chapter 8

B.1 Nominal t' sample in the μ+jets channel

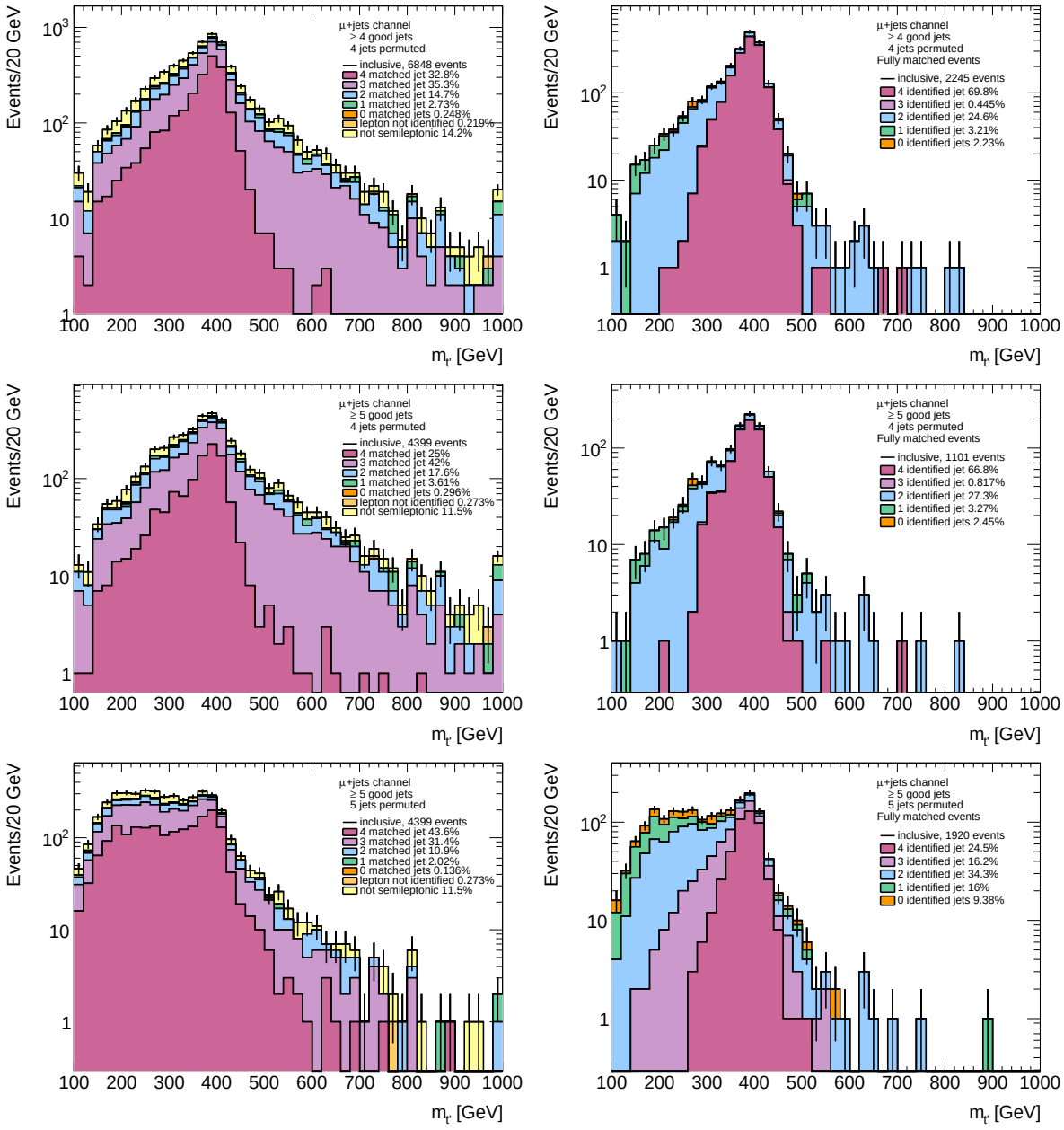


Figure B.1: Distribution of the fitted t'-quark mass decomposed according to: the number truth matches (left) and the number of proper identifications in events with four truth matches (right) shown in stacked plots. The distributions are shown for the samples with $\geqslant 4$ (top) and $\geqslant 5$ (middle and bottom) reconstructed jets in the event in the μ+jets channel. The distributions have been produced by fitting four (top and middle) or five (bottom) leading jets in the nominal t' sample with $m_{t'} = 400$ GeV.

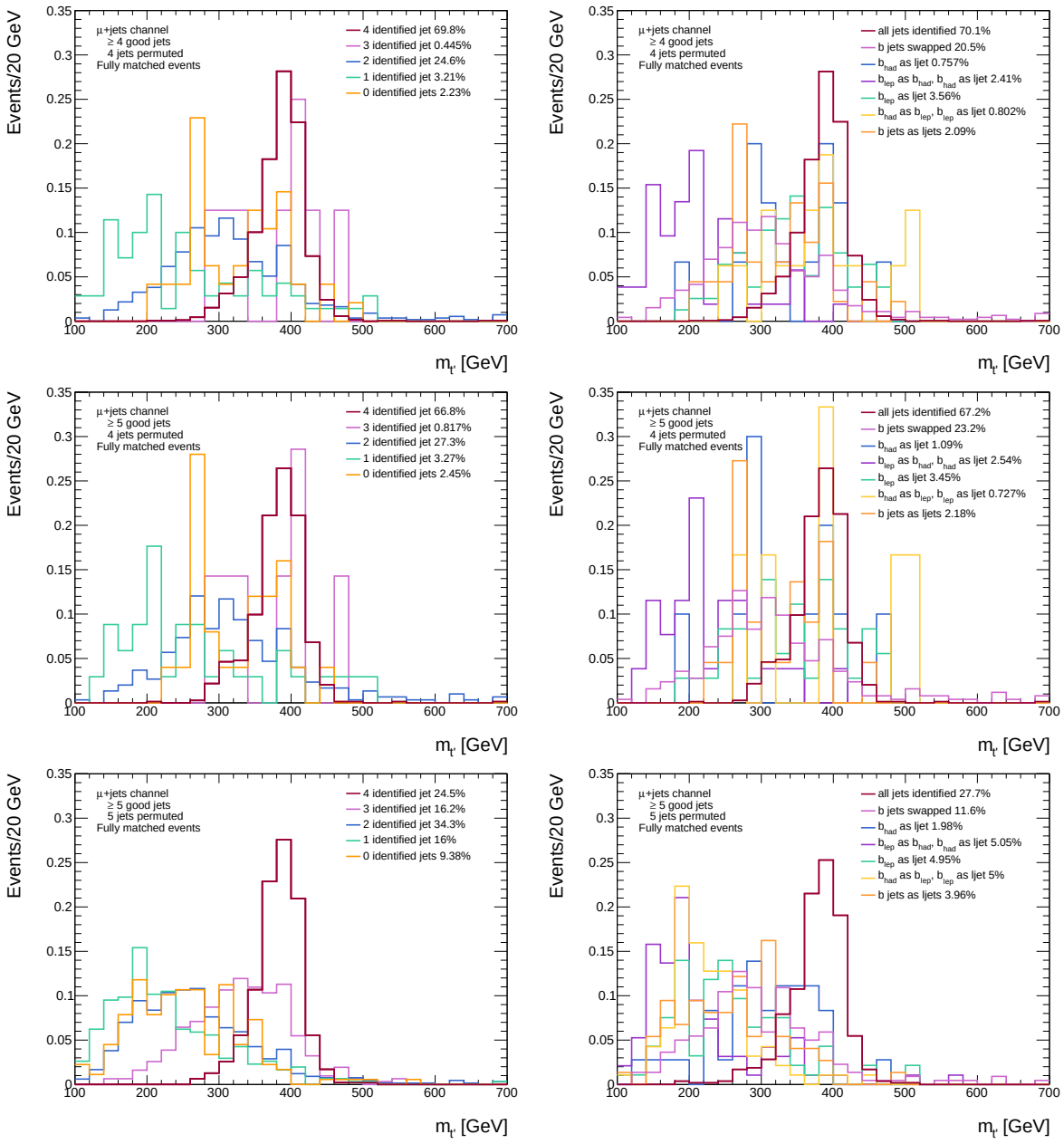


Figure B.2: Shapes of the various components of the fitted t' mass distribution of the fully matched events decomposed according to: the number of identified jets (left) and nature of the mis-identification (right). All histograms are normalised to unity. The distributions are shown for the samples with $\geqslant 4$ (top) and $\geqslant 5$ (middle and bottom) reconstructed jets in the event in the mu+jets channel. The distributions have been produced by fitting four (top and middle) or five (bottom) leading jets in the nominal t' sample with $m_{t'} = 400$ GeV.

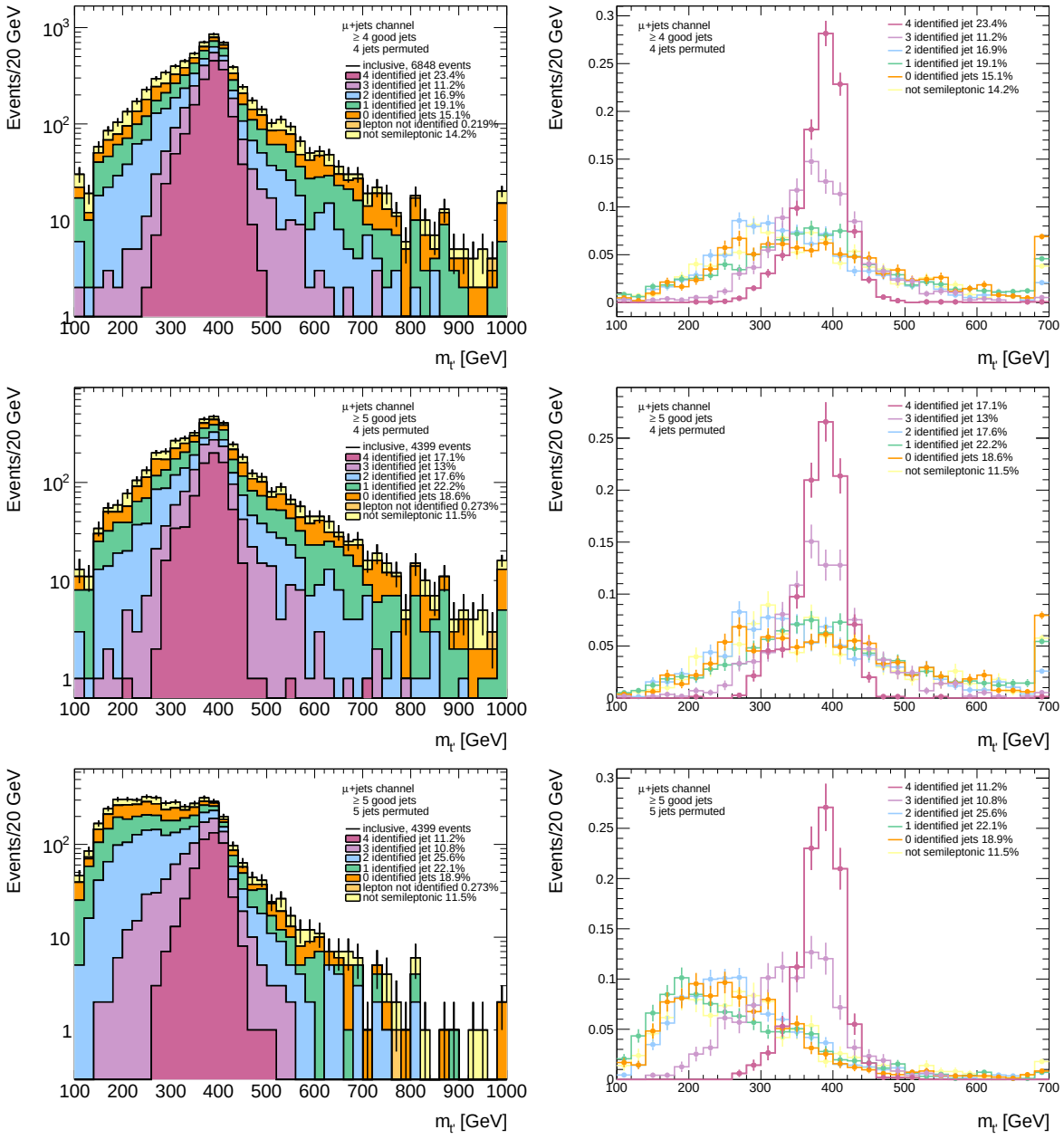


Figure B.3: Distribution of the fitted t'-quark mass decomposed according to the number of identified reconstructed jets shown in a stacked plot (left) and as individual histograms normalised to unity. The distributions are shown for the samples with $\geqslant 4$ (top) and $\geqslant 5$ (middle and bottom) reconstructed jets in the event in the mu+jets channel. The distributions have been produced by fitting four (top and middle) or five (bottom) leading jets in the nominal t' sample with $m_{t'} = 400$ GeV.

B.2 Plots for the $t\bar{t}$ sample

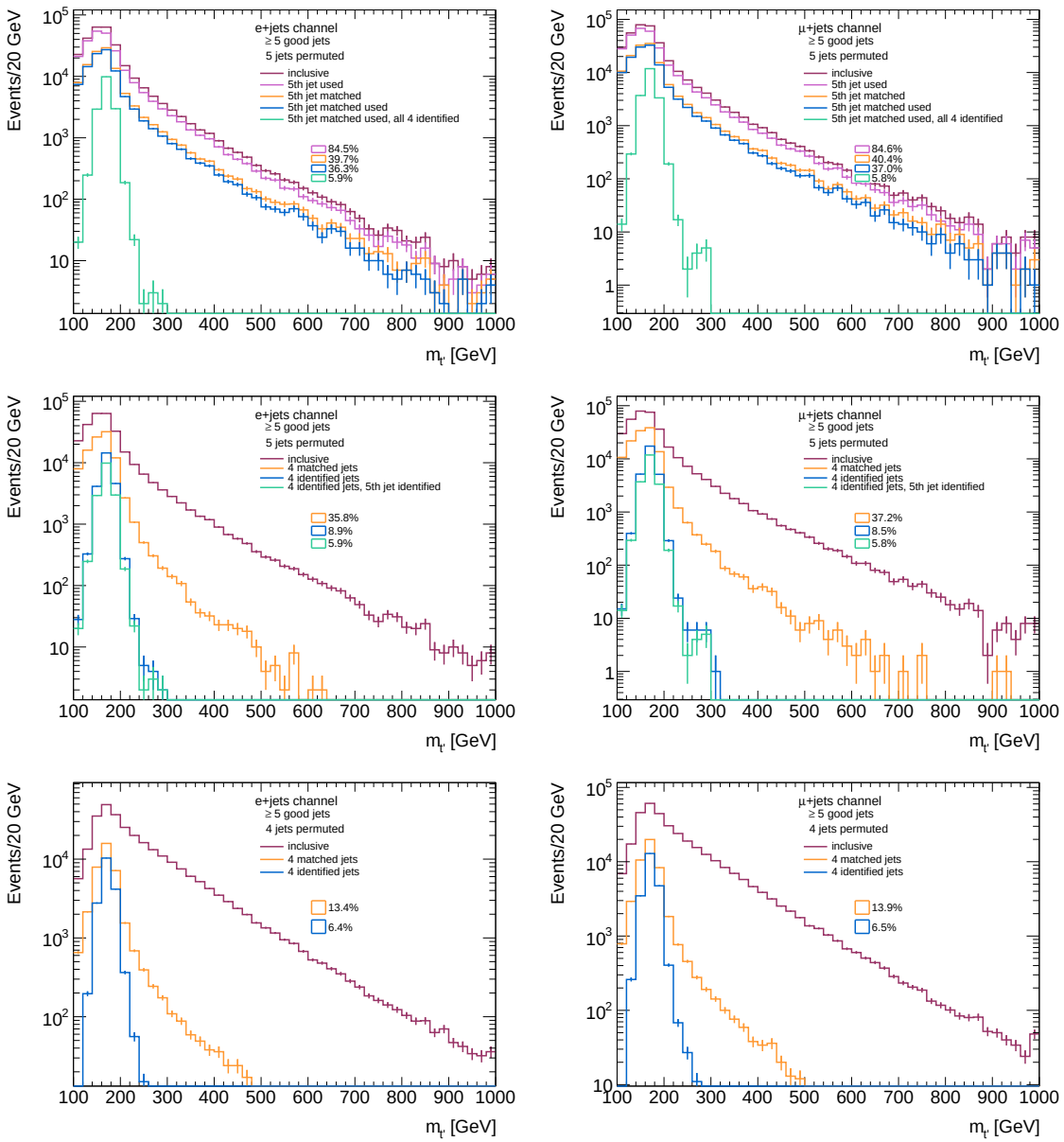


Figure B.4: Decomposition of the fitted top quark mass distributions in the e+jets (left) and μ+jets (right) channels for the events with $\geqslant 5$ reconstructed jets obtained by permuting five (top and middle) or four (bottom) leading jets. The plots in the firs row show the fractions of events in which the fifth leading jet is: used in the mass reconstruction (pink line), matched to a generator level final state jet (orange line), matched and used in the mass reconstruction (blue line), matched and used in events with all four jets properly identified. The second and third row plots show the fractions of events in which: four reconstructed jets are matched to generator level final state jets (orange line) or properly identified (blue line). In addition, the plots in the middle row show the fraction of events with four identified jets among which is also the fifth hardest reconstructed jet.

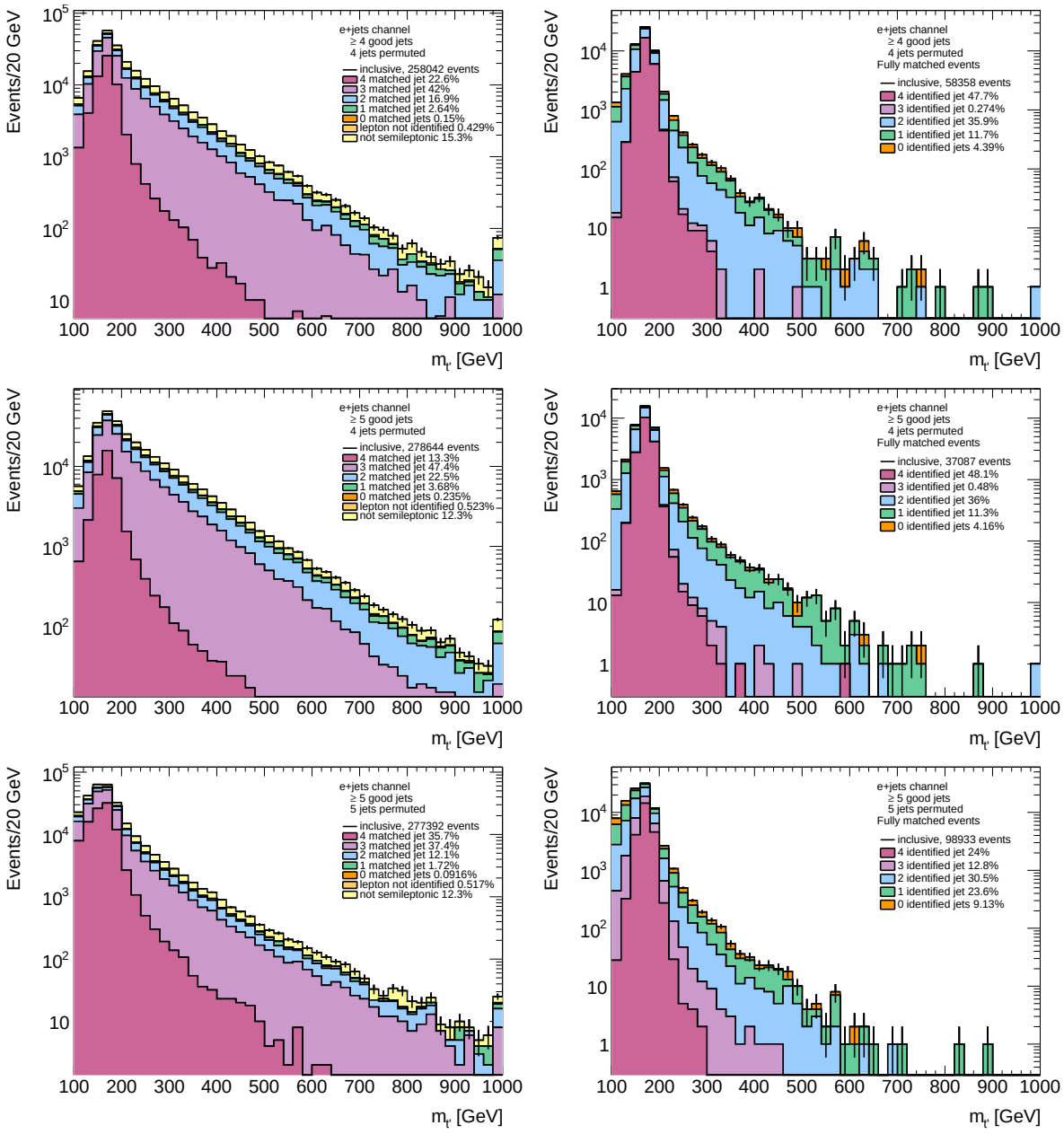


Figure B.5: Distribution of the fitted top quark mass decomposed according to: the number truth matches (left) and the number of proper identifications in events with four truth matches (right) shown in stacked plots. The distributions are shown for the samples with $\geqslant 4$ (top) and $\geqslant 5$ (middle and bottom) reconstructed jets in the event in the e+jets channel. The distributions have been produced by fitting four (top and middle) or five (bottom) leading jets in the nominal $t\bar{t}$ sample.

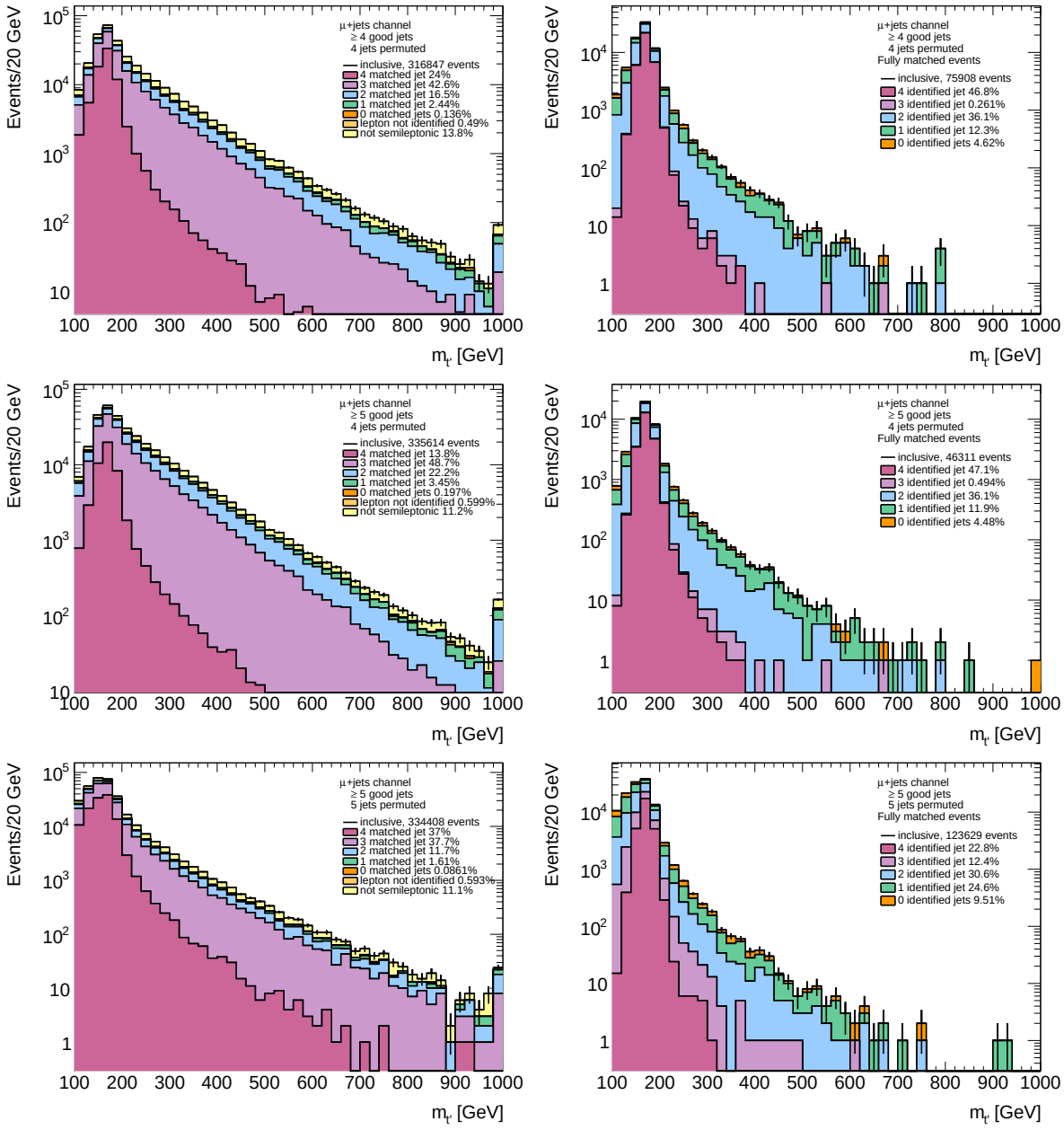


Figure B.6: Distribution of the fitted top quark mass decomposed according to: the number truth matches (left) and the number of proper identifications in events with four truth matches (right) shown in stacked plots. The distributions are shown for the samples with $\geqslant 4$ (top) and $\geqslant 5$ (middle and bottom) reconstructed jets in the event in the μ+jets channel. The distributions have been produced by fitting four (top and middle) or five (bottom) leading jets in the nominal $t\bar{t}$ sample.

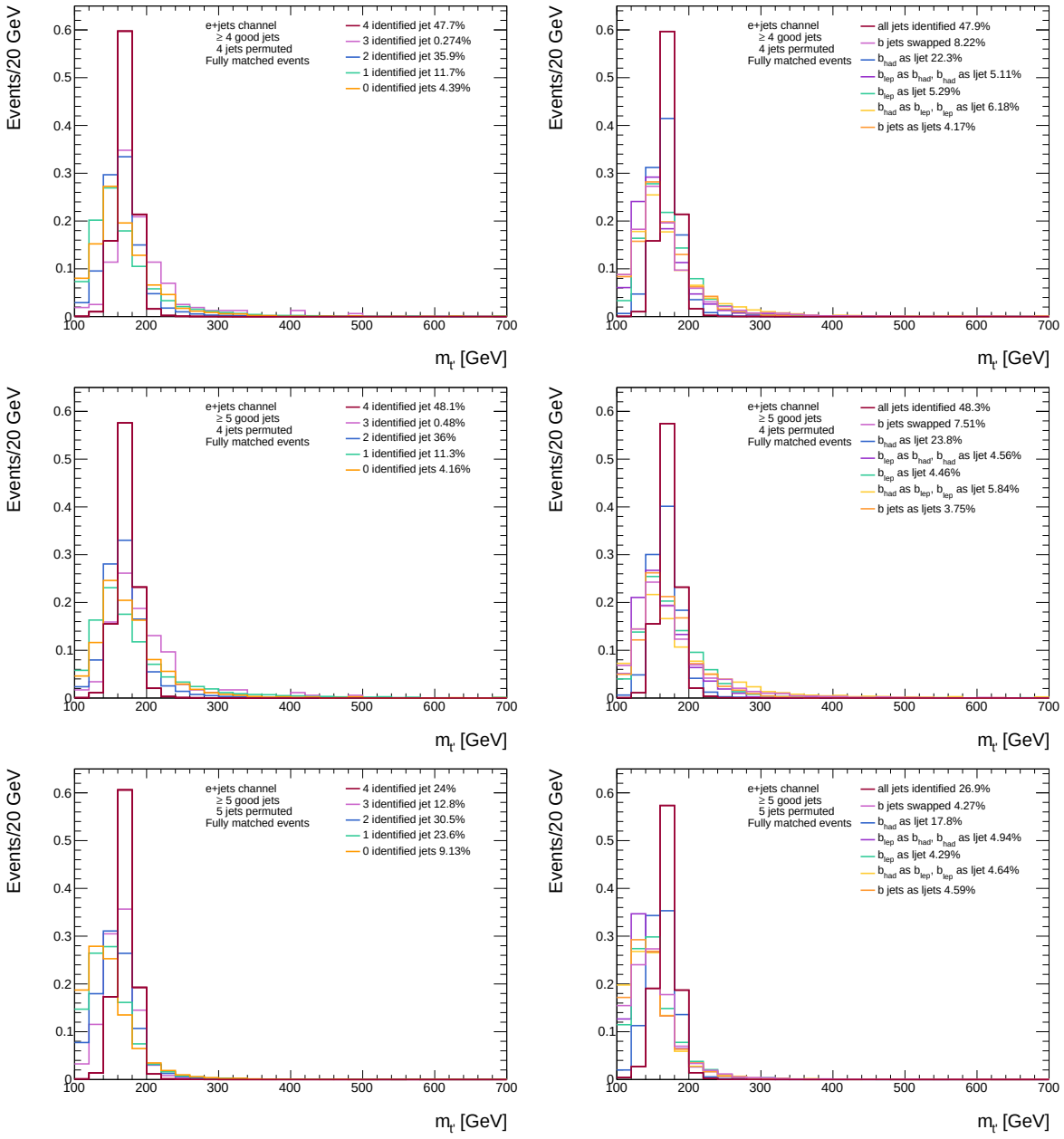


Figure B.7: Shapes of the various components of the fitted top quark mass distribution of the fully matched events decomposed according to: the number of identified jets (left) and nature of the mis-identification(right). All histograms are normalised to unity. The distributions are shown for the samples with $\geqslant 4$ (top) and $\geqslant 5$ (middle and bottom) reconstructed jets in the event in the e+jets channel. The distributions have been produced by fitting four (top and middle) or five (bottom) leading jets in the nominal $t\bar{t}$ sample.

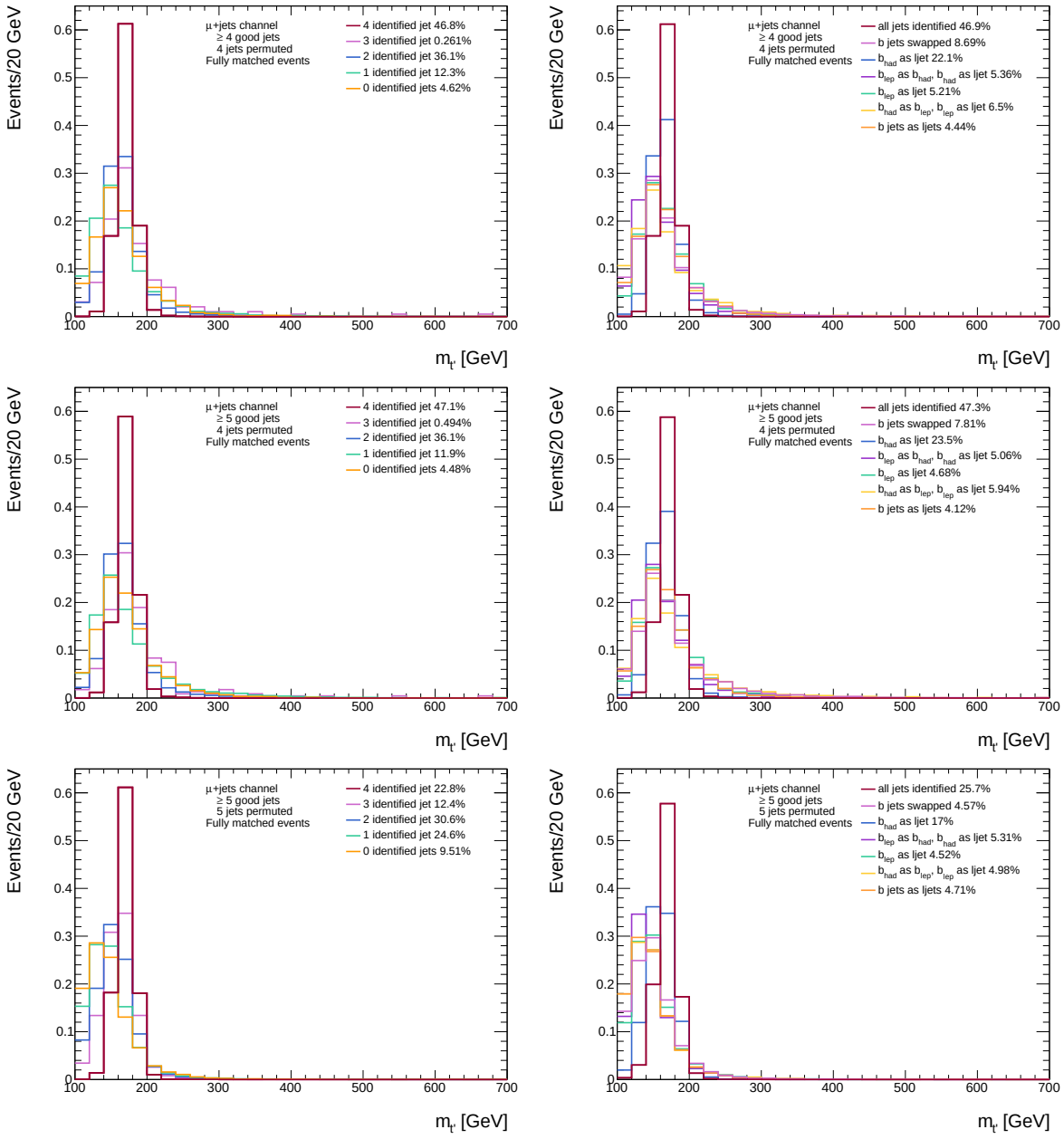


Figure B.8: Shapes of the various components of the fitted top quark mass distribution of the fully matched events decomposed according to: the number of identified jets (left) and nature of the mis-identification(right). All histograms are normalised to unity. The distributions are shown for the samples with $\geqslant 4$ (top) and $\geqslant 5$ (middle and bottom) reconstructed jets in the event in the μ+jets channel. The distributions have been produced by fitting four (top and middle) or five (bottom) leading jets in the nominal $t\bar{t}$ sample.

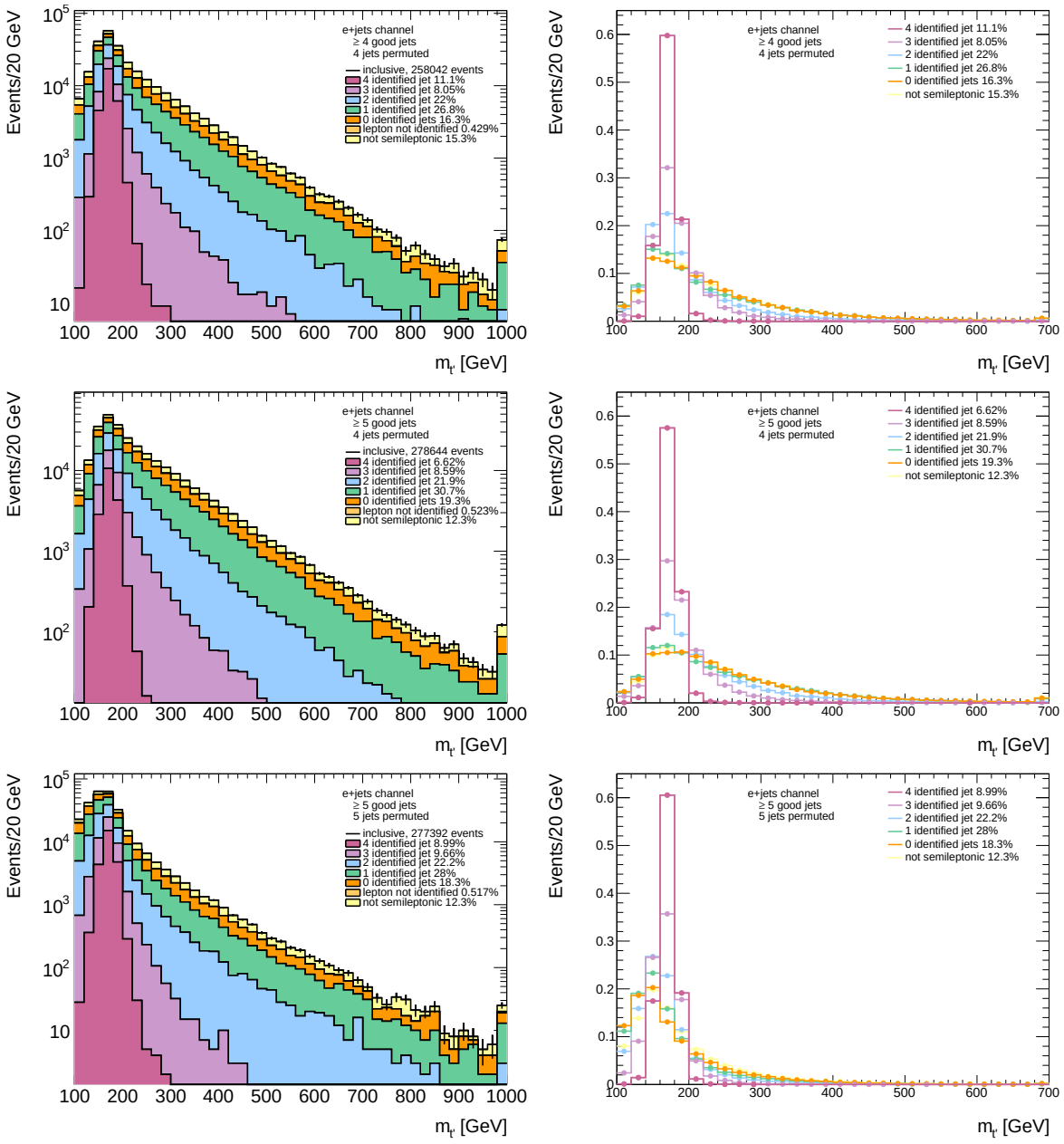


Figure B.9: Distribution of the fitted top quark mass decomposed according to the number of identified reconstructed jets shown in a stacked plot (left) and as individual histograms normalised to unity. The distributions are shown for the samples with $\geqslant 4$ (top) and $\geqslant 5$ (middle and bottom) reconstructed jets in the event in the e+jets channel. The distributions have been produced by fitting four (top and middle) or five (bottom) leading jets in the nominal $t\bar{t}$ sample.

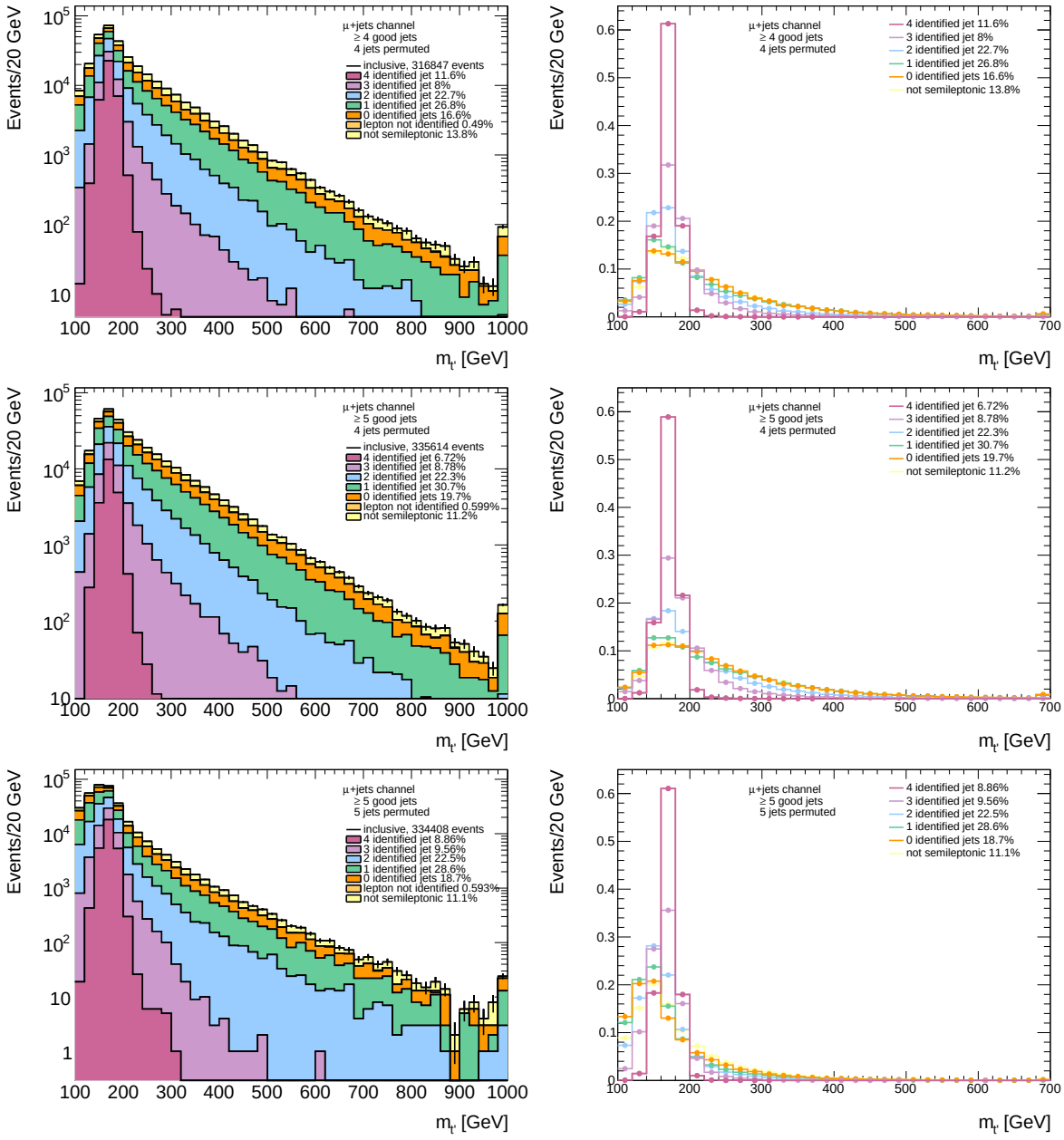


Figure B.10: Distribution of the fitted top quark mass decomposed according to the number of identified reconstructed jets shown in a stacked plot (left) and as individual histograms normalised to unity. The distributions are shown for the samples with $\geqslant 4$ (top) and $\geqslant 5$ (middle and bottom) reconstructed jets in the event in the μ+jets channel. The distributions have been produced by fitting four (top and middle) or five (bottom) leading jets in the nominal $t\bar{t}$ sample.

Appendix C

Additional Information for Chapter 9

C.1 List of Used MC Samples

The following simulation samples have been used in this analyses in the NTUP_TOP ntuple format:

- t' Signal samples:

  mc11_7TeV.115420.Pythia_u4u4_350_b.merge.NTUP_TOP.e972_s1310_s1300_r3043_r2993_p834
  mc11_7TeV.115423.Pythia_u4u4_400_b.merge.NTUP_TOP.e972_s1310_s1300_r3043_r2993_p834
  mc11_7TeV.115426.Pythia_u4u4_450_b.merge.NTUP_TOP.e972_s1310_s1300_r3043_r2993_p834
  mc11_7TeV.115429.Pythia_u4u4_500_b.merge.NTUP_TOP.e972_s1310_s1300_r3043_r2993_p834
  mc11_7TeV.119682.Pythia_u4u4_550_b.merge.NTUP_TOP.e997_s1372_s1370_r3043_r2993_p834
  mc11_7TeV.119685.Pythia_u4u4_600_b.merge.NTUP_TOP.e997_s1372_s1370_r3043_r2993_p834
  mc11_7TeV.119688.Pythia_u4u4_650_b.merge.NTUP_TOP.e997_s1372_s1370_r3043_r2993_p937
  mc11_7TeV.158100.Pythia_u4u4_700_b.merge.NTUP_TOP.e1301_s1372_s1370_r3108_r3109_p937
  mc11_7TeV.158101.Pythia_u4u4_750_b.merge.NTUP_TOP.e1301_s1372_s1370_r3108_r3109_p937

- $t\bar{t}$ samples:

  Nominal sample:
  mc11_7TeV.105200.T1_McAtNlo_Jimmy.merge.NTUP_TOP.e835_s1272_s1274_r3043_r2993

  Systematic samples:
  mc11_7TeV.105200.T1_McAtNlo_Jimmy.merge.NTUP_TOP.e835_a131_s1353_a139_r2900_p834
  mc11_7TeV.117209.AcerMCttbar_MorePS.merge.NTUP_TOP.e1029_a131_s1353_a145_r2993_p834
  mc11_7TeV.117210.AcerMCttbar_LessPS.merge.NTUP_TOP.e1029_a131_s1353_a145_r2993_p834
  mc11_7TeV.105860.TTbar_PowHeg_Jimmy.merge.NTUP_TOP.e873_a131_s1353_a139_r2900_p834
  mc11_7TeV.105861.TTbar_PowHeg_Pythia.merge.NTUP_TOP.e873_a131_s1353_a139_r2900_p834

- W+jets samples:
  mc11_7TeV.107280.AlpgenJimmyWbbFullNp0_pt20.merge.NTUP_TOP.e887_s1310_s1300_r3043_r2993_p834
  mc11_7TeV.107281.AlpgenJimmyWbbFullNp1_pt20.merge.NTUP_TOP.e887_s1310_s1300_r3043_r2993_p834
  mc11_7TeV.107282.AlpgenJimmyWbbFullNp2_pt20.merge.NTUP_TOP.e887_s1310_s1300_r3043_r2993_p834

mc11_7TeV.107283.AlpgenJimmyWbbFullNp3_pt20.merge.NTUP_TOP.e887_s1310_s1300_r3043_r2993_p834
mc11_7TeV.107680.AlpgenJimmyWenuNp0_pt20.merge.NTUP_TOP.e825_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107681.AlpgenJimmyWenuNp1_pt20.merge.NTUP_TOP.e825_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107682.AlpgenJimmyWenuNp2_pt20.merge.NTUP_TOP.e825_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107683.AlpgenJimmyWenuNp3_pt20.merge.NTUP_TOP.e825_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107684.AlpgenJimmyWenuNp4_pt20.merge.NTUP_TOP.e825_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107685.AlpgenJimmyWenuNp5_pt20.merge.NTUP_TOP.e825_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107690.AlpgenJimmyWmunuNp0_pt20.merge.NTUP_TOP.e825_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107691.AlpgenJimmyWmunuNp1_pt20.merge.NTUP_TOP.e825_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107692.AlpgenJimmyWmunuNp2_pt20.merge.NTUP_TOP.e825_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107693.AlpgenJimmyWmunuNp3_pt20.merge.NTUP_TOP.e825_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107694.AlpgenJimmyWmunuNp4_pt20.merge.NTUP_TOP.e825_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107695.AlpgenJimmyWmunuNp5_pt20.merge.NTUP_TOP.e825_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107700.AlpgenJimmyWtaunuNp0_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107701.AlpgenJimmyWtaunuNp1_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107702.AlpgenJimmyWtaunuNp2_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107703.AlpgenJimmyWtaunuNp3_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107704.AlpgenJimmyWtaunuNp4_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107705.AlpgenJimmyWtaunuNp5_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.117284.AlpgenWccFullNp0_pt20.merge.NTUP_TOP.e887_s1310_s1300_r3043_r2993_p834
mc11_7TeV.117285.AlpgenWccFullNp1_pt20.merge.NTUP_TOP.e887_s1310_s1300_r3043_r2993_p834
mc11_7TeV.117286.AlpgenWccFullNp2_pt20.merge.NTUP_TOP.e887_s1310_s1300_r3043_r2993_p834
mc11_7TeV.117287.AlpgenWccFullNp3_pt20.merge.NTUP_TOP.e887_s1310_s1300_r3043_r2993_p834
mc11_7TeV.117293.AlpgenWcNp0_pt20.merge.NTUP_TOP.e887_s1310_s1300_r3043_r2993_p834
mc11_7TeV.117294.AlpgenWcNp1_pt20.merge.NTUP_TOP.e887_s1310_s1300_r3043_r2993_p834
mc11_7TeV.117295.AlpgenWcNp2_pt20.merge.NTUP_TOP.e887_s1310_s1300_r3043_r2993_p834
mc11_7TeV.117296.AlpgenWcNp3_pt20.merge.NTUP_TOP.e887_s1310_s1300_r3043_r2993_p834
mc11_7TeV.117297.AlpgenWcNp4_pt20.merge.NTUP_TOP.e887_s1310_s1300_r3043_r2993_p834

- Z+jets samples:

mc11_7TeV.109300.AlpgenJimmyZeebbNp0_nofilter.merge.NTUP_TOP.e835_s1310_s1300_r3043_r2993_p834
mc11_7TeV.109301.AlpgenJimmyZeebbNp1_nofilter.merge.NTUP_TOP.e835_s1310_s1300_r3043_r2993_p834
mc11_7TeV.109302.AlpgenJimmyZeebbNp2_nofilter.merge.NTUP_TOP.e835_s1310_s1300_r3043_r2993_p834
mc11_7TeV.109303.AlpgenJimmyZeebbNp3_nofilter.merge.NTUP_TOP.e835_s1310_s1300_r3043_r2993_p834

mc11_7TeV.107650.AlpgenJimmyZeeNp0_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107651.AlpgenJimmyZeeNp1_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107652.AlpgenJimmyZeeNp2_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107653.AlpgenJimmyZeeNp3_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107654.AlpgenJimmyZeeNp4_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107655.AlpgenJimmyZeeNp5_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834

mc11_7TeV.109305.AlpgenJimmyZmumubbNp0_nofilter.merge.NTUP_TOP.e835_s1310_s1300_r3043_r2993_p834
mc11_7TeV.109306.AlpgenJimmyZmumubbNp1_nofilter.merge.NTUP_TOP.e835_s1310_s1300_r3043_r2993_p834
mc11_7TeV.109307.AlpgenJimmyZmumubbNp2_nofilter.merge.NTUP_TOP.e835_s1310_s1300_r3043_r2993_p834
mc11_7TeV.109308.AlpgenJimmyZmumubbNp3_nofilter.merge.NTUP_TOP.e835_s1310_s1300_r3043_r2993_p834

mc11_7TeV.107660.AlpgenJimmyZmumuNp0_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107661.AlpgenJimmyZmumuNp1_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107662.AlpgenJimmyZmumuNp2_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107663.AlpgenJimmyZmumuNp3_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107664.AlpgenJimmyZmumuNp4_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107665.AlpgenJimmyZmumuNp5_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.109310.AlpgenJimmyZtautaubbNp0_nofilter.merge.NTUP_TOP.e835_s1310_s1300_r3043_r2993_p834
mc11_7TeV.109311.AlpgenJimmyZtautaubbNp1_nofilter.merge.NTUP_TOP.e835_s1310_s1300_r3043_r2993_p834
mc11_7TeV.109312.AlpgenJimmyZtautaubbNp2_nofilter.merge.NTUP_TOP.e835_s1310_s1300_r3043_r2993_p834

mc11_7TeV.109313.AlpgenJimmyZtautaubbNp3_nofilter.merge.NTUP_TOP.e835_s1310_s1300_r3043_r2993_p834
mc11_7TeV.107670.AlpgenJimmyZtautauNp0_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107671.AlpgenJimmyZtautauNp1_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107672.AlpgenJimmyZtautauNp2_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107673.AlpgenJimmyZtautauNp3_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107674.AlpgenJimmyZtautauNp4_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834
mc11_7TeV.107675.AlpgenJimmyZtautauNp5_pt20.merge.NTUP_TOP.e835_s1299_s1300_r3043_r2993_p834

- Single top samples:
  mc11_7TeV.108343.st_schan_enu_McAtNlo_Jimmy.merge.NTUP_TOP.e825_s1310_s1300_r3043_r2993_p834
  mc11_7TeV.108344.st_schan_munu_McAtNlo_Jimmy.merge.NTUP_TOP.e825_s1310_s1300_r3043_r2993_p834
  mc11_7TeV.108345.st_schan_taunu_McAtNlo_Jimmy.merge.NTUP_TOP.e835_s1310_s1300_r3043_r2993_p834
  mc11_7TeV.117360.st_tchan_enu_AcerMC.merge.NTUP_TOP.e835_s1310_s1300_r3043_r2993_p834
  mc11_7TeV.117361.st_tchan_munu_AcerMC.merge.NTUP_TOP.e835_s1310_s1300_r3043_r2993_p834
  mc11_7TeV.117362.st_tchan_taunu_AcerMC.merge.NTUP_TOP.e825_s1310_s1300_r3043_r2993_p834
  mc11_7TeV.108346.st_Wt_McAtNlo_Jimmy.merge.NTUP_TOP.e835_s1310_s1300_r3043_r2993_p834
- Di-boson samples:
  mc11_7TeV.105985.WW_Herwig.merge.NTUP_TOP.e825_s1310_s1300_r3043_r2993_p834
  mc11_7TeV.105987.WZ_Herwig.merge.NTUP_TOP.e825_s1310_s1300_r3043_r2993_p834
  mc11_7TeV.105986.ZZ_Herwig.merge.NTUP_TOP.e825_s1310_s1300_r3043_r2993_p834

C.2 Splitting Validation

The splitting of the MC sample passing the W_{had}-Selection selection into a training (A) and a testing (B) sample has been performed as described in Section 9.3.2. A comparison between the distributions of the NN input variables and the M_{reco} discriminant for the samples A and B is demonstrated in the plots in Figures C.1-C.3. The distribution of each variable is shown for unweighted events on the left and the expected as well as observed event yield on the right. Most bins of all variables agree within $1\sigma_{stat}$. The bins with the discrepancy slightly higher than $1\sigma_{stat}$ are mostly those with poor MC statistics, such as those of the high tail of the $\Delta R^{l\nu}$ variable. These bins are however not expected to harm the NN training in the signal region as they are in the region with very few signal events.

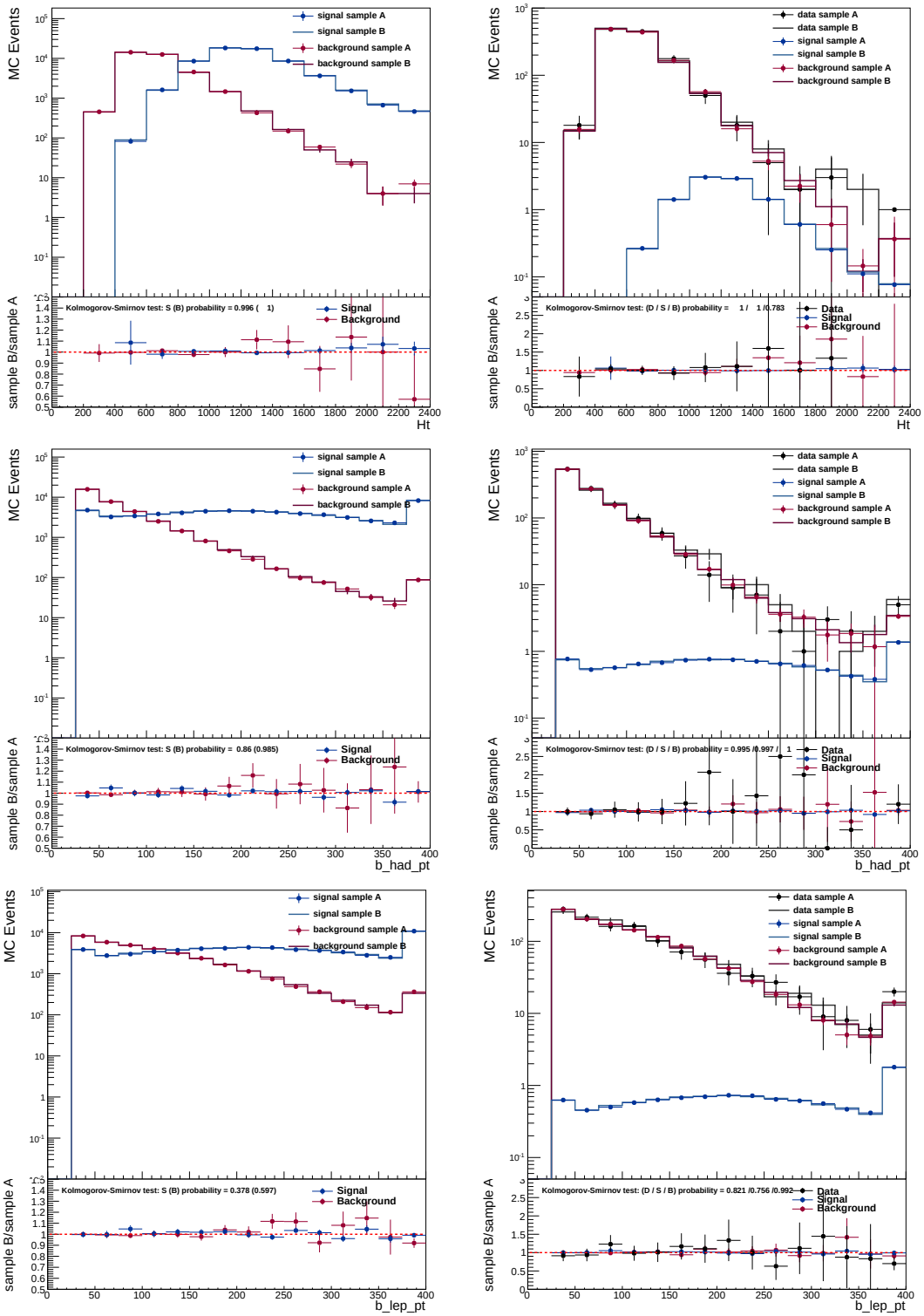


Figure C.1: Comparison of the distributions of the NN input variables in the subsamples A and B for the data, the signal with the T-quark mass of 650 GeV and the combined MC backgrounds separately

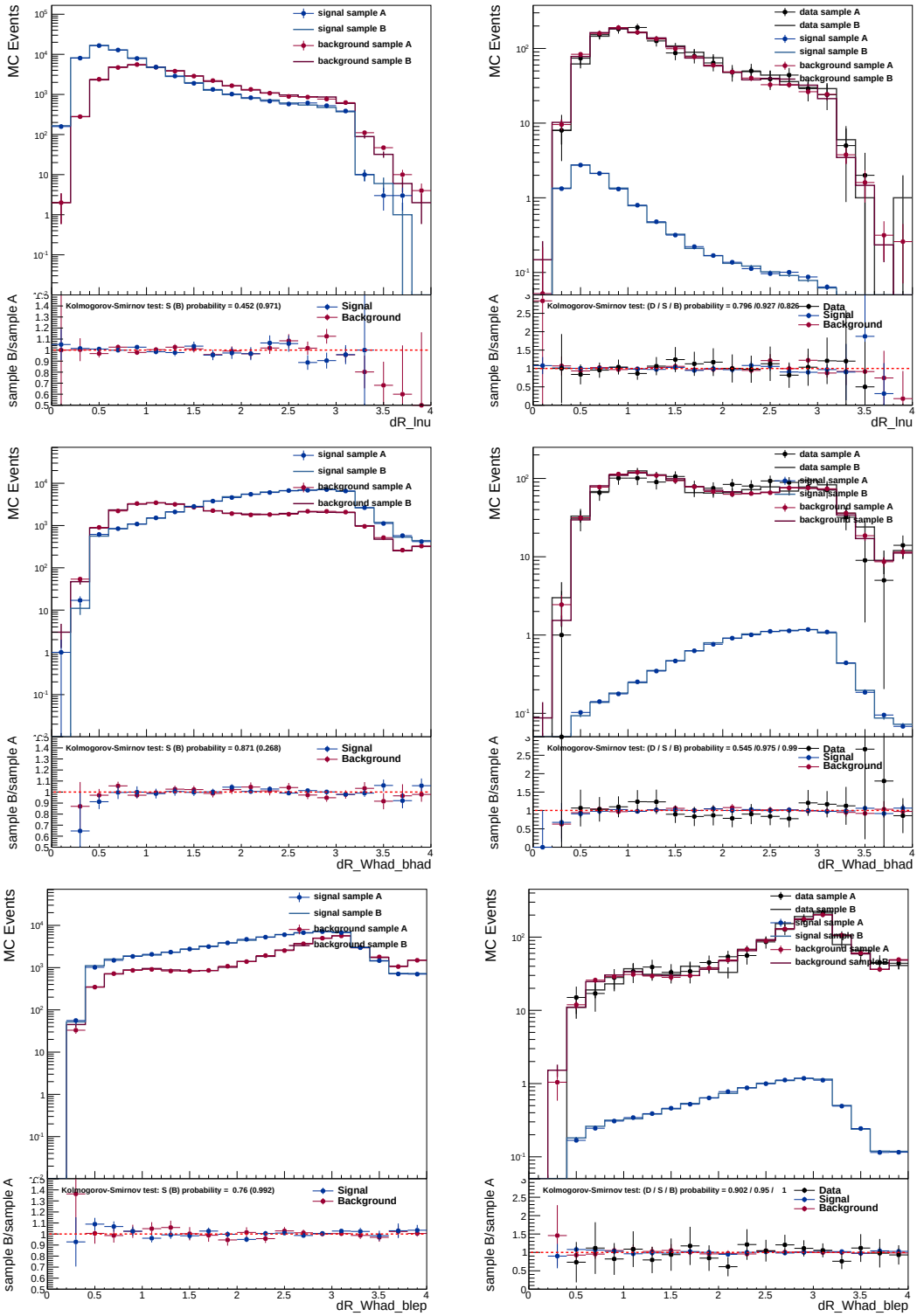


Figure C.2: Comparison of the distributions of the NN input variables in the sub-samples A and B for the data, the signal with the T-quark mass of 650 GeV and the combined MC backgrounds separately.

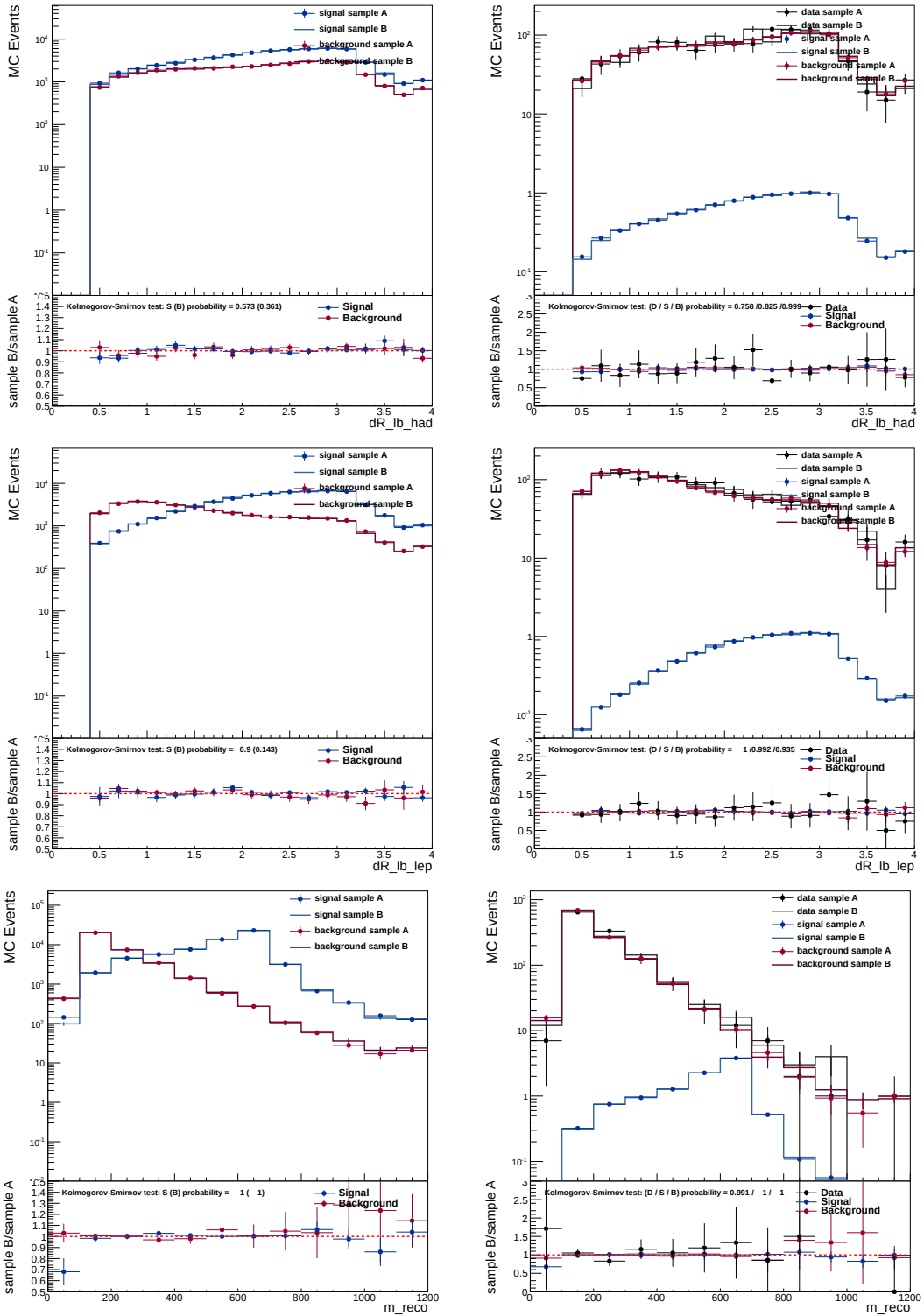


Figure C.3: Comparison of the distributions of the NN input variables in the subsamples A and B for the data, the signal with the T-quark mass of 650 GeV and the combined MC backgrounds separately.

C.3 Training Statistics Study

For the purpose of the NN training a privately produced $T\bar{T}$ signal sample of 1M events has been used instead of the officially produced sample of 75k events. In this section the motivation for the usage of a high statistic sample is explained.

Table C.1 summarises the sample information as well as the training statistics. In case of a training with the low statistics signal sample only 9.2k MC signal events are available which are split equally into the training and testing samples. Having such low statistics also limits the training statistics of the background. In case of a too large statistics asymmetry between the signal and background samples in the training, the estimator calculation (Equation 6.12) either breaks down or produces an unreasonable outcome. Thus only a portion of the available background statistics can be used for the training and the rest for testing. In case of the training with the high statistics signal sample however, one half of the available statistics is used for the training and the other one for testing for both signal and background.

		$T\bar{T}$ 650 high stat	$T\bar{T}$ 650 low stat	MC bgd
Sample info	Initial sample stat	1M	75k	104M
	Pretraining selection stat	122k	9.2k	68k
	Exected event yield	20.2	20.3	2370.9
low stat training	training events	–	4.6k	6.1k
	testing events	–	4.6k	61.9k
high stat training	training events	61k	–	34k
	testing events	61k	–	34k

Table C.1: MC statistics, event yields and the training statistics of the samples used in training: combined MC backgrounds as well as high and low statistics $T\bar{T}$ samples with the T mass of 650 GeV.

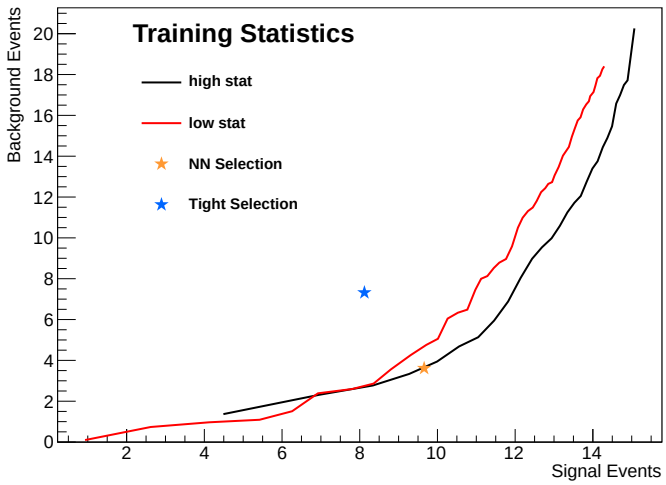


Figure C.4: Response function comparison between trainings with high (black) and low (red) signal training statistics.

The difference in the performance between the low and high statistics trainings is shown in terms of the response function, i.e. the background acceptance as a function of the signal acceptance

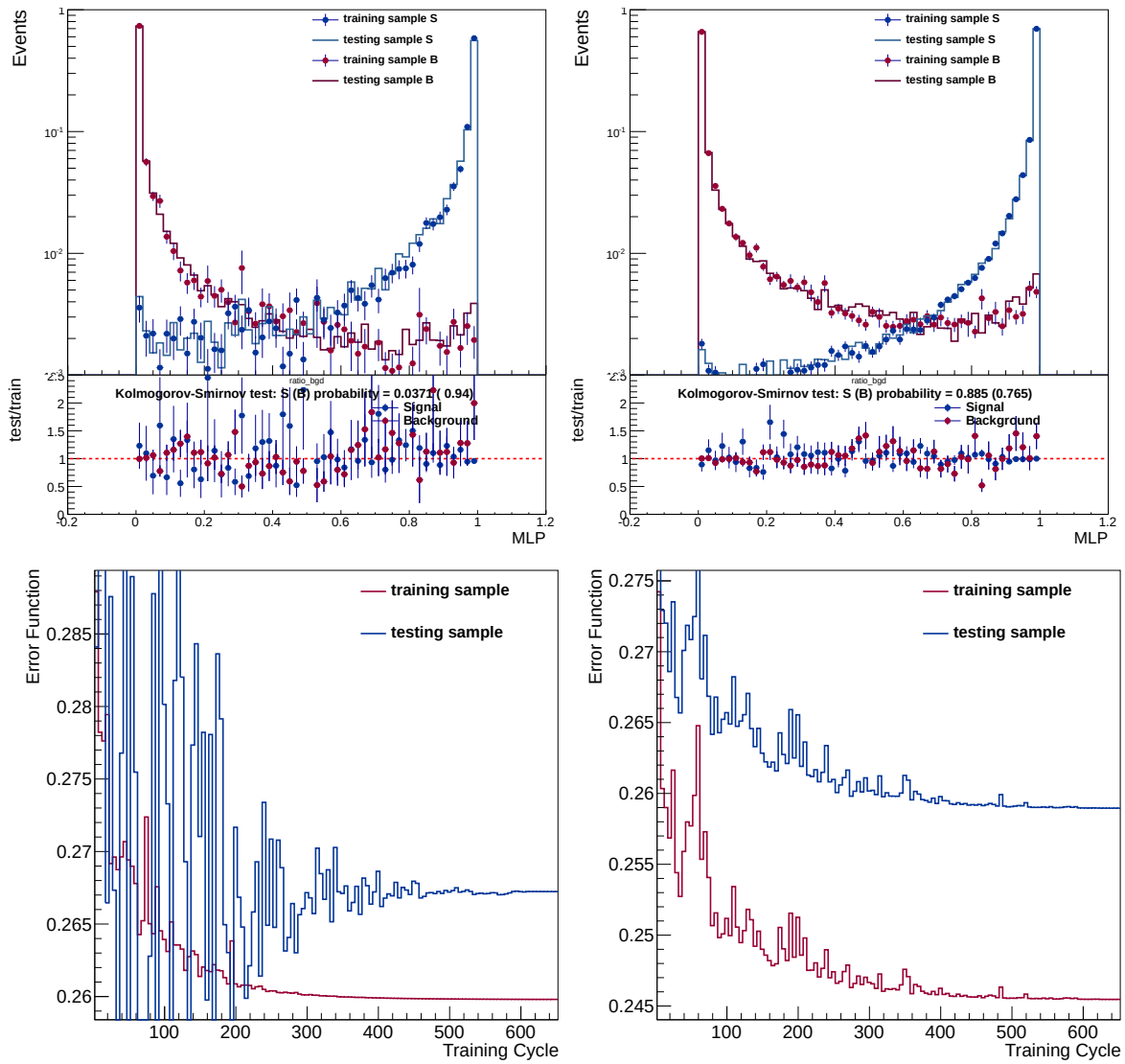


Figure C.5: The NN response (top row) and the error function evolution (bottom row) for high (left) and low (right) training statistics.

in Figure C.4. Although the response curves in both cases have seemingly similar response, the curve corresponding to the low statistics training has a more irregular shape while the curve corresponding to the high statistics training looks much smoother.

That can be better understood from the NN response distribution (Figure C.5, top row) as well as from the error function evolution through the training cycles (Figure C.5, bottom row). The top row plots clearly show that the NN response distribution has significantly smaller statistical fluctuations and a better agreement between the training and testing samples for the training with the high statistics samples. The bottom row plots demonstrate that the training with the low statistics samples makes more dramatic oscillations before it converges.

According to these results, with the given set of variables (8 variables in total) the low statistics training could be used, but the high statistics training is more reliable in sense of statistical fluctuations.

C.4 Data/MC Comparisons

C.4.1 NN Input Variables, Control Region: Pre-selection, $= 0\,b$-tags

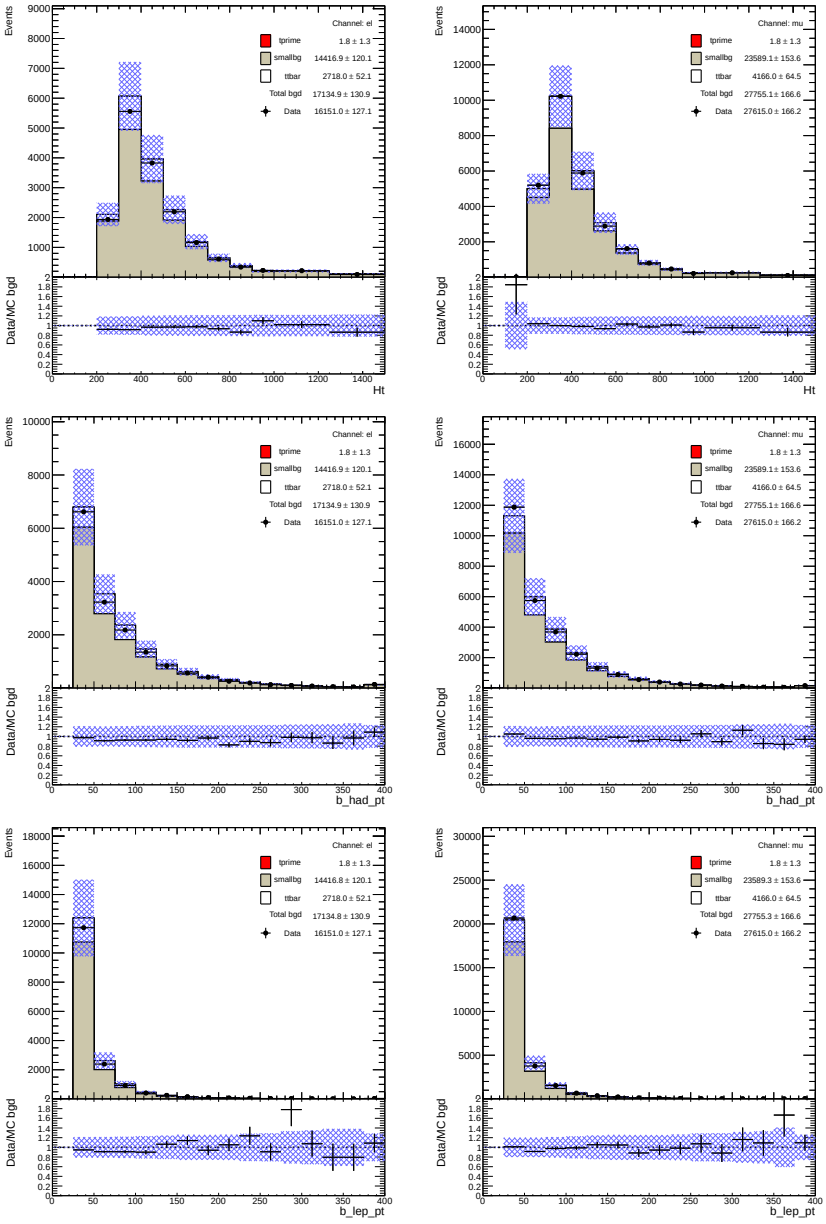


Figure C.6: Data/MC comparison of the NN input variables in the Control Region Pre-selection, $= 0\,b$-tags for the e-Channel (left) and μ-Channel (right) separately: total transverse momentum H_{T} (top), p_{T} of the leptonic b (middle), p_{T} of the hadronically decaying b (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

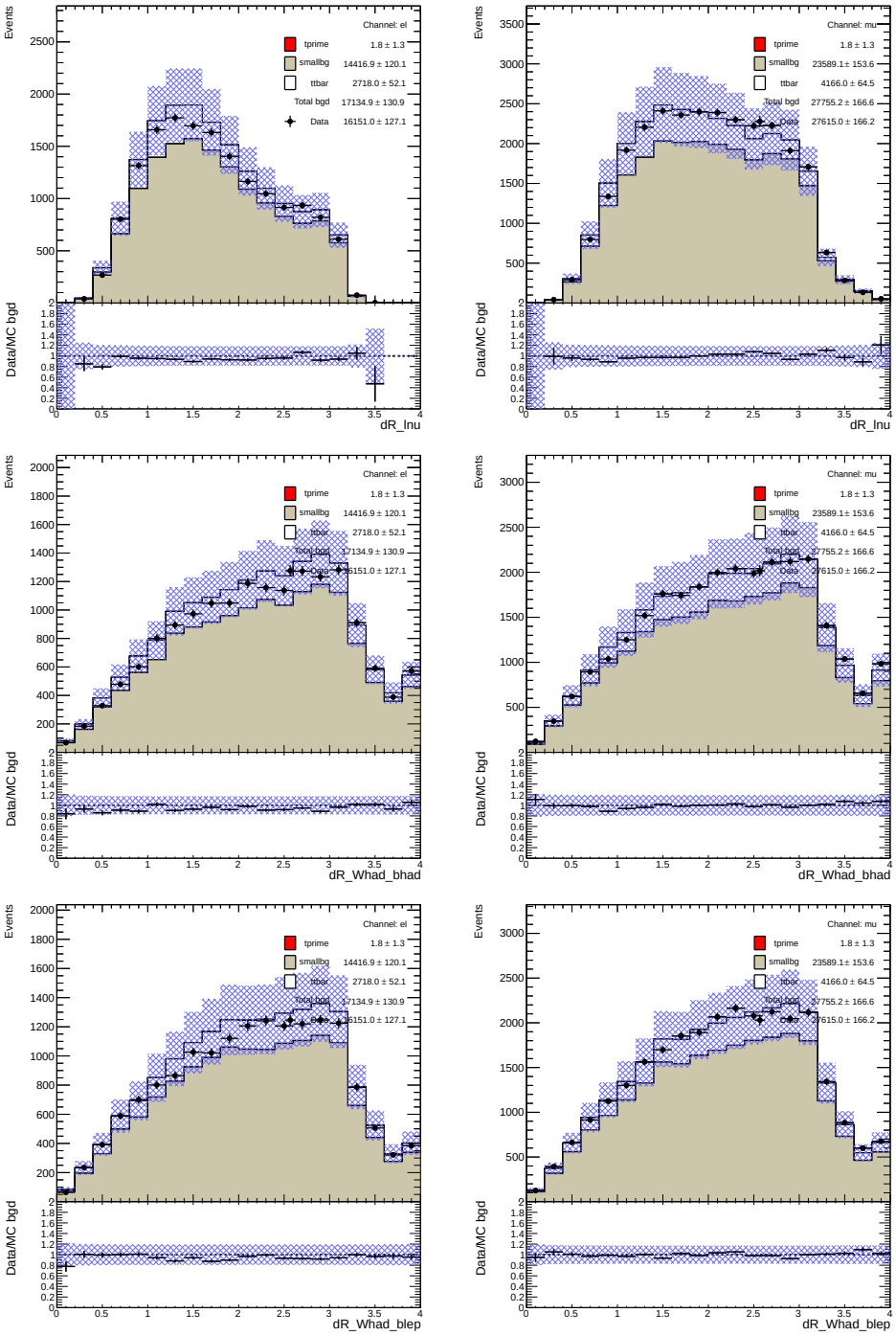


Figure C.7: Data/MC comparison of the NN input variables in the Control Region Pre-selection, $= 0\,b$-tags for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{l\nu}$ (top), $\Delta R^{W_{\mathrm{had}}b_{\mathrm{had}}}$ (middle), $\Delta R^{W_{\mathrm{had}}b_{\mathrm{lep}}}$ (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

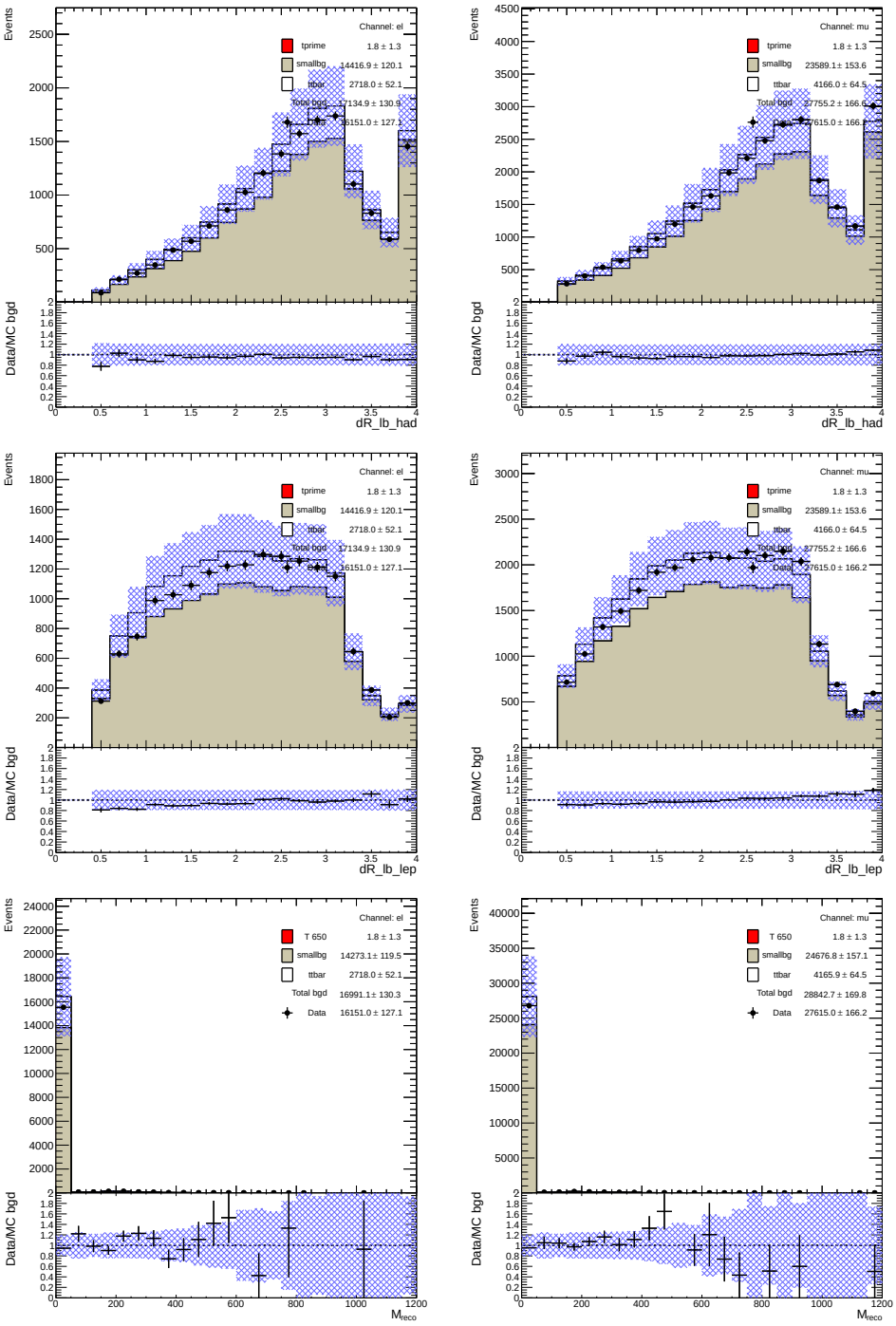


Figure C.8: Data/MC comparison of the NN input variables in the Control Region Pre-selection, $= 0\,b$-tags for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{lb_{\mathrm{had}}}$ (top), $\Delta R^{lb_{\mathrm{lep}}}$ (middle), M_{reco} (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

C.4.2 NN Input Variables, Control Region: Preselection, $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $M_{\text{reco}} < 350$ GeV

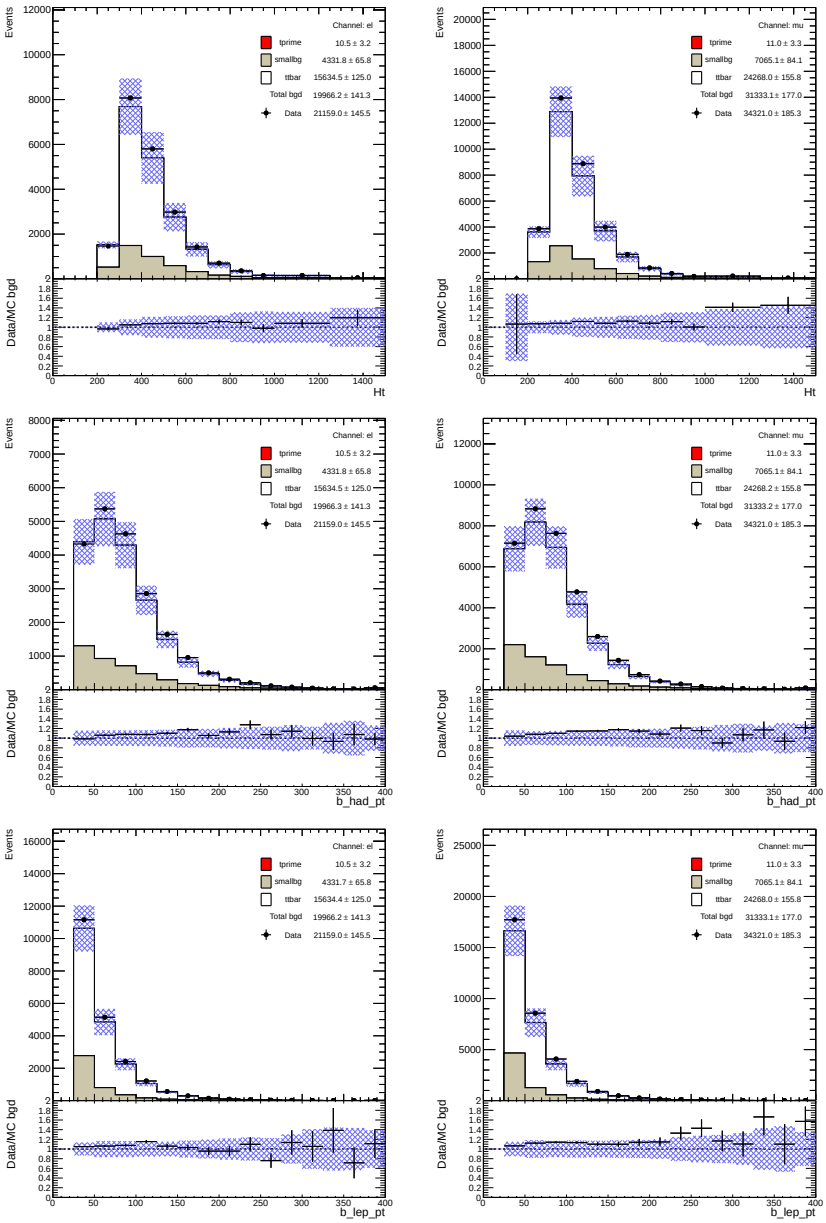


Figure C.9: Data/MC comparison of the NN input variables in the Control Region Pre-selection, $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $M_{\text{reco}} < 350$ GeV for the e-Channel (left) and μ-Channel (right) separately: total transverse momentum H_{T} (top), p_{T} of the leptonic b (middle), p_{T} of the hadronically decaying b (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

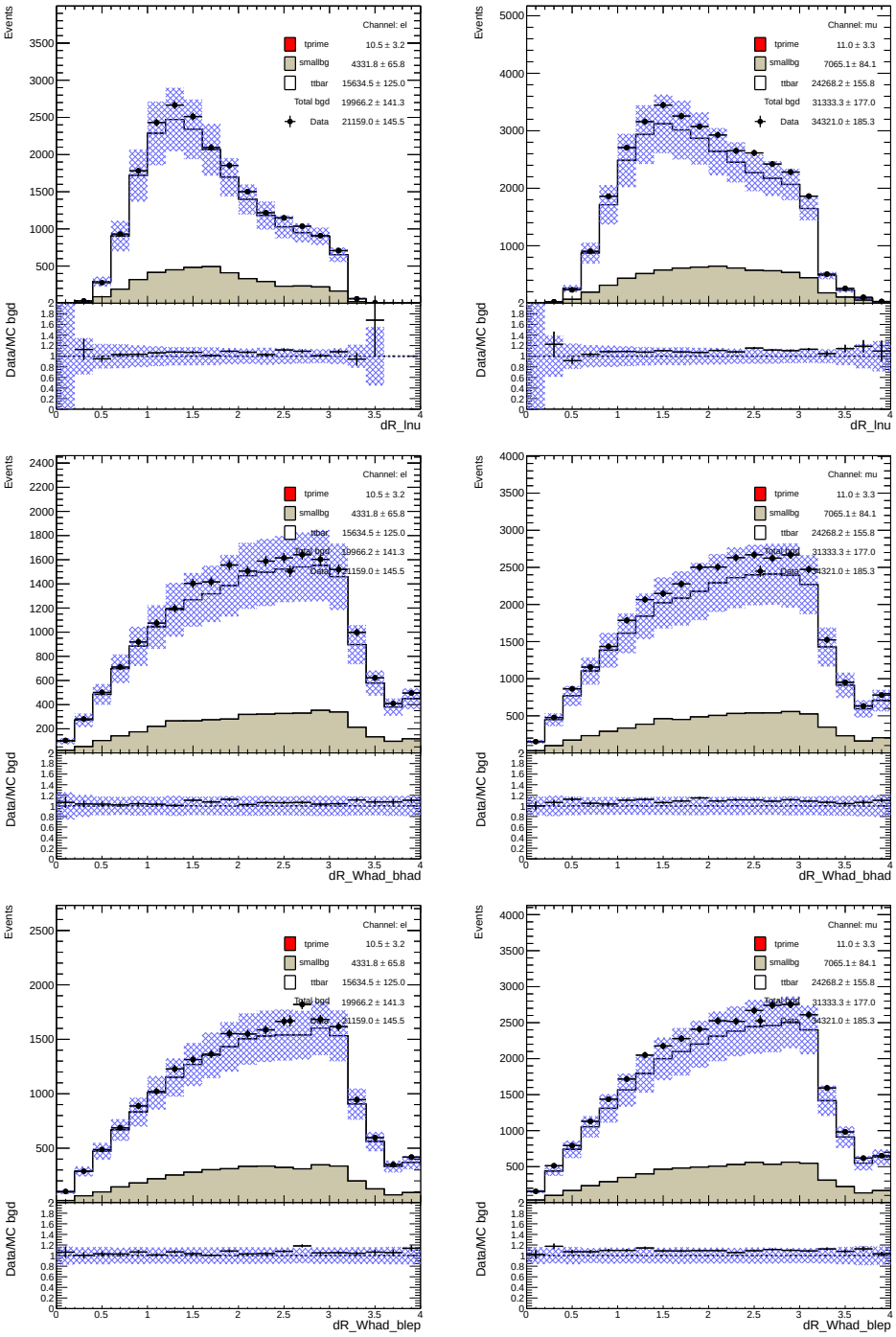


Figure C.10: Data/MC comparison of the NN input variables in the Control Region Preselection, $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $M_{\text{reco}} < 350$ GeV for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{l\nu}$ (top), $\Delta R^{W_{\text{had}}b_{\text{had}}}$ (middle), $\Delta R^{W_{\text{had}}b_{\text{lep}}}$ (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

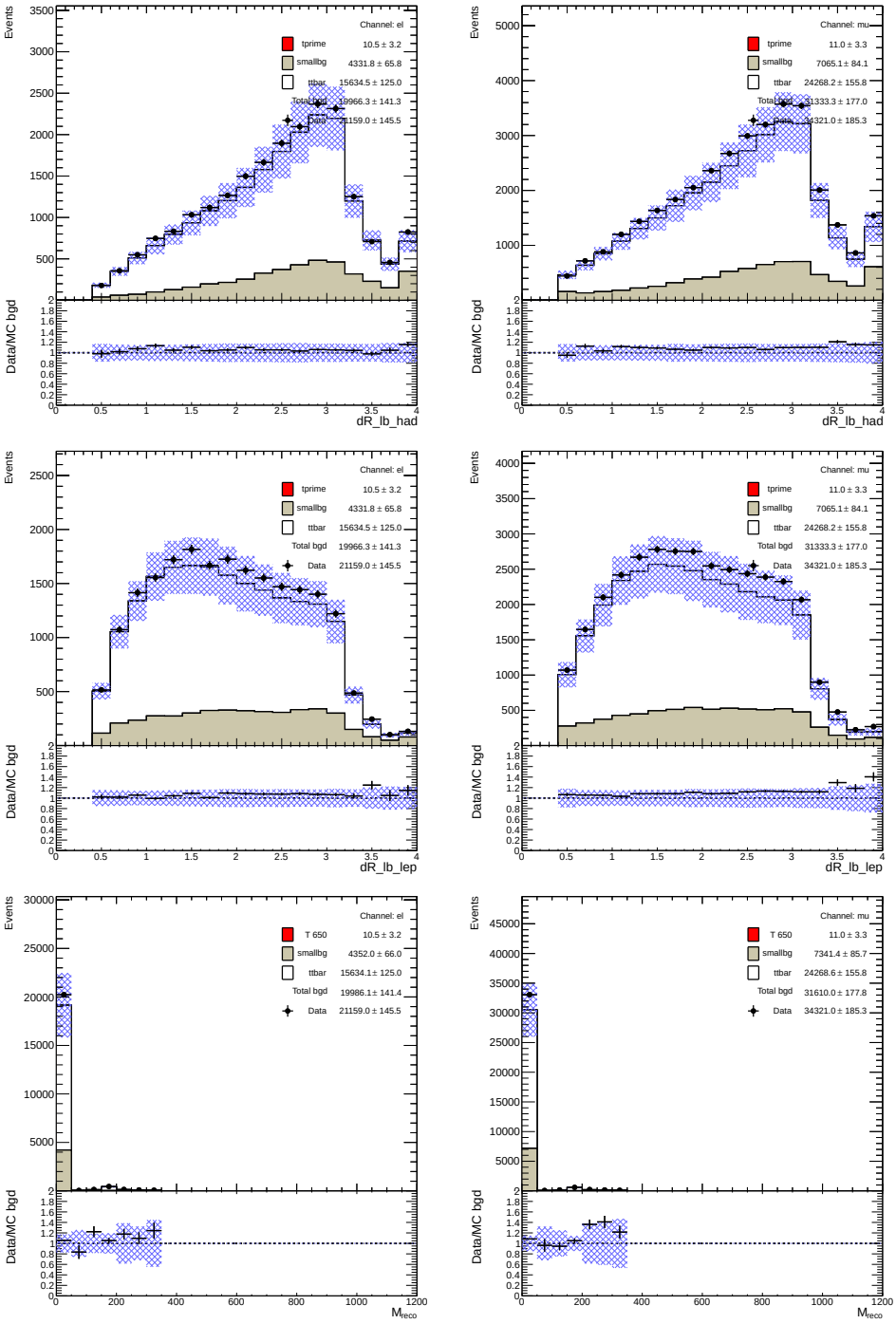


Figure C.11: Data/MC comparison of the NN input variables in the Control Region Preselection, $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $M_{\text{reco}} < 350$ GeV for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{lb_{\text{had}}}$ (top), $\Delta R^{lb_{\text{lep}}}$ (middle), M_{reco} (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

C.4.3 NN Input Variables, Control Region: Pre-selection, $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $H_{\mathrm{T}} < 700$ GeV

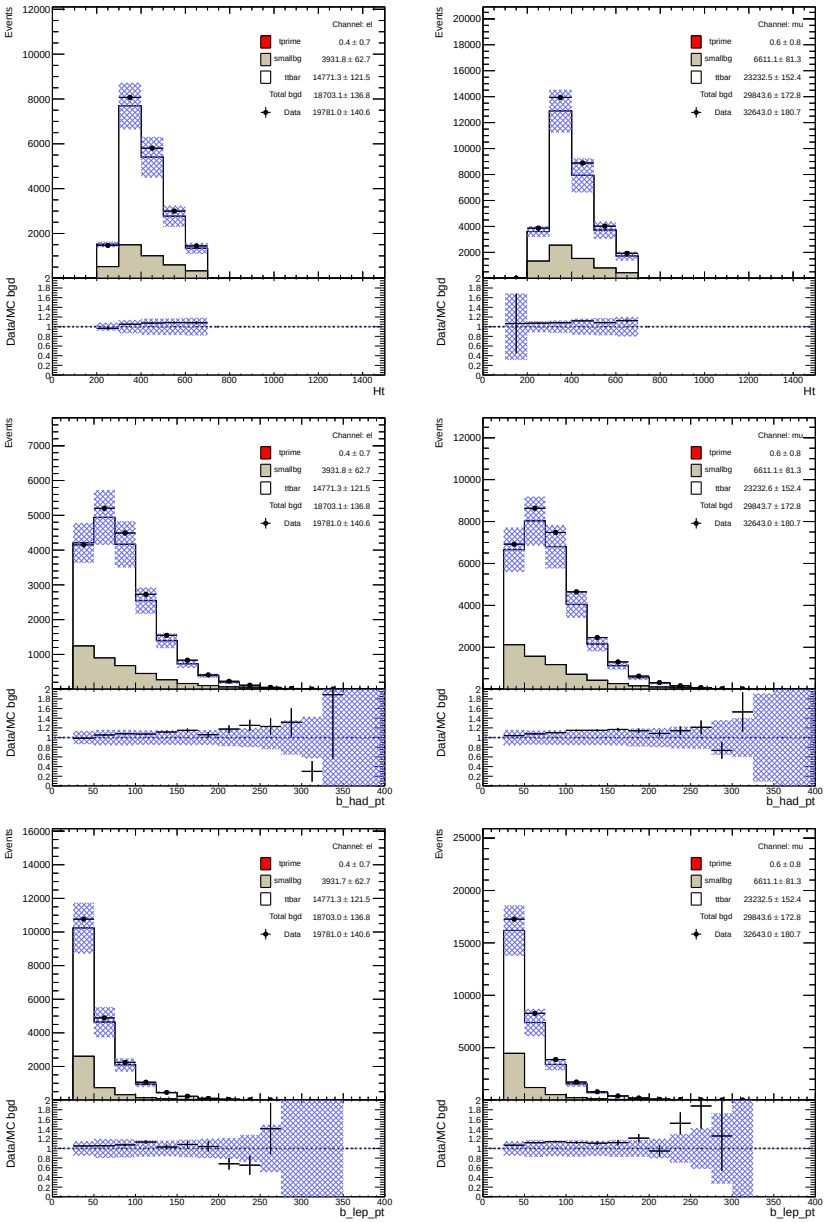


Figure C.12: Data/MC comparison of the NN input variables in the Control Region Pre-selection, $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $H_{\mathrm{T}} < 700$ GeV for the e-Channel (left) and μ-Channel (right) separately: total transverse momentum H_{T} (top), p_{T} of the leptonic b (middle), p_{T} of the hadronically decaying b (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

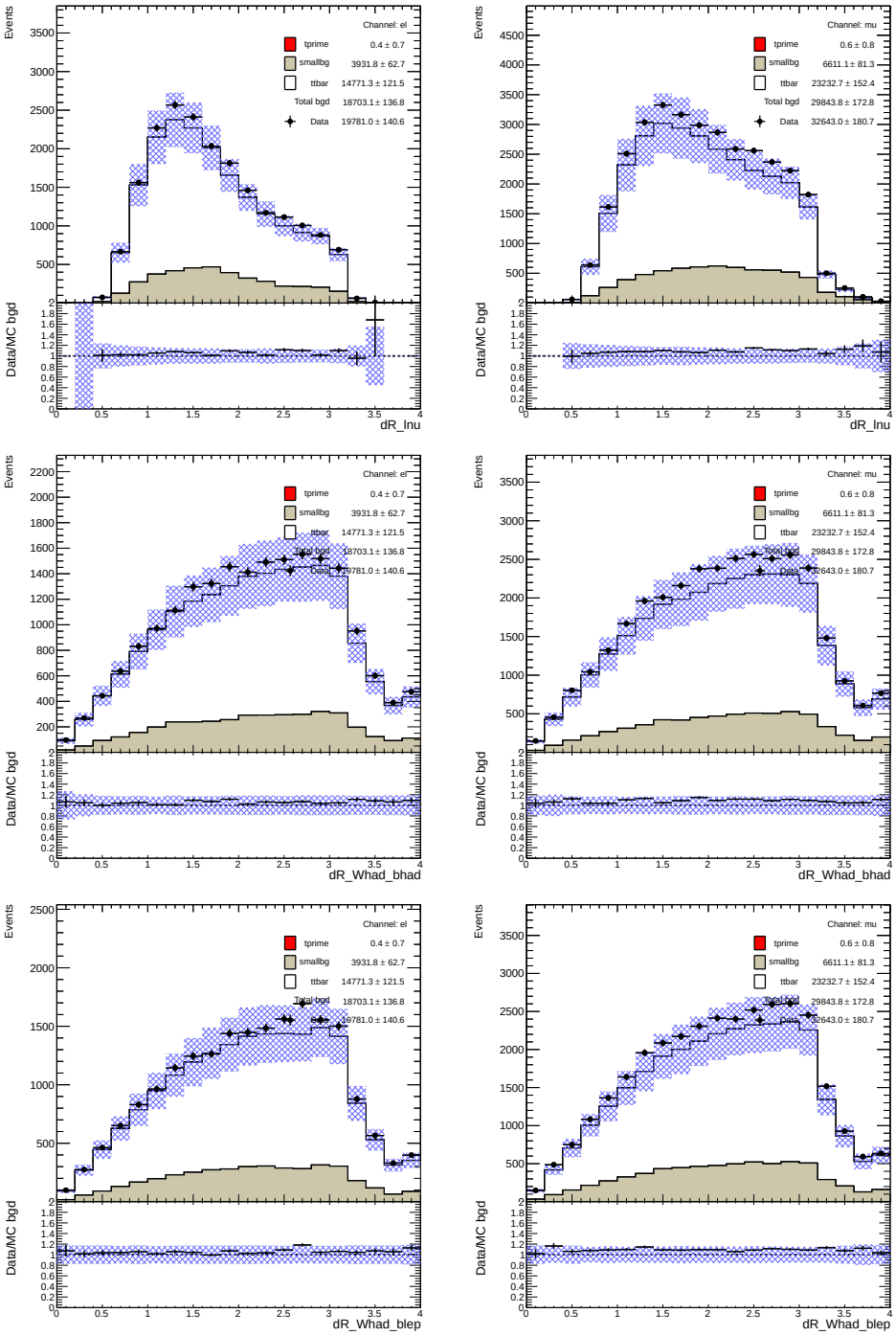


Figure C.13: Data/MC comparison of the NN input variables in the Control Region Preselection, $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $H_{\mathrm{T}} < 700$ GeV for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{l\nu}$ (top), $\Delta R^{W_{\mathrm{had}}b_{\mathrm{had}}}$ (middle), $\Delta R^{W_{\mathrm{had}}b_{\mathrm{lep}}}$ (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

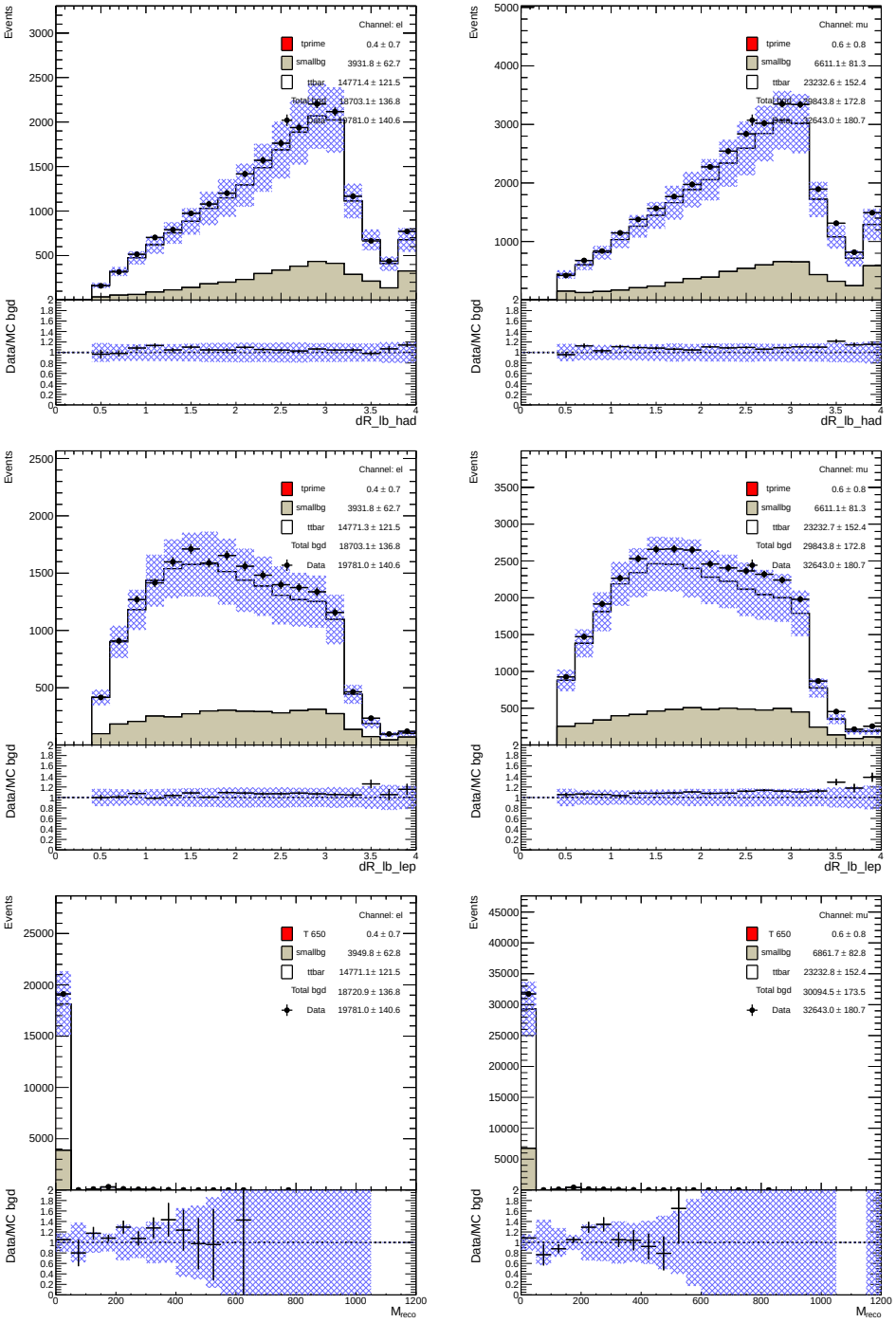


Figure C.14: Data/MC comparison of the NN input variables in the Control Region Preselection, $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $H_{\mathrm{T}} < 700$ GeV for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{lb_{\mathrm{had}}}$ (top), $\Delta R^{lb_{\mathrm{lep}}}$ (middle), M_{reco} (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

C.4.4 NN Input Variables, Pre-selection, full region

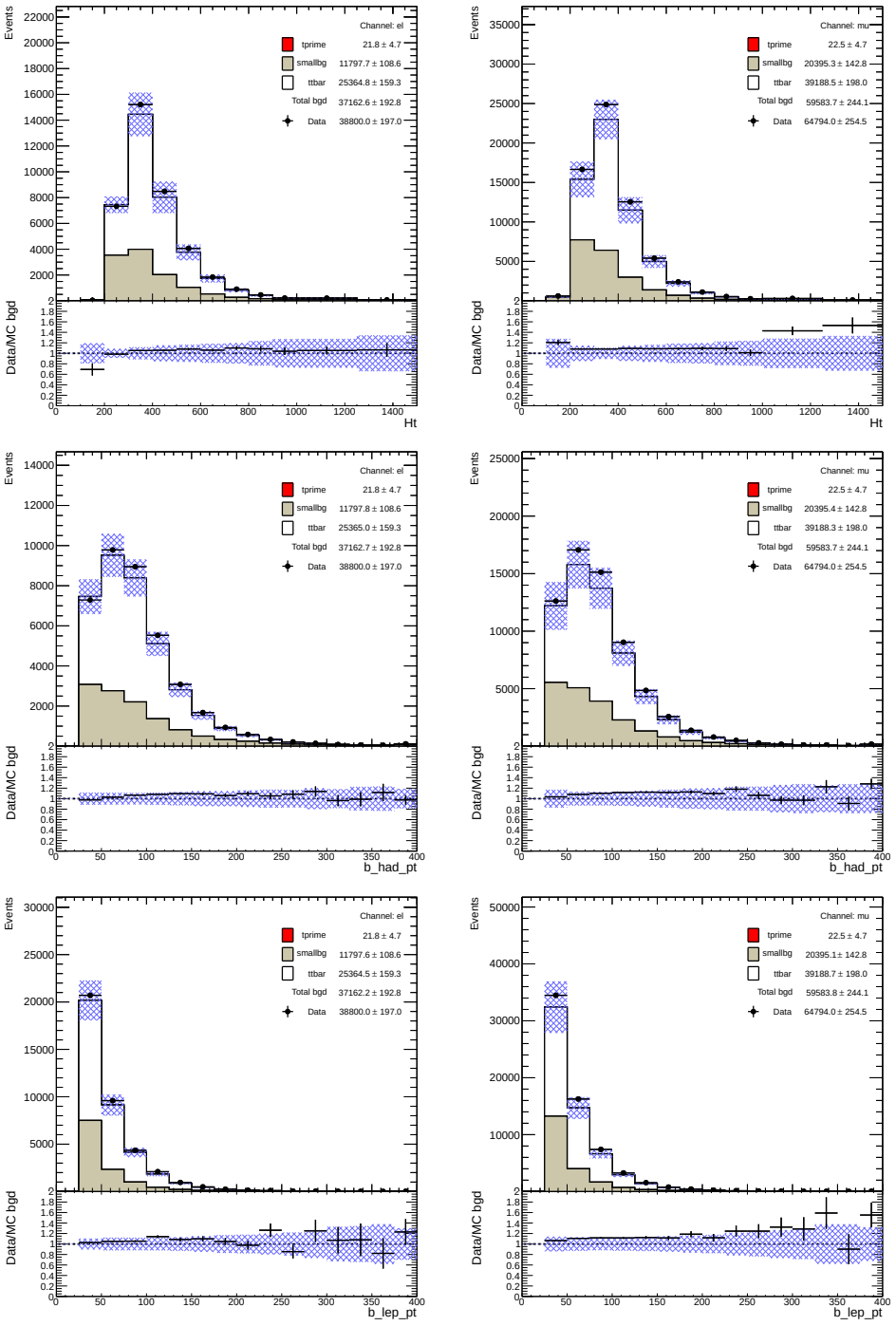


Figure C.15: Data/MC comparison of the NN input variables after Pre-selection for the e-Channel (left) and μ-Channel (right) separately: total transverse momentum H_{T} (top), p_{T} of the leptonic b (middle), p_{T} of the hadronically decaying b (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

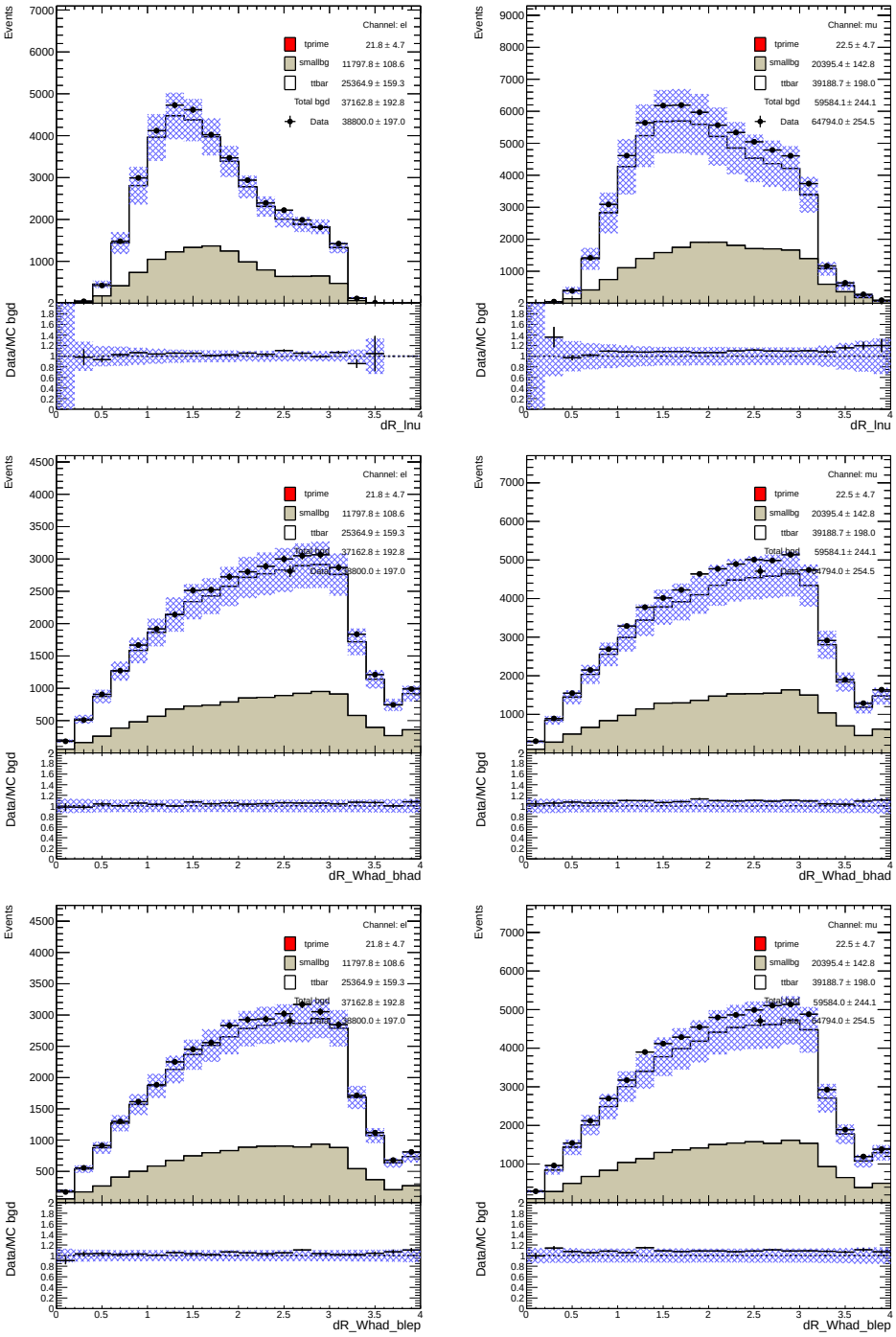


Figure C.16: Data/MC comparison of the NN input variables after Pre-selection for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{l\nu}$ (top), $\Delta R^{W_{\mathrm{had}}b_{\mathrm{had}}}$ (middle), $\Delta R^{W_{\mathrm{had}}b_{\mathrm{lep}}}$ (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

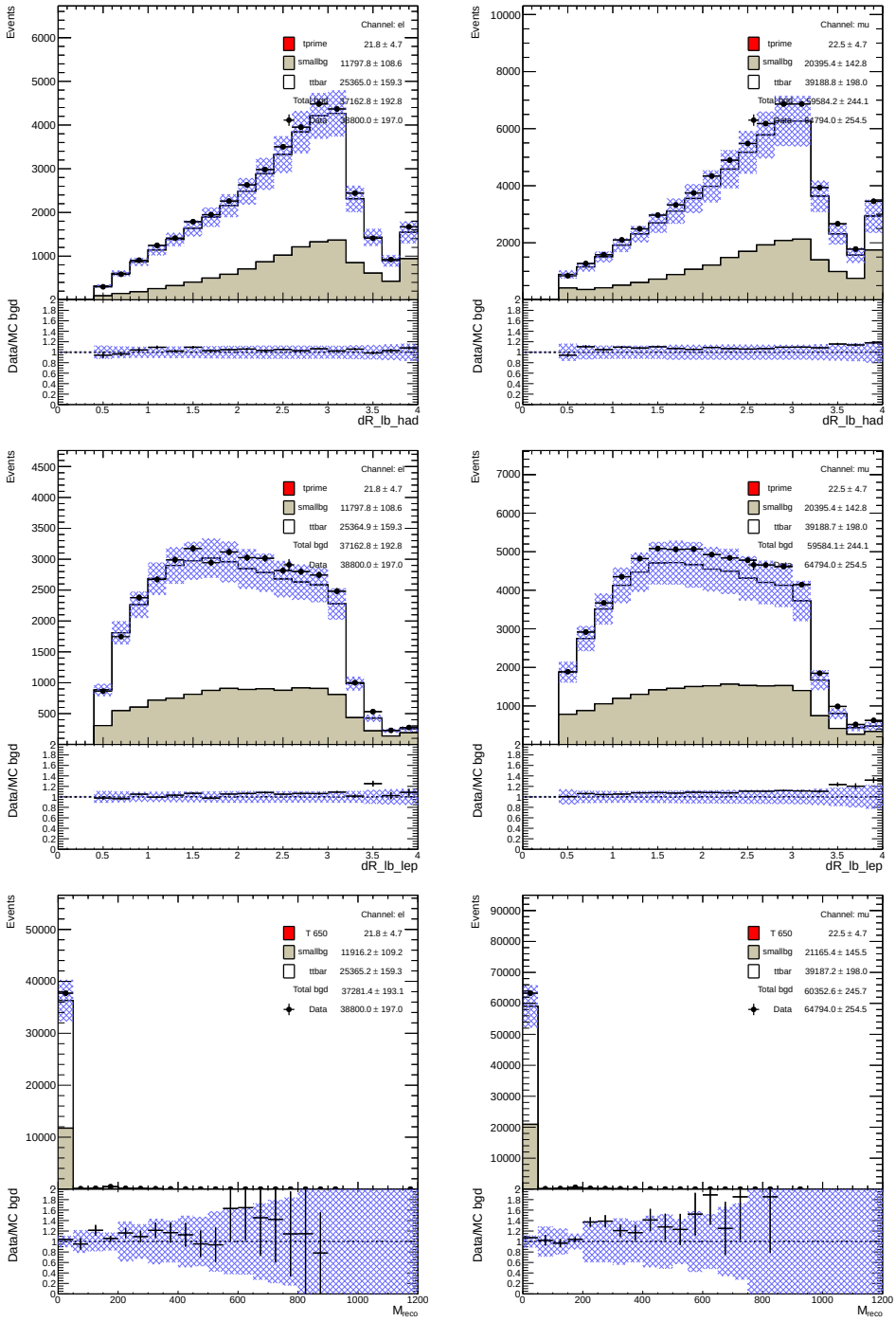


Figure C.17: Data/MC comparison of the NN input variables after Pre-selection for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{lb_{\mathrm{had}}}$ (top), $\Delta R^{lb_{\mathrm{lep}}}$ (middle), M_{reco} (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

C.4.5 NN Input Variables, Control Region: W_{had}-Selection, $= 0\,b$-tags

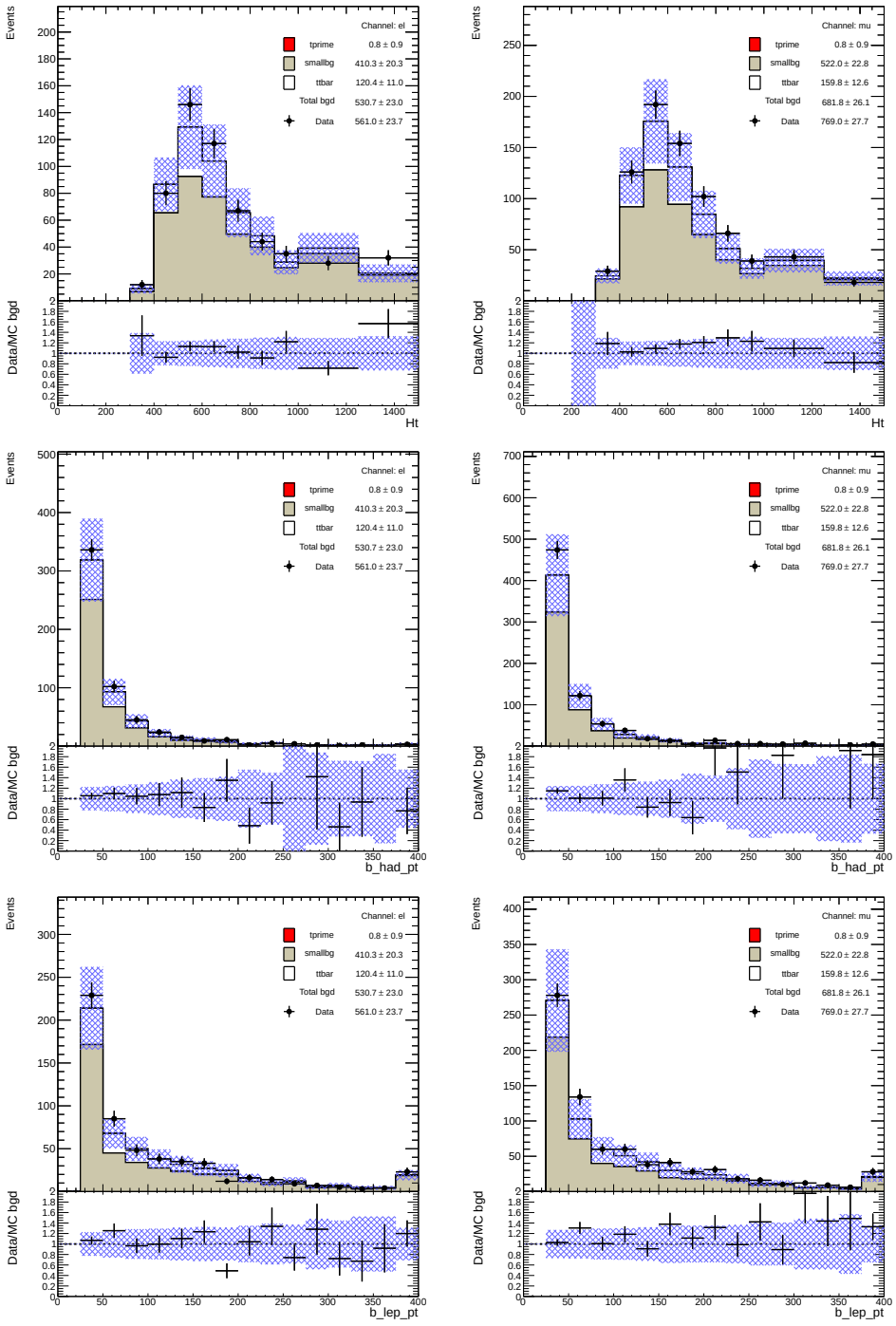


Figure C.18: Data/MC comparison of the NN input variables in the Control Region W_{had}-Selection, $= 0\,b$-tags for the e-Channel (left) and μ-Channel (right) separately: total transverse momentum H_{T} (top), p_{T} of the leptonic b (middle), p_{T} of the hadronically decaying b (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

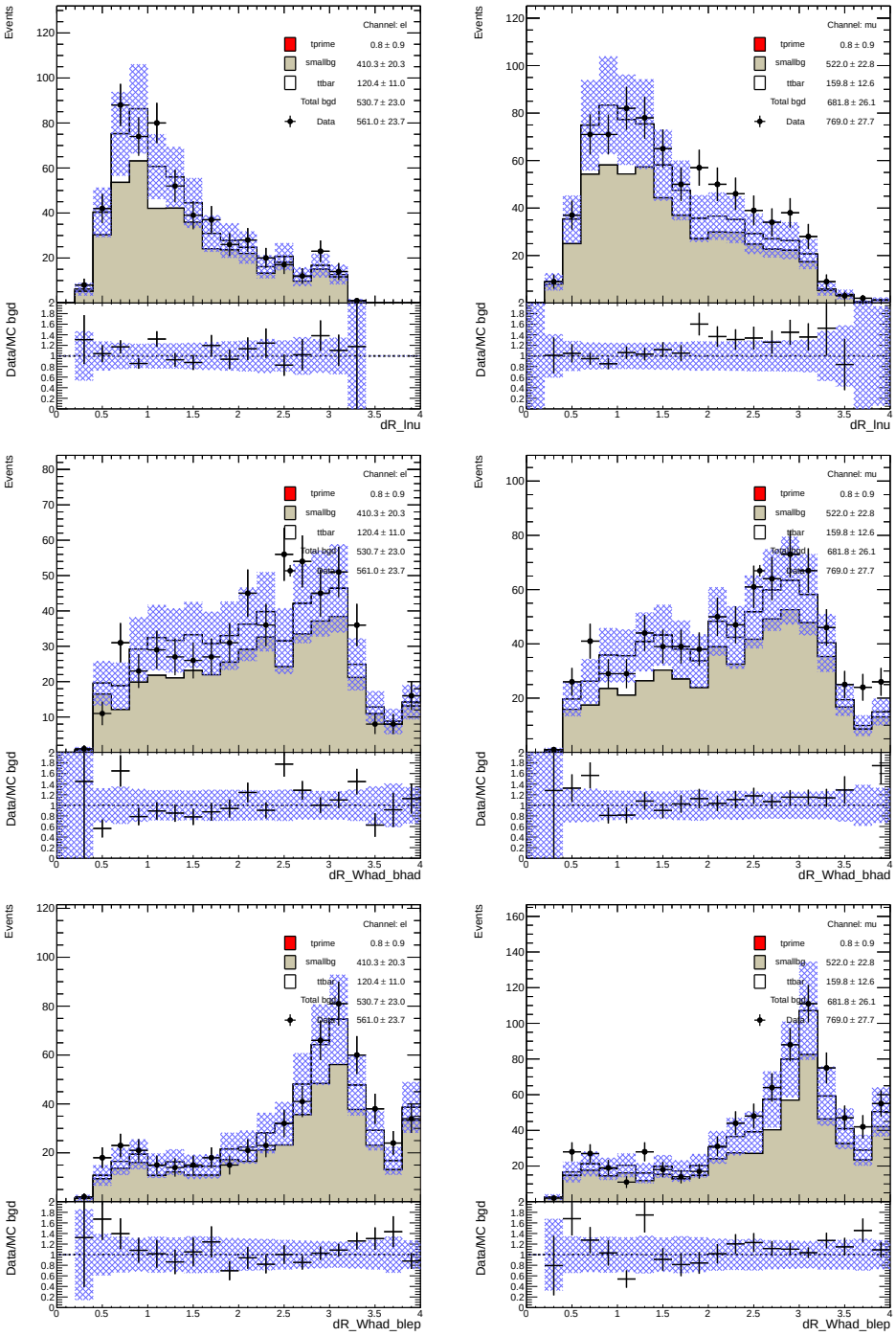


Figure C.19: Data/MC comparison of the NN input variables in the Control Region W_{had}-Selection, $= 0\,b$-tags for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{l\nu}$ (top), $\Delta R^{W_{\mathrm{had}}b_{\mathrm{had}}}$ (middle), $\Delta R^{W_{\mathrm{had}}b_{\mathrm{lep}}}$ (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

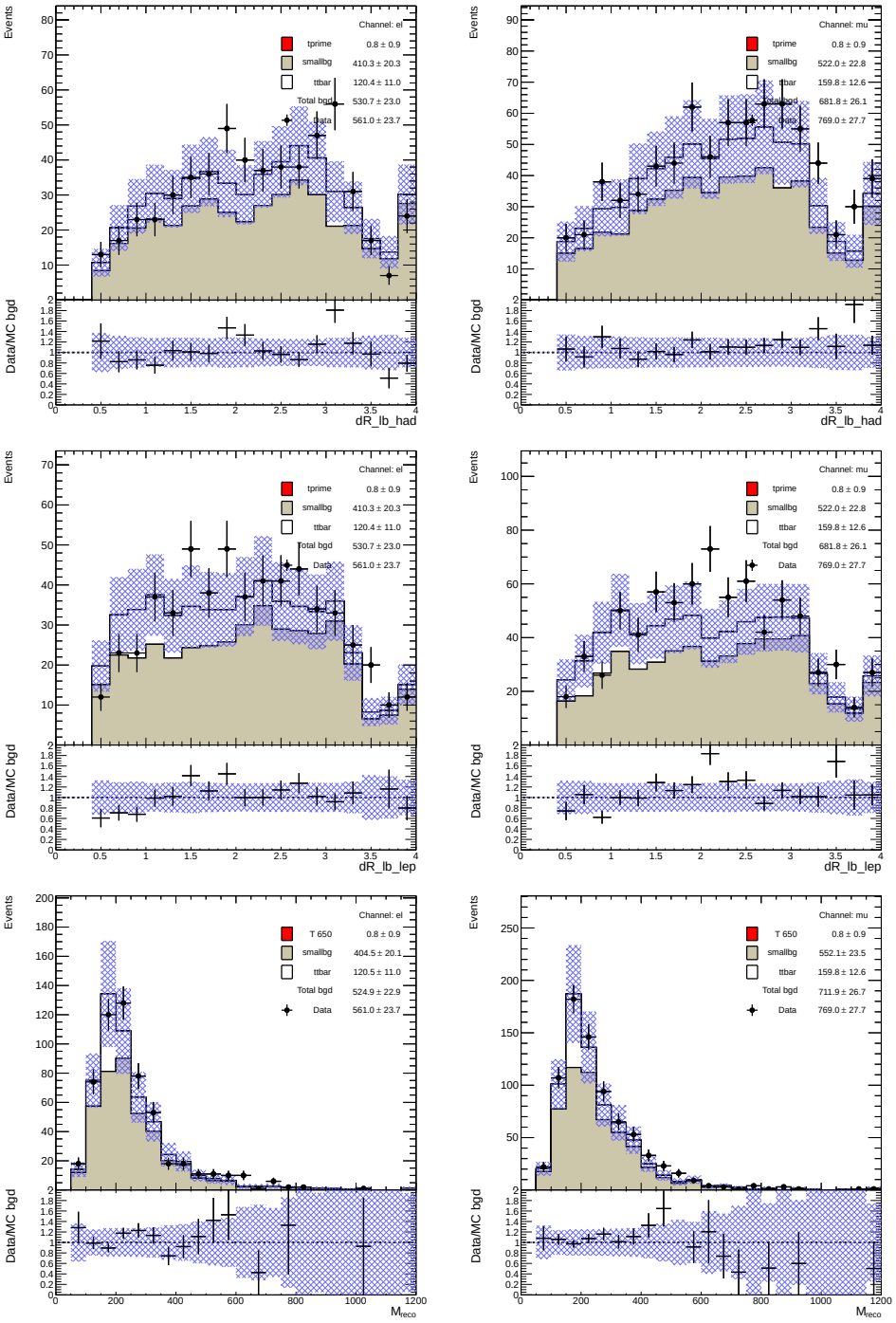


Figure C.20: Data/MC comparison of the NN input variables in the Control Region W_{had}-Selection, $= 0\,b$-tags for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{lb_{\text{had}}}$ (top), $\Delta R^{lb_{\text{lep}}}$ (middle), M_{reco} (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

C.4.6 NN Input Variables, Control Region W_{had}-Selection, $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $M_{\mathrm{reco}} < 350$ GeV

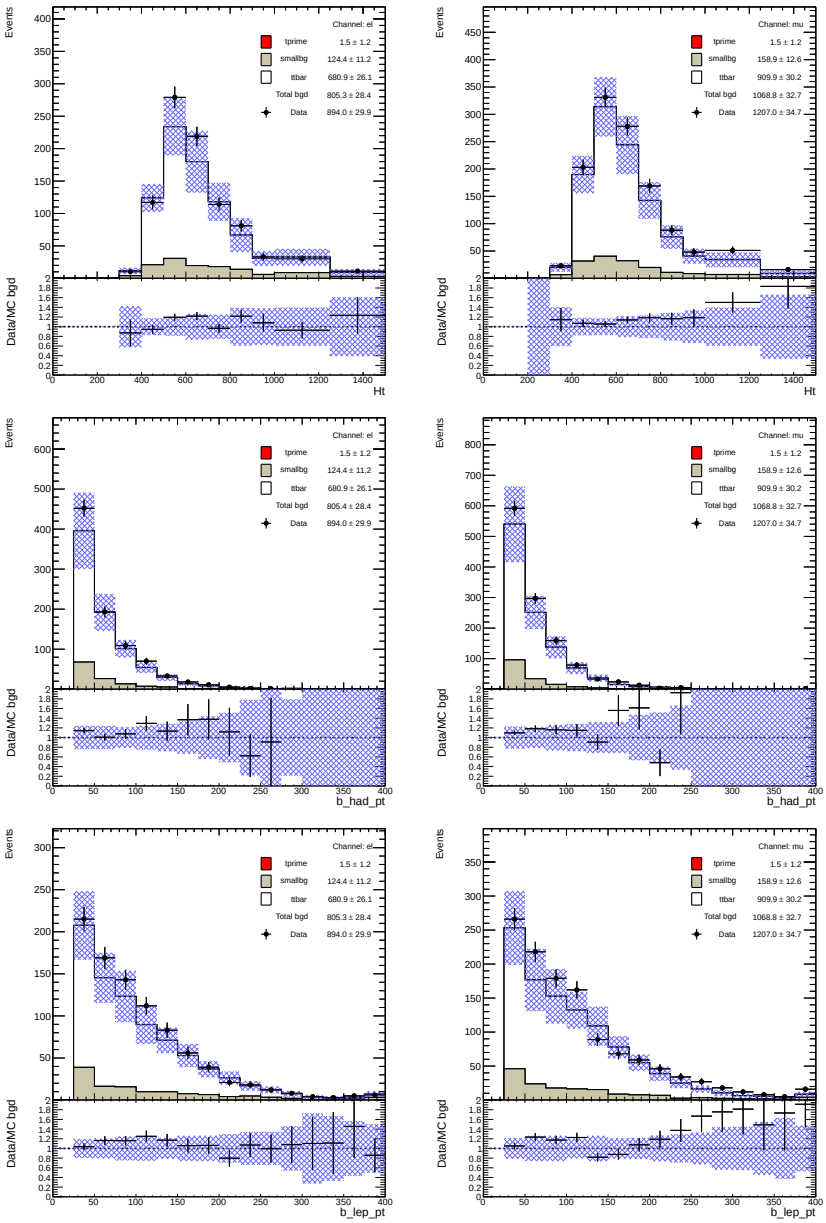


Figure C.21: Data/MC comparison of the NN input variables in the Control Region W_{had}-Selection, $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $M_{\mathrm{reco}} < 350$ GeV for the e-Channel (left) and μ-Channel (right) separately: total transverse momentum H_{T} (top), p_{T} of the leptonic b (middle), p_{T} of the hadronically decaying b (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

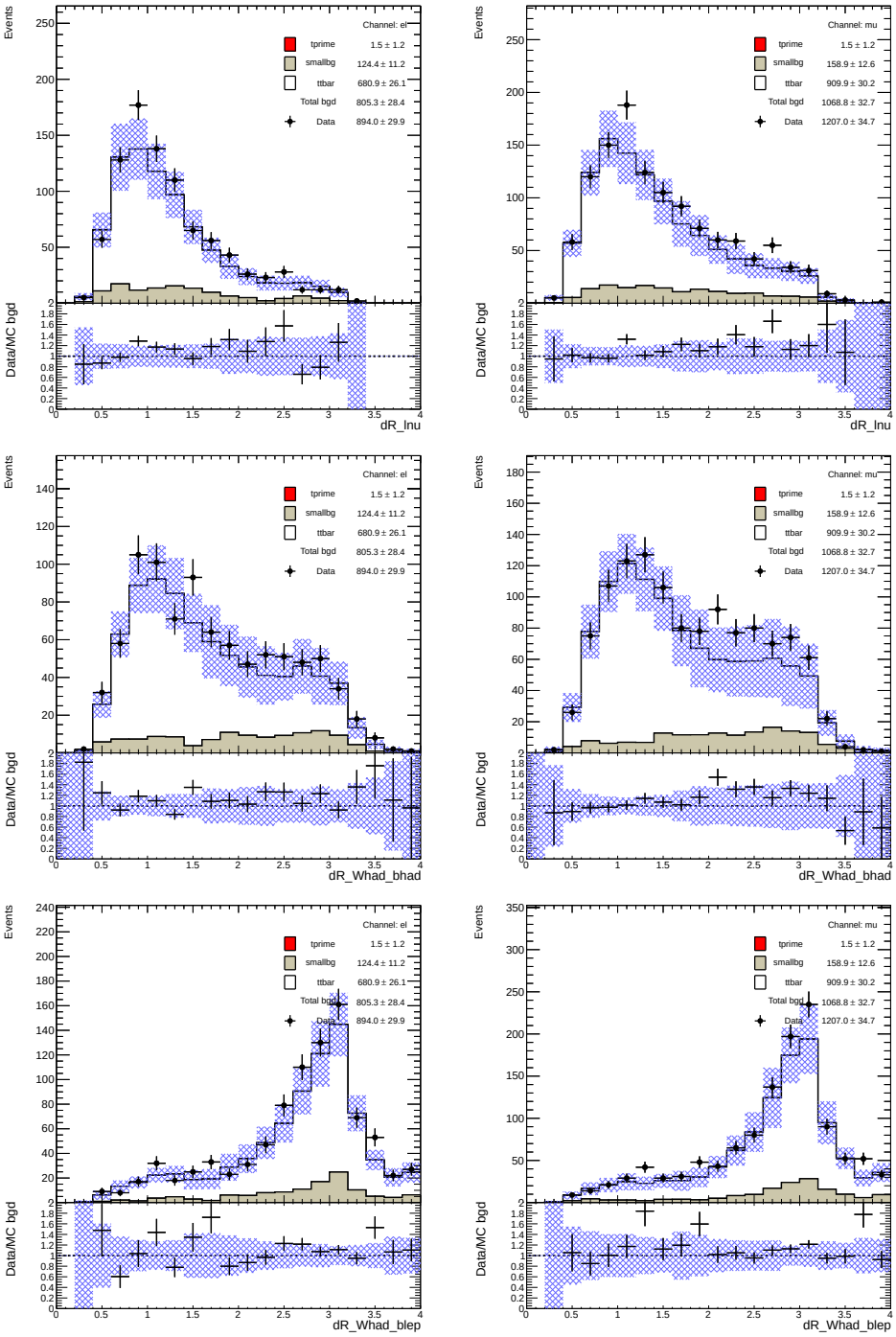


Figure C.22: Data/MC comparison of the NN input variables in the Control Region W_{had}-Selection, $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $M_{\mathrm{reco}} < 350$ GeV for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{l\nu}$ (top), $\Delta R^{W_{\mathrm{had}} b_{\mathrm{had}}}$ (middle), $\Delta R^{W_{\mathrm{had}} b_{\mathrm{lep}}}$ (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

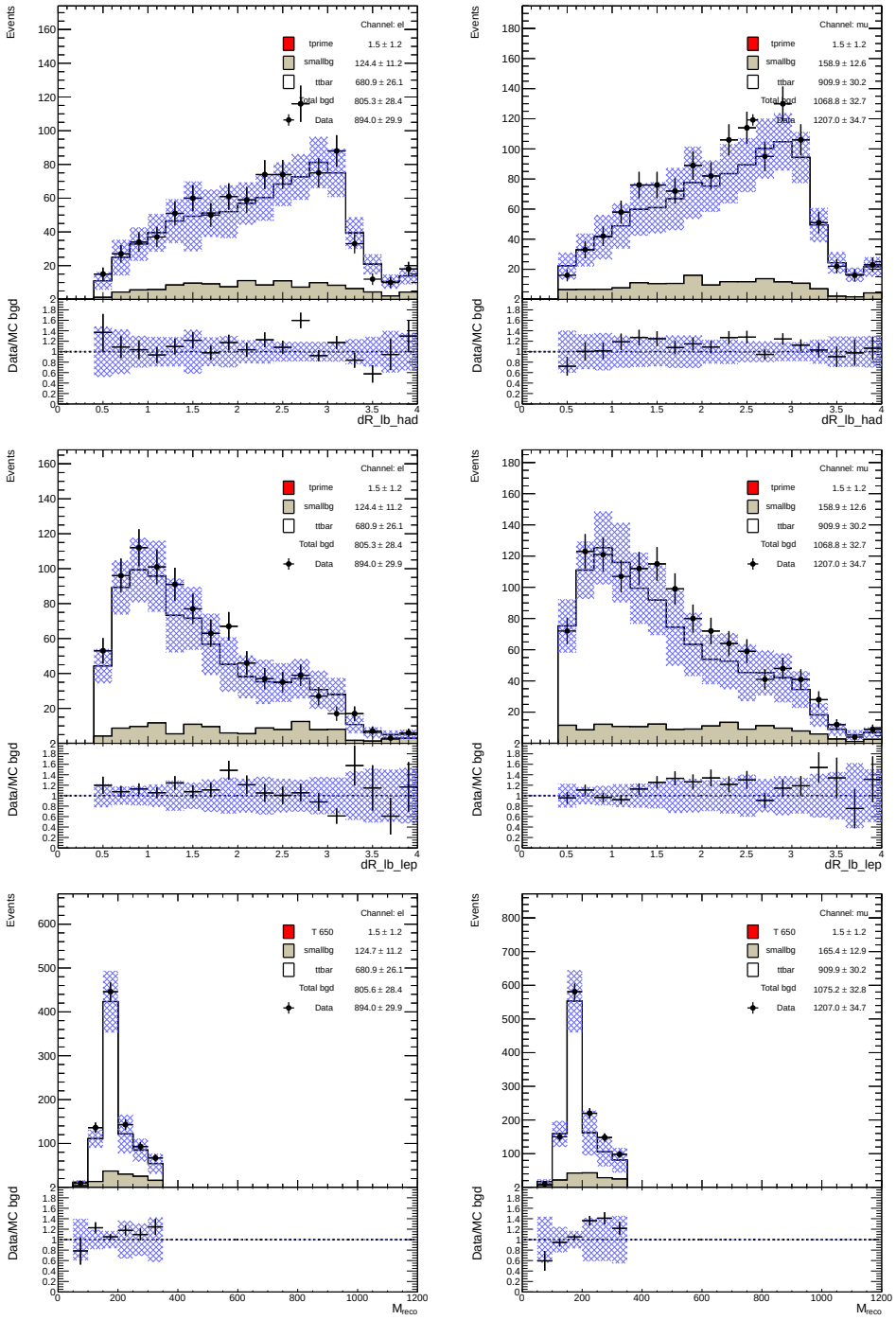


Figure C.23: Data/MC comparison of the NN input variables in the Control Region W_{had}-Selection, $\geqslant 4$ jets, $\geqslant 1$ b-tags, $M_{\text{reco}} < 350$ GeV for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{lb_{\text{had}}}$ (top), $\Delta R^{lb_{\text{lep}}}$ (middle), M_{reco} (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

C.4.7 NN Input Variables, Control Region: W_{had}-Selection, $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $H_{\text{T}} < 700$ GeV

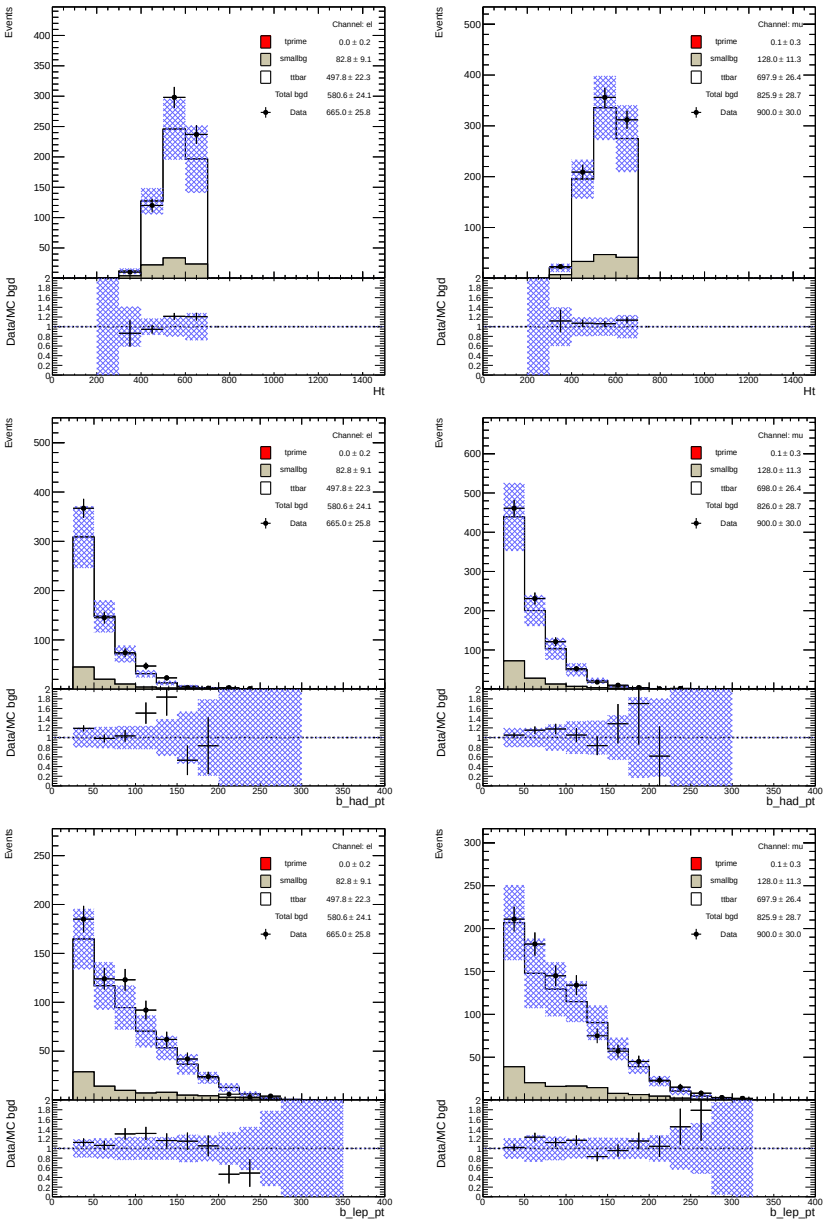


Figure C.24: Data/MC comparison of the NN input variables in the Control Region W_{had}-Selection, $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $H_{\text{T}} < 700$ GeV for the e-Channel (left) and μ-Channel (right) separately: total transverse momentum H_{T} (top), p_{T} of the leptonic b (middle), p_{T} of the hadronically decaying b (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

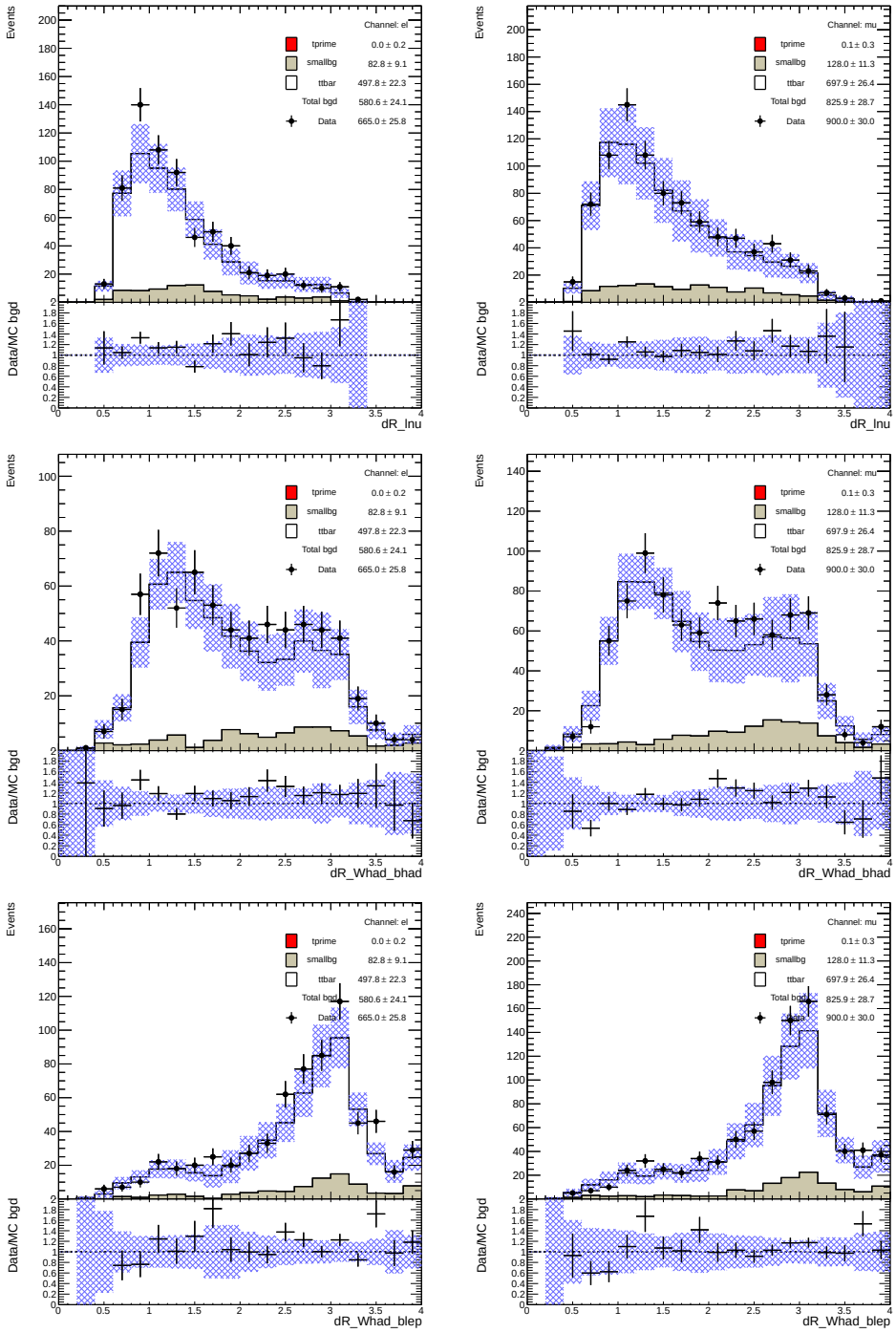


Figure C.25: Data/MC comparison of the NN input variables in the Control Region W_{had}-Selection, $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $H_{\mathrm{T}} < 700$ GeV for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{l\nu}$ (top), $\Delta R^{W_{\mathrm{had}}b_{\mathrm{had}}}$ (middle), $\Delta R^{W_{\mathrm{had}}b_{\mathrm{lep}}}$ (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

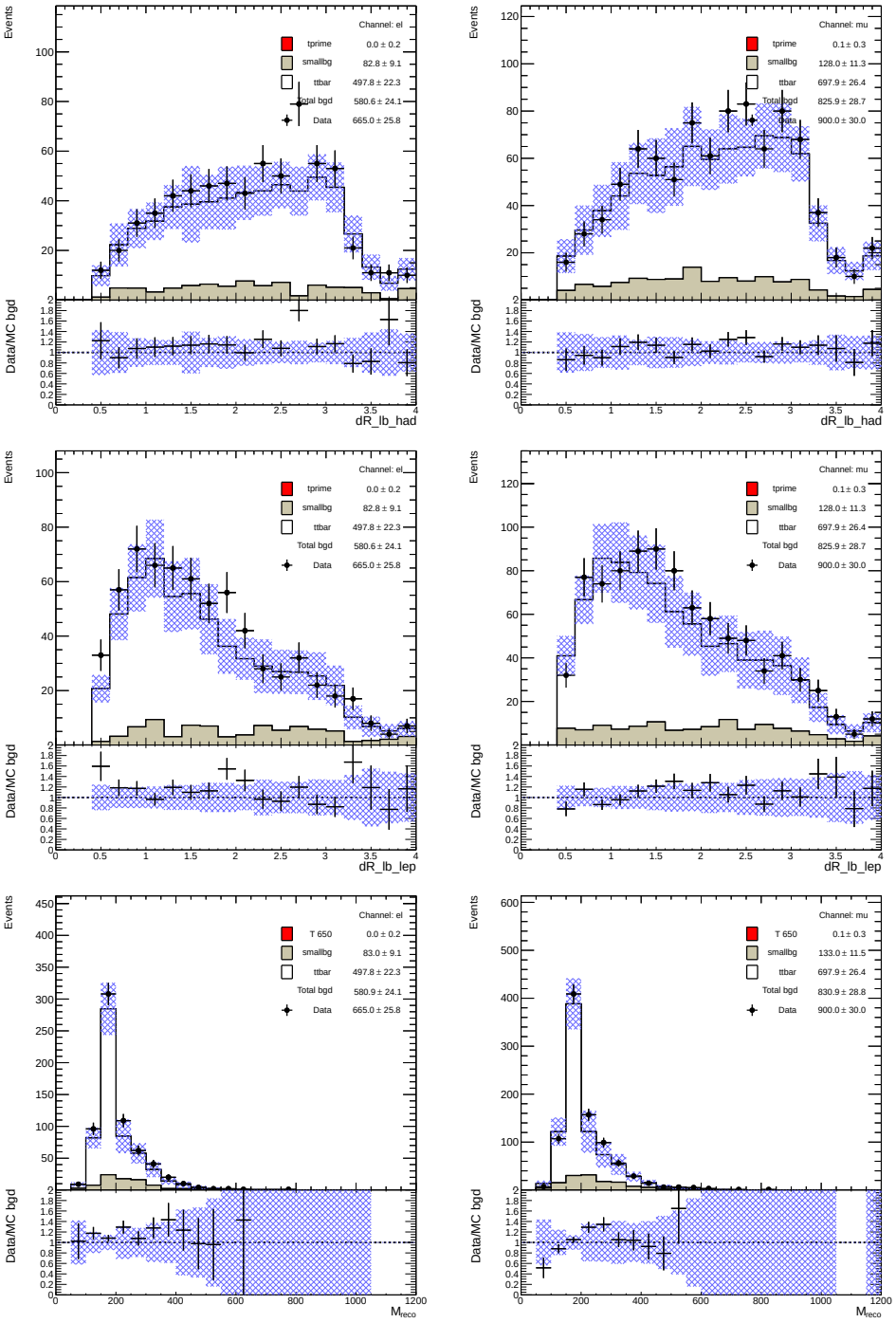


Figure C.26: Data/MC comparison of the NN input variables in the Control Region W_{had}-Selection, $\geqslant 4$ jets, $\geqslant 1\,b$-tags, $H_{\text{T}} < 700$ GeV for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{lb_{\text{had}}}$ (top), $\Delta R^{lb_{\text{lep}}}$ (middle), M_{reco} (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

C.4.8 NN Input Variables, Control Region: W_{had}-Selection, $\geqslant 4$ jets, $\geqslant 1\,b$-tags, NN<0.8

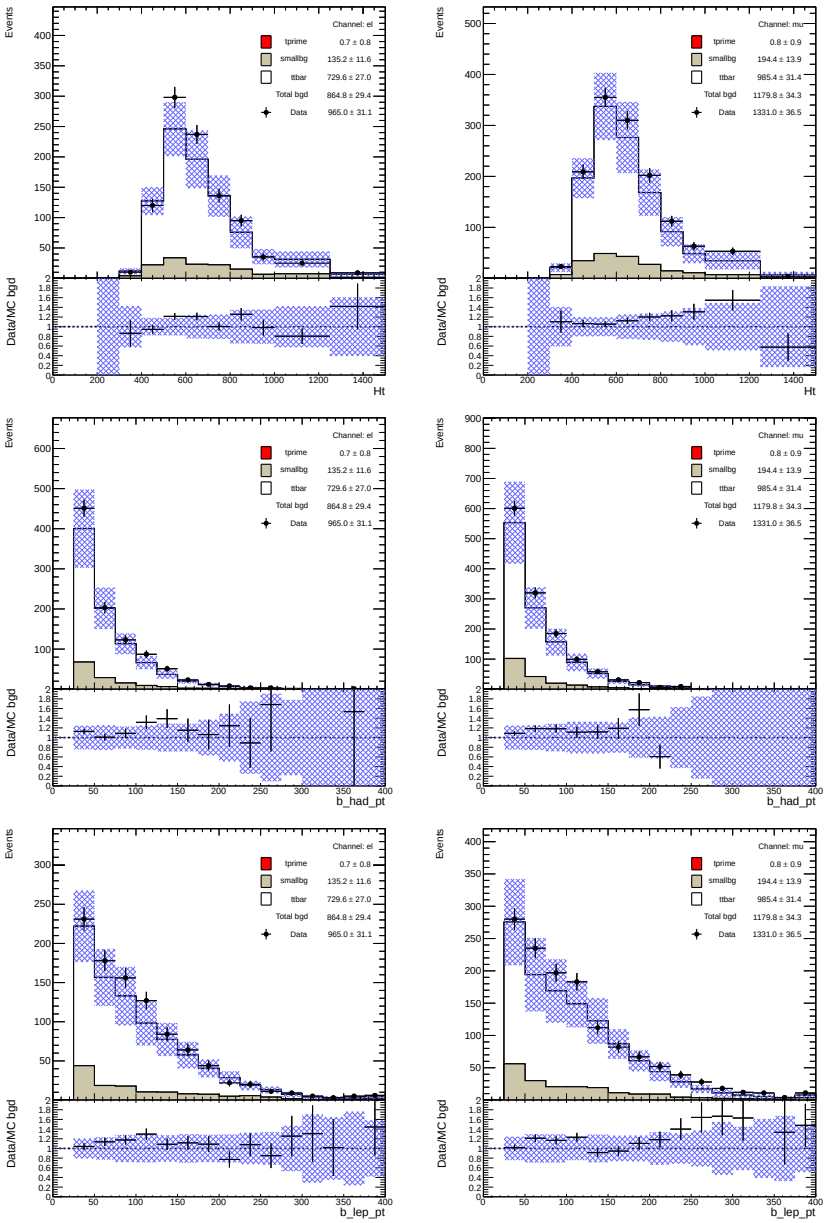


Figure C.27: Data/MC comparison of the NN input variables in the Control Region W_{had}-Selection, $\geqslant 4$ jets, $\geqslant 1\,b$-tags, NN<0.8 for the e-Channel (left) and μ-Channel (right) separately: total transverse momentum H_{T} (top), p_{T} of the leptonic b (middle), p_{T} of the hadronically decaying b (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

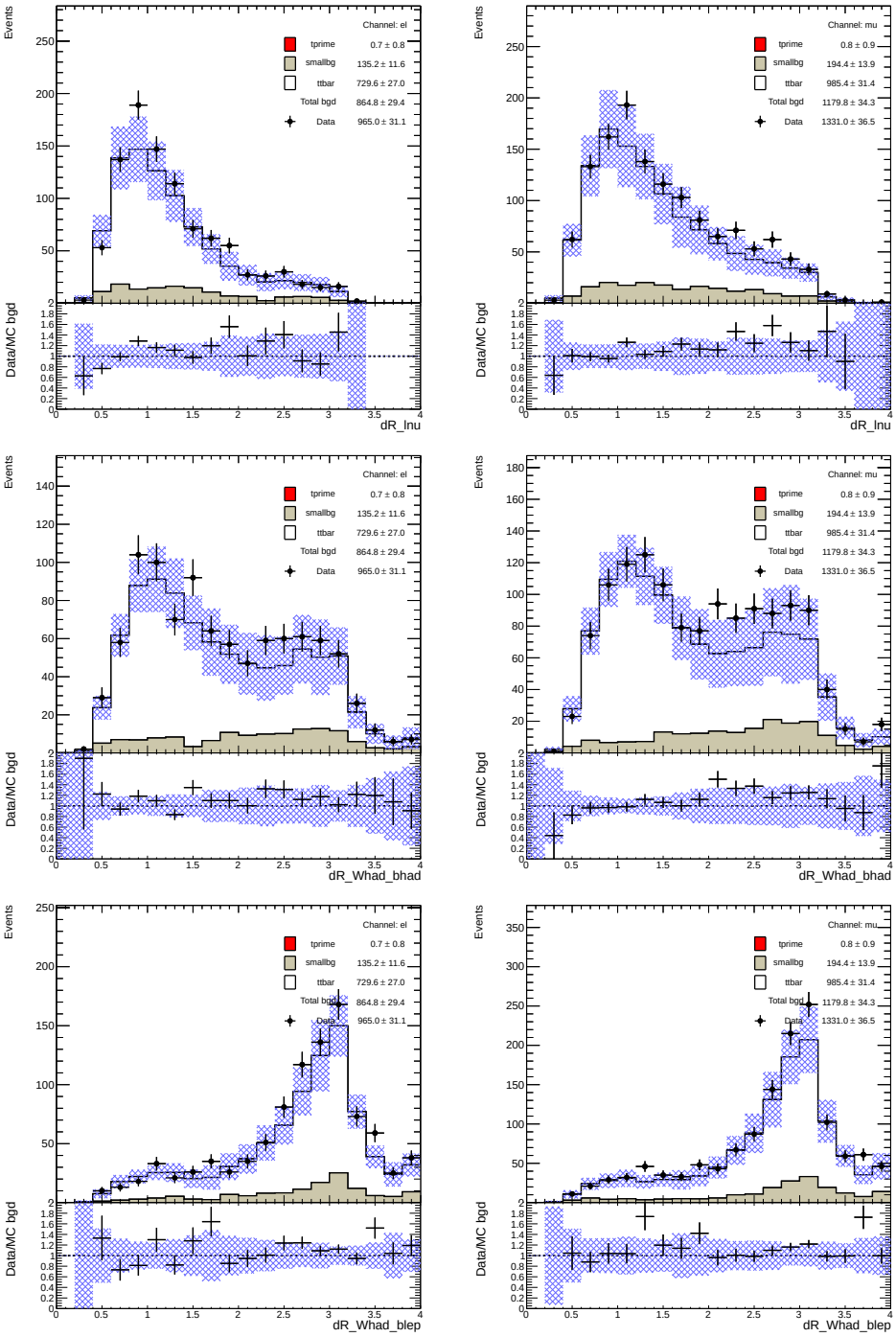


Figure C.28: Data/MC comparison of the NN input variables in the Control Region W_{had}-Selection $\geqslant 4$ jets, $\geqslant 1$ b-tags, NN<0.8 for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{l\nu}$ (top), $\Delta R^{W_{\mathrm{had}}b_{\mathrm{had}}}$ (middle), $\Delta R^{W_{\mathrm{had}}b_{\mathrm{lep}}}$ (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

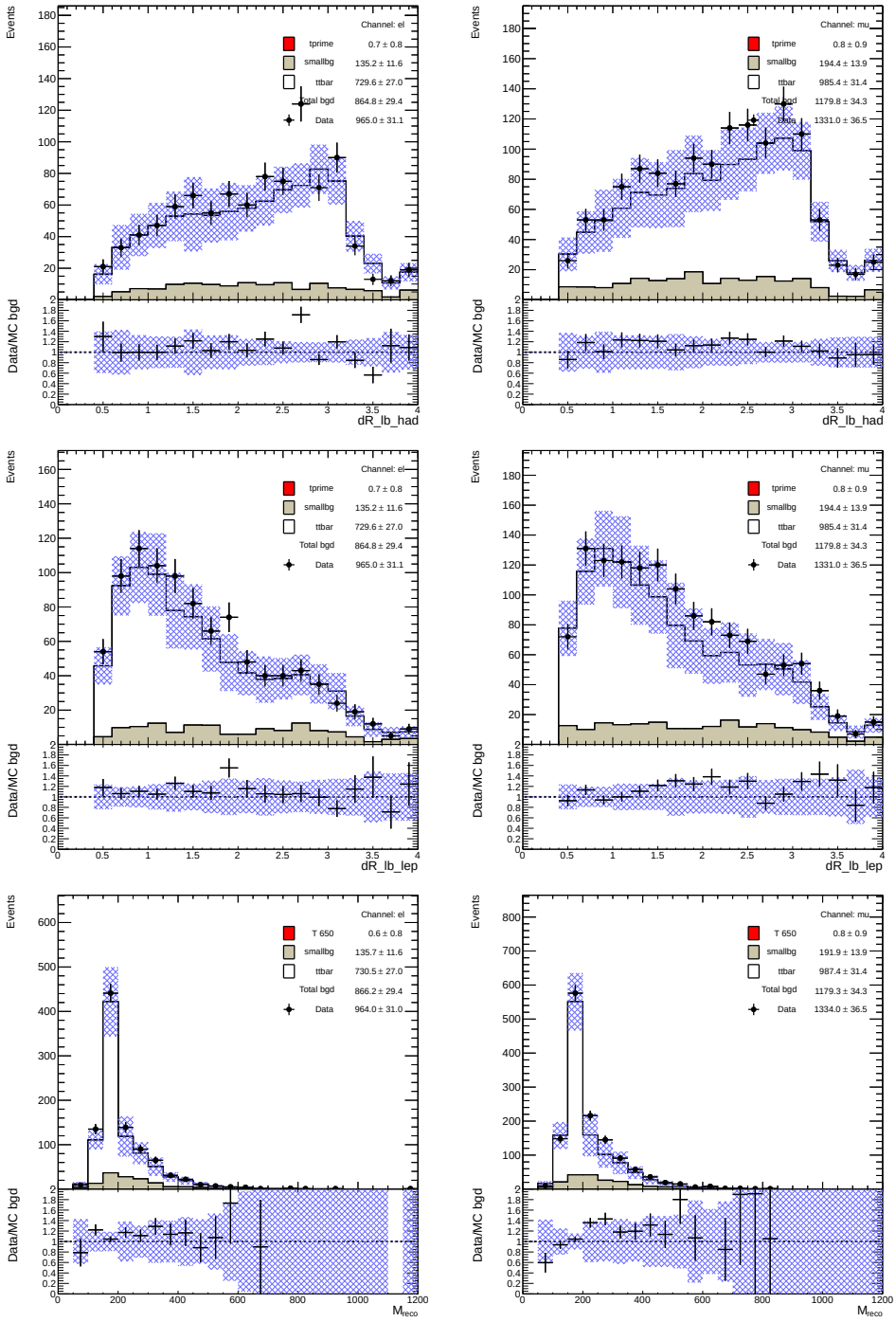


Figure C.29: Data/MC comparison of the NN input variables in the Control Region W_{had}-Selection, $\geqslant 4$ jets, $\geqslant 1\,b$-tags, NN<0.8 for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{lb_{\text{had}}}$ (top), $\Delta R^{lb_{\text{lep}}}$ (middle), M_{reco} (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

C.4.9 NN Input Variables, W_{had}-Selection, full region

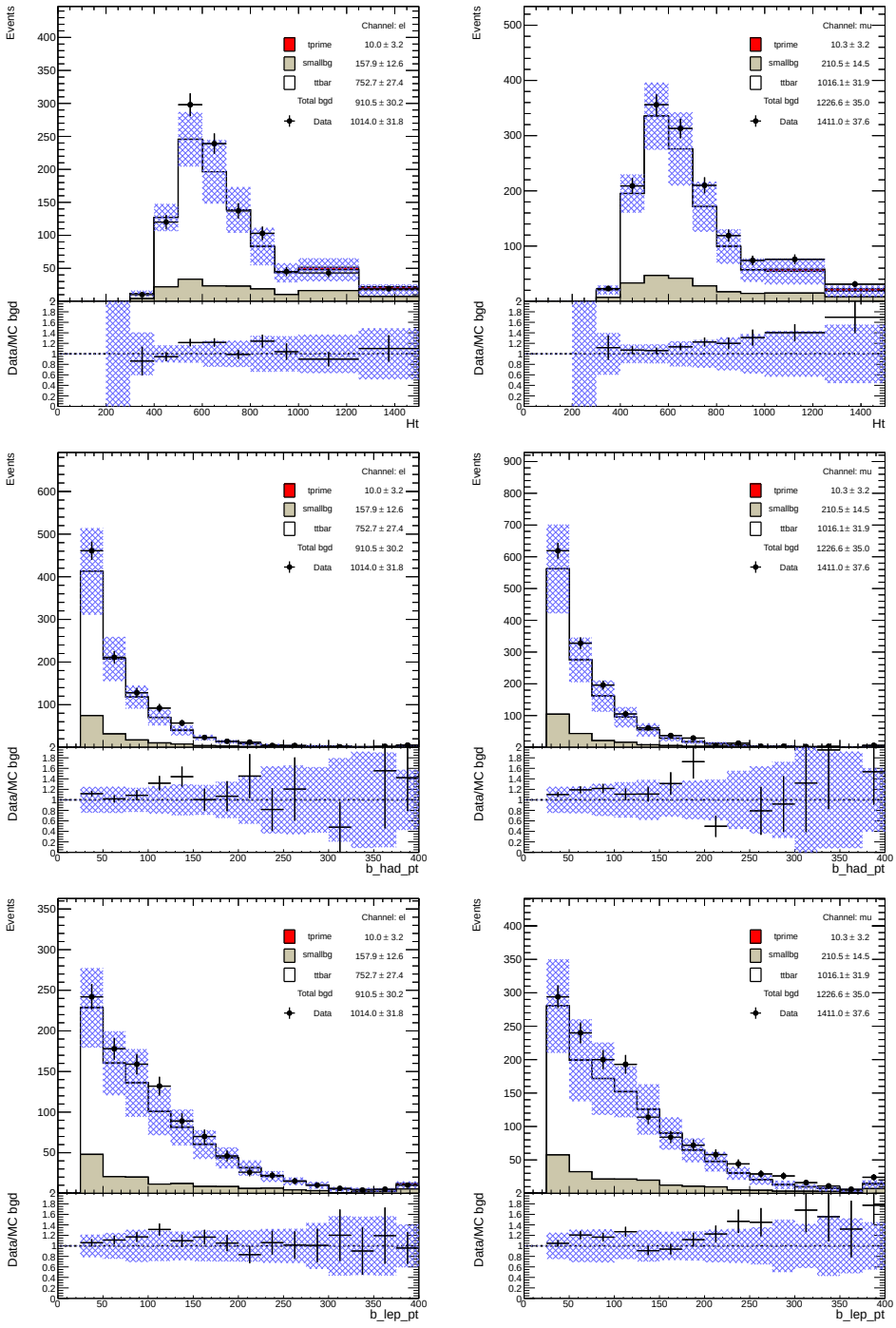


Figure C.30: Data/MC comparison of the NN input variables after W_{had}-Selection for the e-Channel (left) and μ-Channel (right) separately: total transverse momentum H_{T} (top), p_{T} of the leptonic b (middle), p_{T} of the hadronically decaying b (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

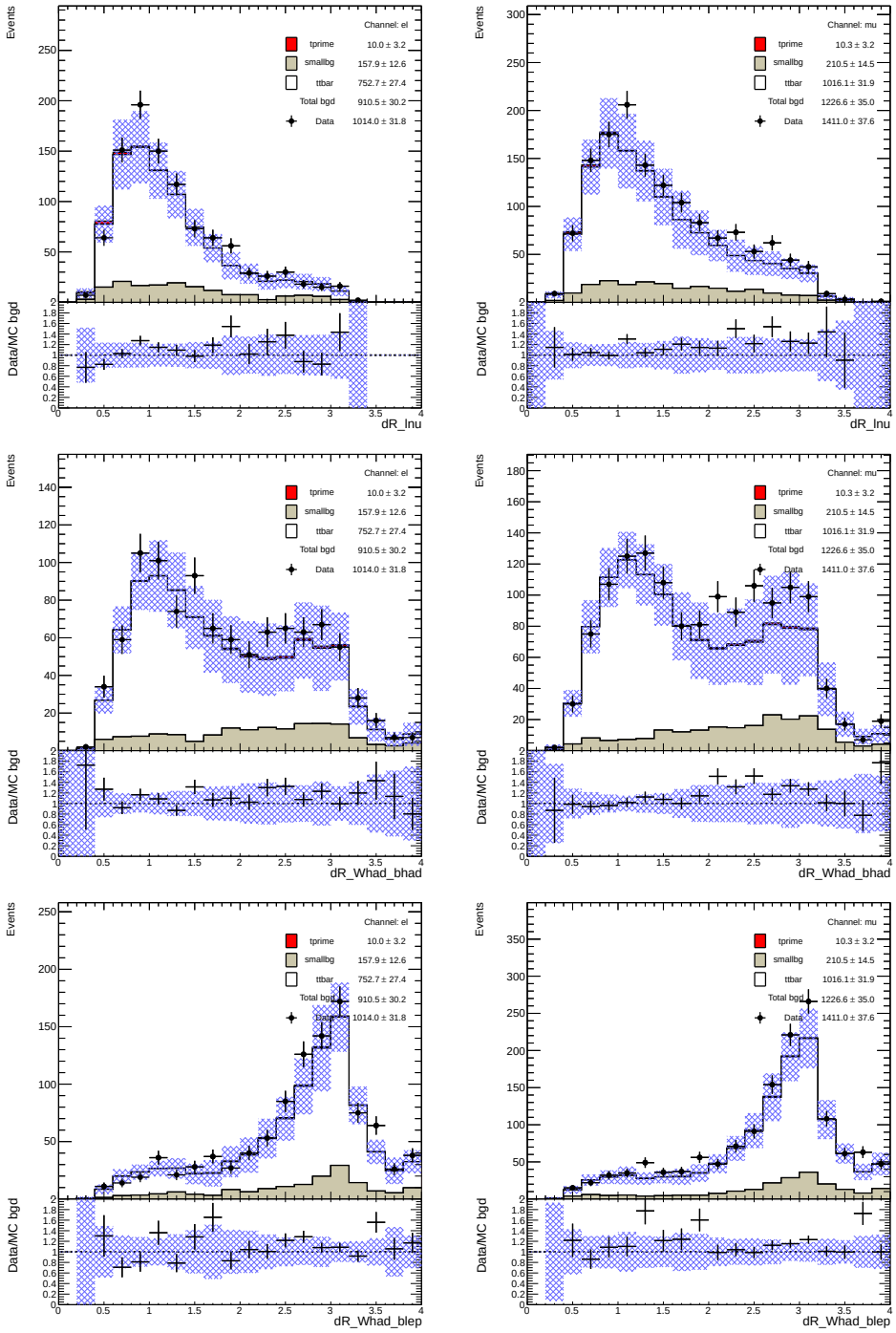


Figure C.31: Data/MC comparison of the NN input variables after W_{had}-Selection for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{l\nu}$ (top), $\Delta R^{W_{\mathrm{had}}b_{\mathrm{had}}}$ (middle), $\Delta R^{W_{\mathrm{had}}b_{\mathrm{lep}}}$ (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

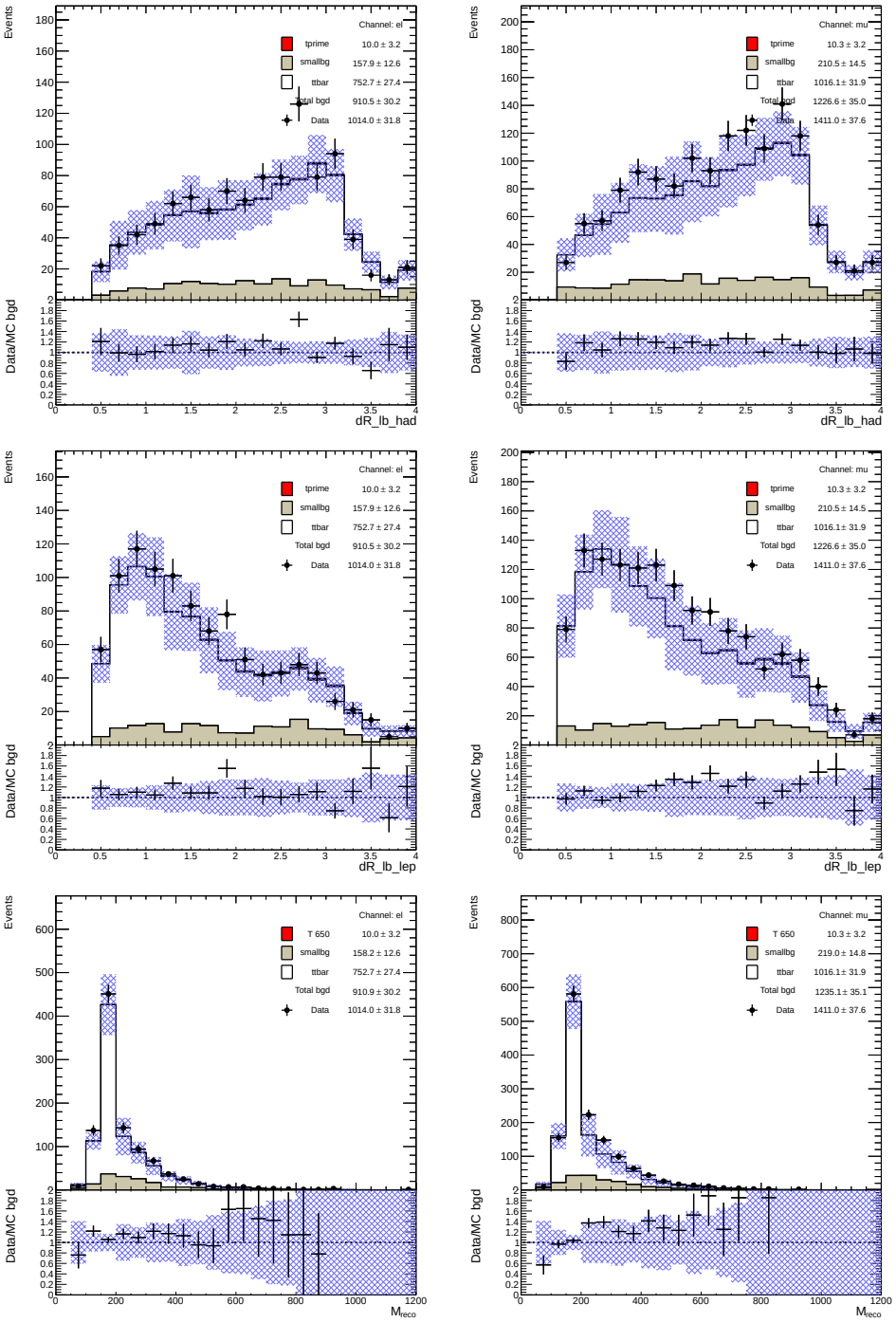


Figure C.32: Data/MC comparison of the NN input variables after W_{had}-Selection for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{lb_{\text{had}}}$ (top), $\Delta R^{lb_{\text{lep}}}$ (middle), M_{reco} (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

C.4.10 NN Input Variables, TIGHT Selection

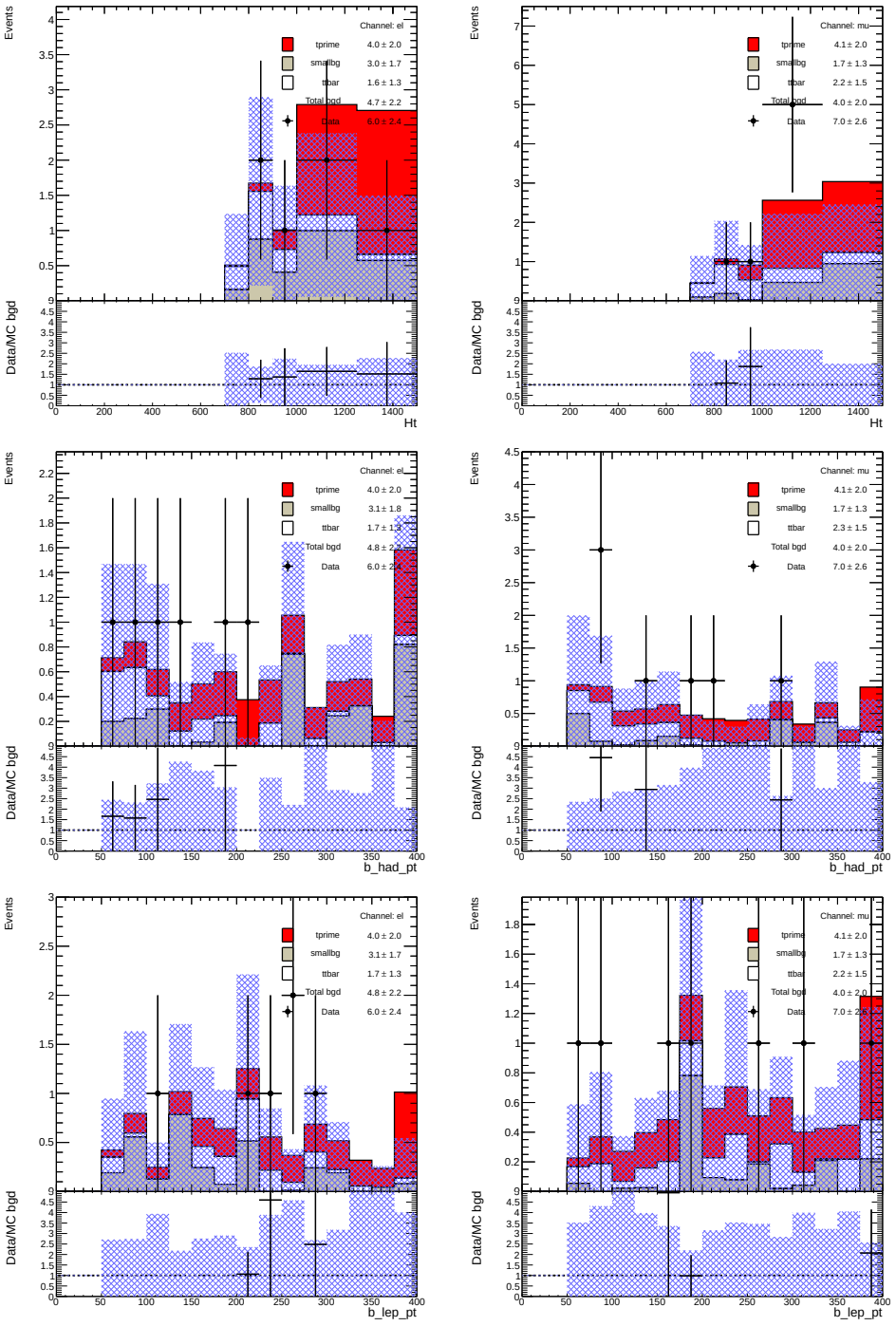


Figure C.33: Data/MC comparison of the NN input variables after the TIGHT Selection for the e-Channel (left) and μ-Channel (right) separately: total transverse momentum H_{T} (top), p_{T} of the leptonic b (middle), p_{T} of the hadronically decaying b (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

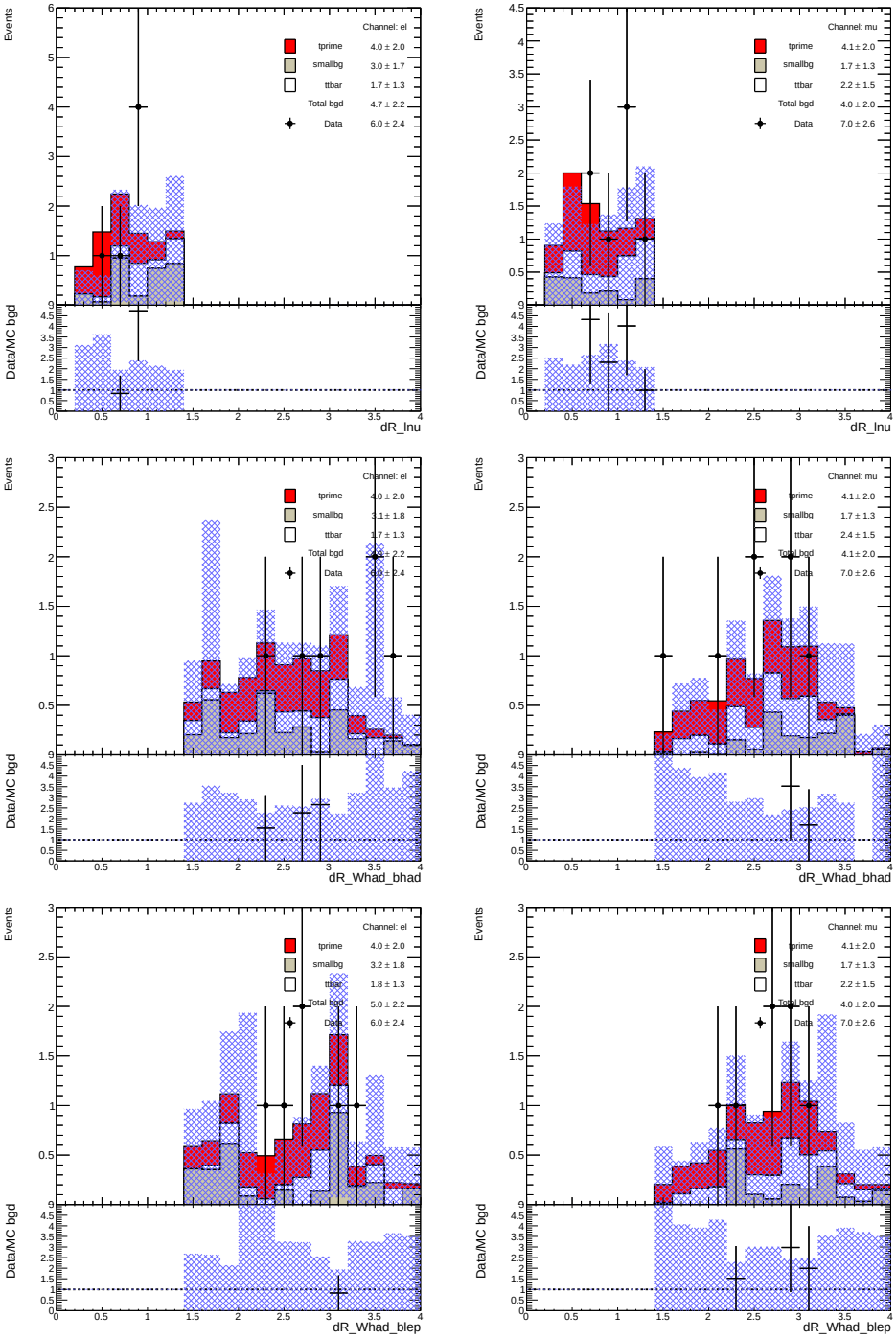


Figure C.34: Data/MC comparison of the NN input variables after the TIGHT Selection for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{l\nu}$ (top), $\Delta R^{W_{\mathrm{had}}b_{\mathrm{had}}}$ (middle), $\Delta R^{W_{\mathrm{had}}b_{\mathrm{lep}}}$ (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

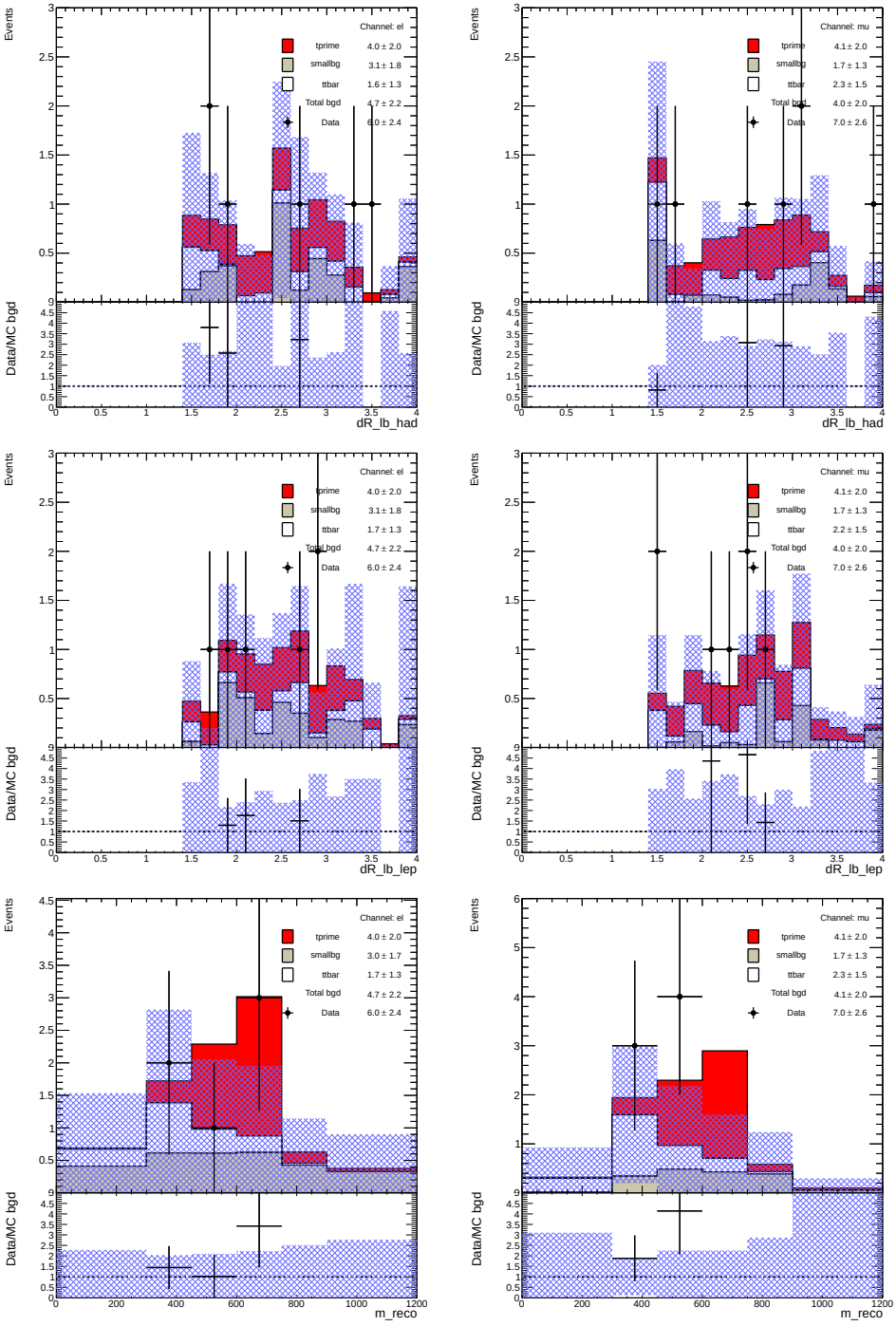


Figure C.35: Data/MC comparison of the NN input variables after the TIGHT Selection for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{lb_{\text{had}}}$ (top), $\Delta R^{lb_{\text{lep}}}$ (middle), M_{reco} (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

C.4.11 NN Input Variables, NN>0.9965

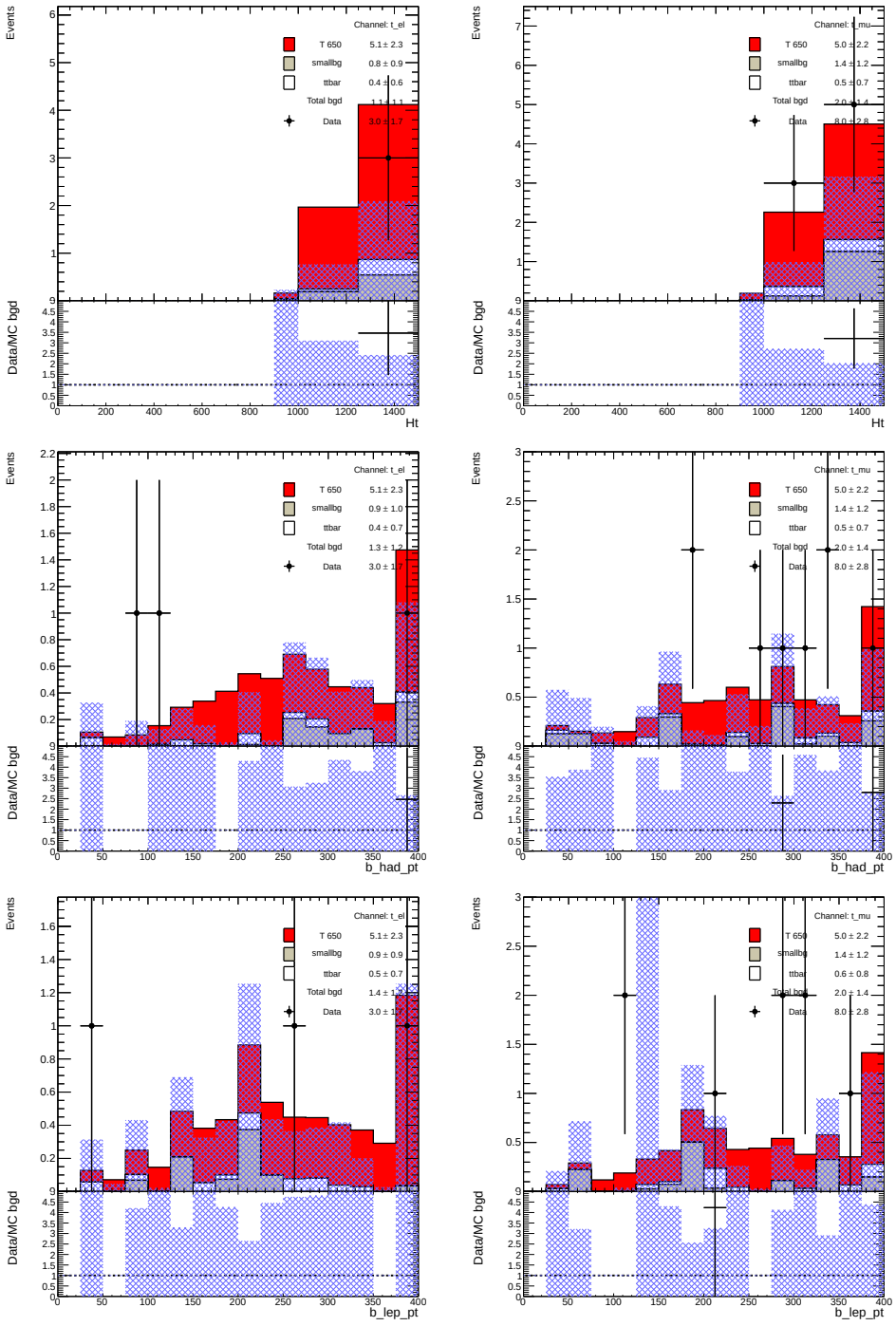


Figure C.36: Data/MC comparison of the NN input variables after the NN Selection for the e-Channel (left) and μ-Channel (right) separately: total transverse momentum H_{T} (top), p_{T} of the leptonic b (middle), p_{T} of the hadronically decaying b (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

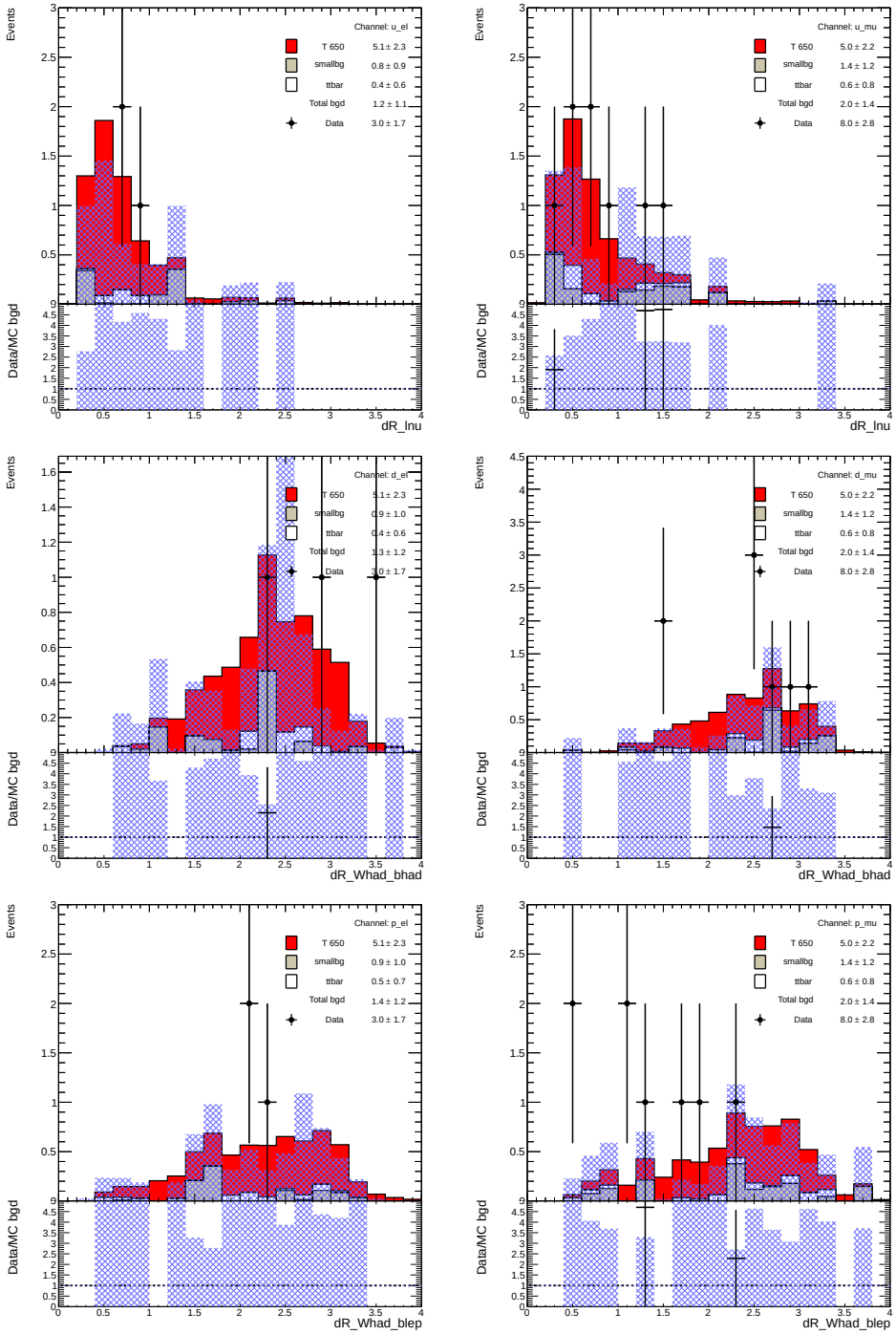


Figure C.37: Data/MC comparison of the NN input variables after the NN Selection for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{l\nu}$ (top), $\Delta R^{W_{\mathrm{had}} b_{\mathrm{had}}}$ (middle), $\Delta R^{W_{\mathrm{had}} b_{\mathrm{lep}}}$ (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

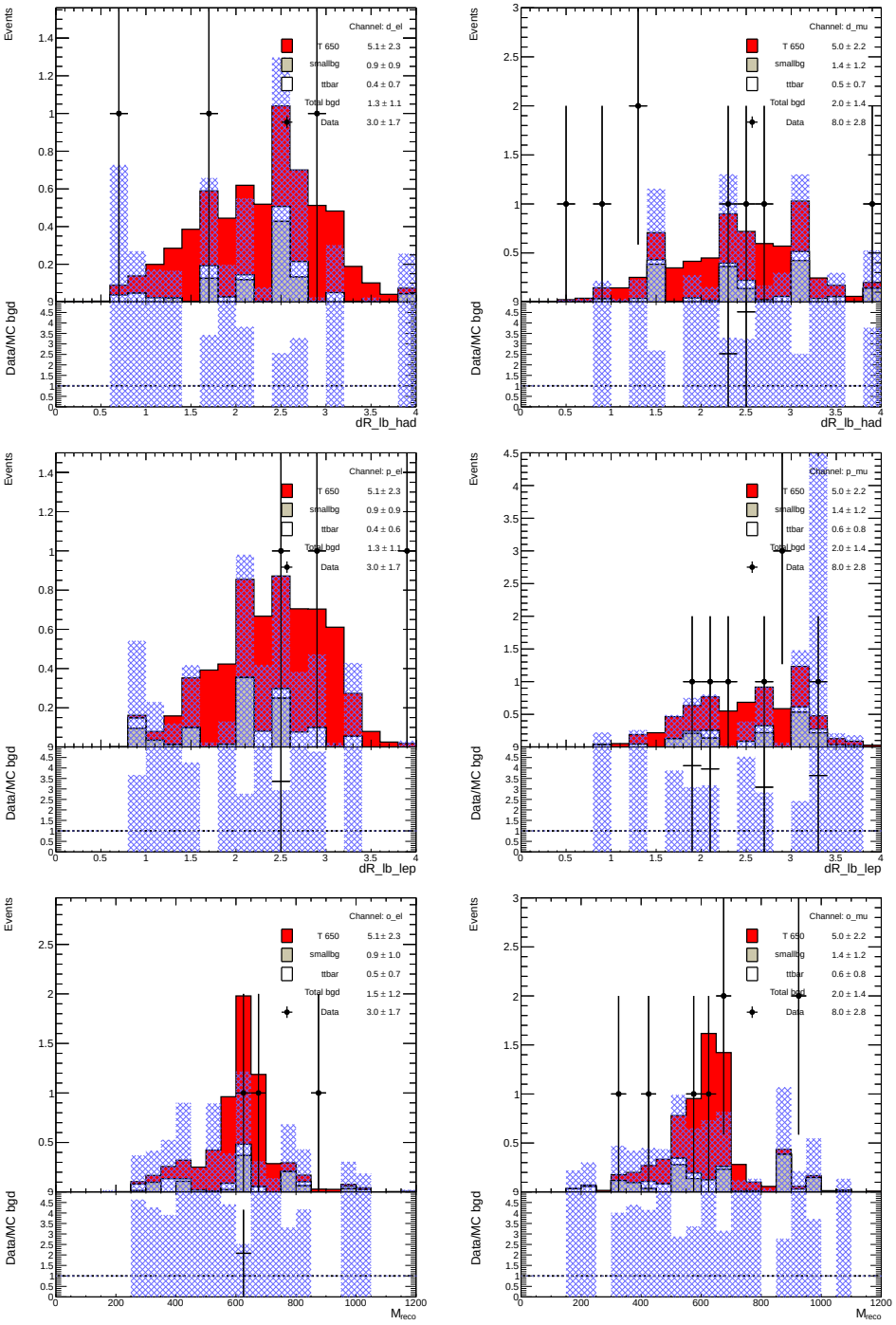


Figure C.38: Data/MC comparison of the NN input variables after the NN Selection for the e-Channel (left) and μ-Channel (right) separately: $\Delta R^{lb_{\mathrm{had}}}$ (top), $\Delta R^{lb_{\mathrm{lep}}}$ (middle), M_{reco} (bottom). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

C.5 Investigation of the Muon Channel Excess

The muon channel exhibits an excess of data in the high tail of the H_{T} variable after the W_{had}-selection, as shown in the plot in the top left corner of Figure C.39. The plots following six plots show that this excess is dominated by the jet contribution. The plot in the middle of the bottom raw shows that this excess originates form the events with a high jet multiplicity. The plot in the bottom right corner shows that the $\min\Delta R(l,j)$ variable is decently reconstructed, with the excess distributed across the full range of the variable. This excludes the possibility that the excess is due to the events with the fully hadronic final state, in which the muon emerges a b-jet, in which case a peak at the low values of the variable would be expected.

To understand the role of the nominal $t\bar{t}$ sample in the excess in the H_{T} variable, this excess has been investigated using the $t\bar{t}$ samples generated with the POWHEG event generator with the parton shower modelled with either PYTHIA or HERWIG. These two $t\bar{t}$ samples provide fewer expected events, bringing to a higher overall excess, which is, however, differently distributed across the H_{T} spectrum. Instead of raising steadily from the low to the high tail as in case of the nominal $t\bar{t}$ sample, the data to simulation ratio varies from $\geqslant 1$ in the low region and high region to $\leqslant 1$ in the medium region if the POWHEG samples are used, as shown in Figure C.40. The excess is more uniformly distributed across all inspected variables, especially in the distributions of the p_{T} of the jets, as shown in Figure C.41. As both POWHEG samples show similar behaviour, the problematic behaviour of the nominal sample is unlikely to be due to the hadronisation model used (HERWIG). The lack of events in the MC events in the high tails of the transverse momenta of the jets is likely to reflect a problem of the NLO matrix element in MC@NLO.

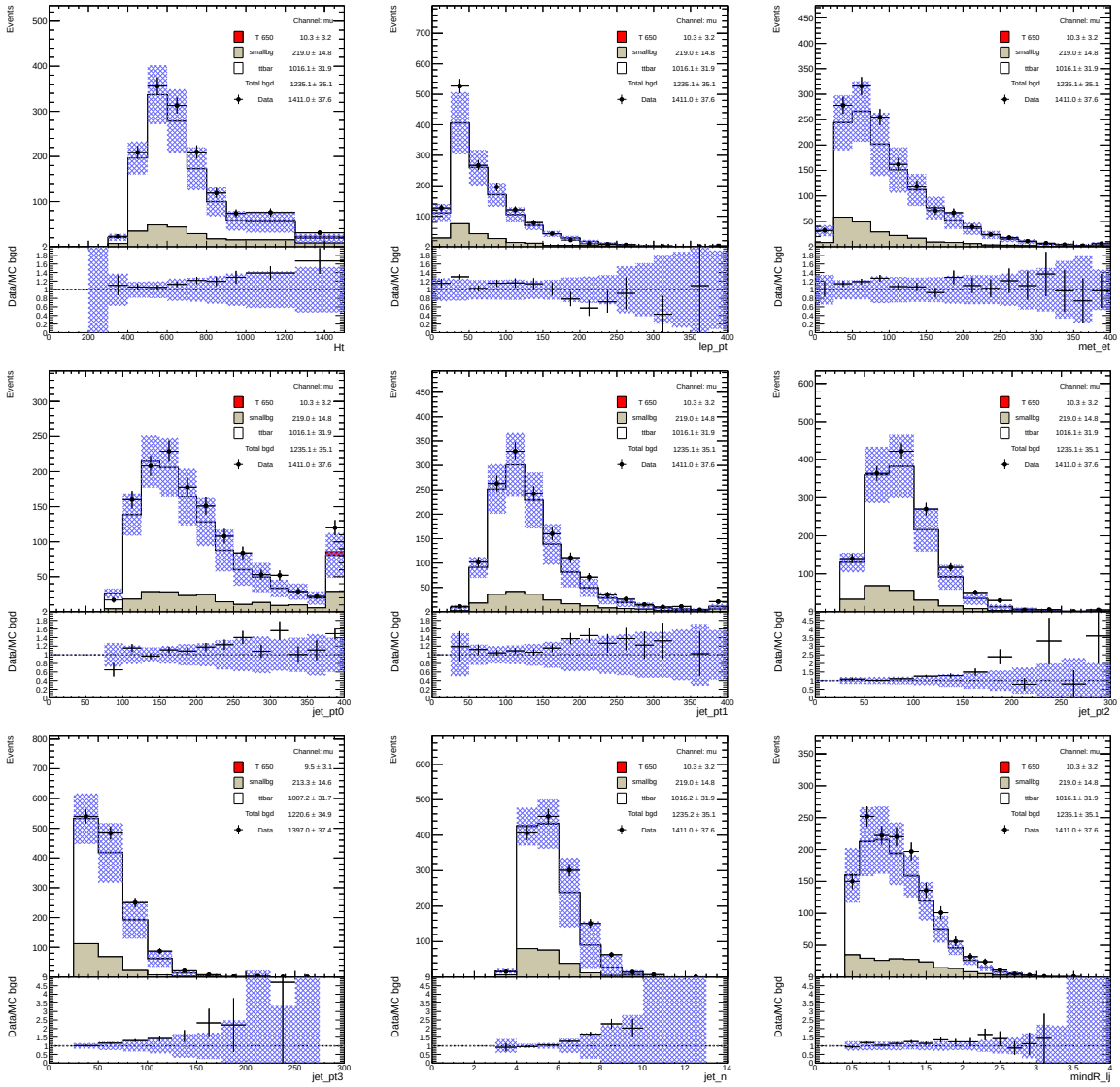


Figure C.39: Data/MC comparison of the H_{T} variable, its components, the jet multiplicity and the minimal space angle between the lepton and a jet in the muon channel after the W_{had}-selection. The blue shaded area represents the 1σ systematic error bands on the expected event yields.

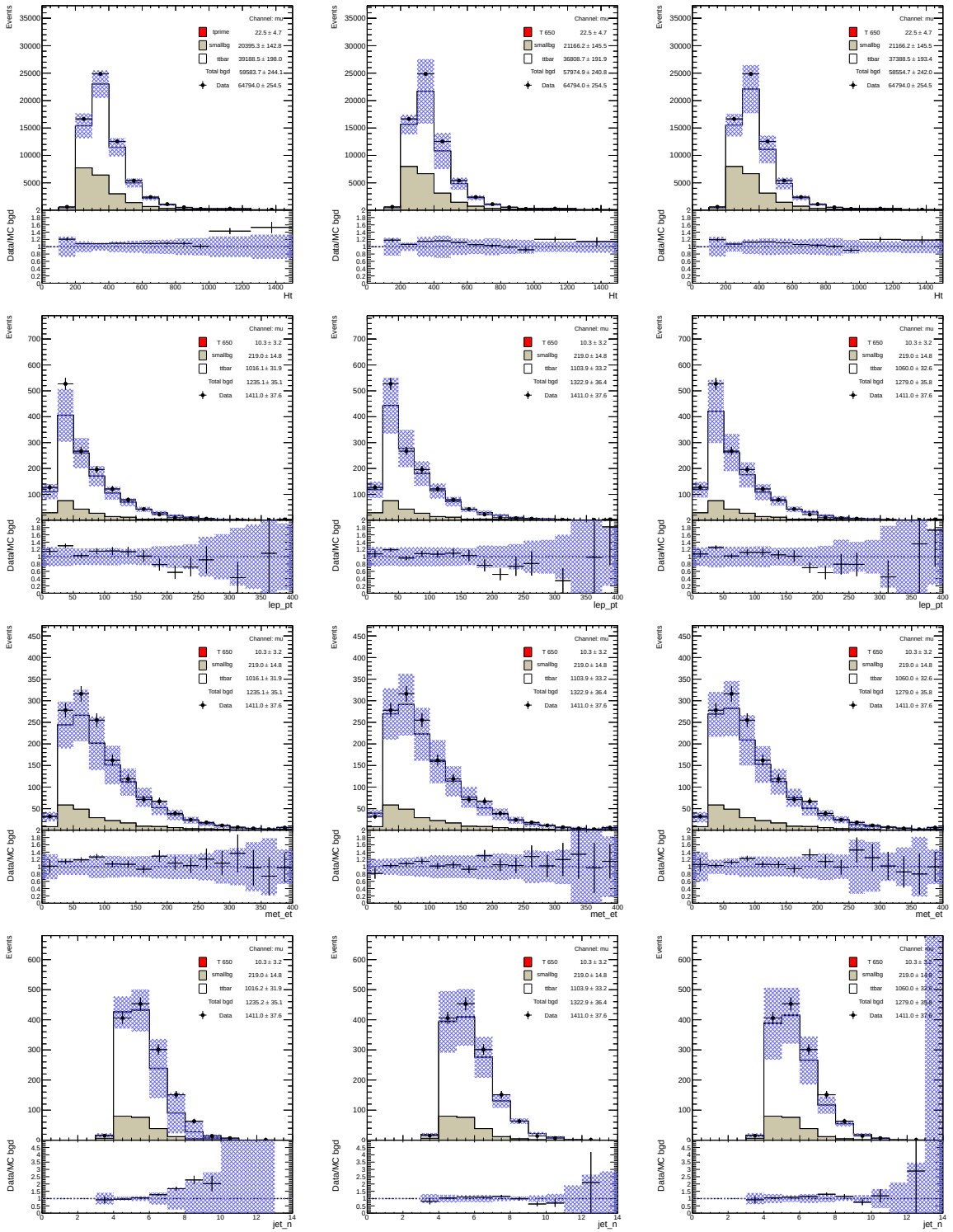


Figure C.40: Data/MC comparison of the H_{T} variable, p_{T} of the lepton, missing transverse energy and jet multiplicity after the W_{had}-selection for three samples generated using different generators: MC@NLO (left column), Powheg-Pythia (middle column), Powheg-Jimmy (right column). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

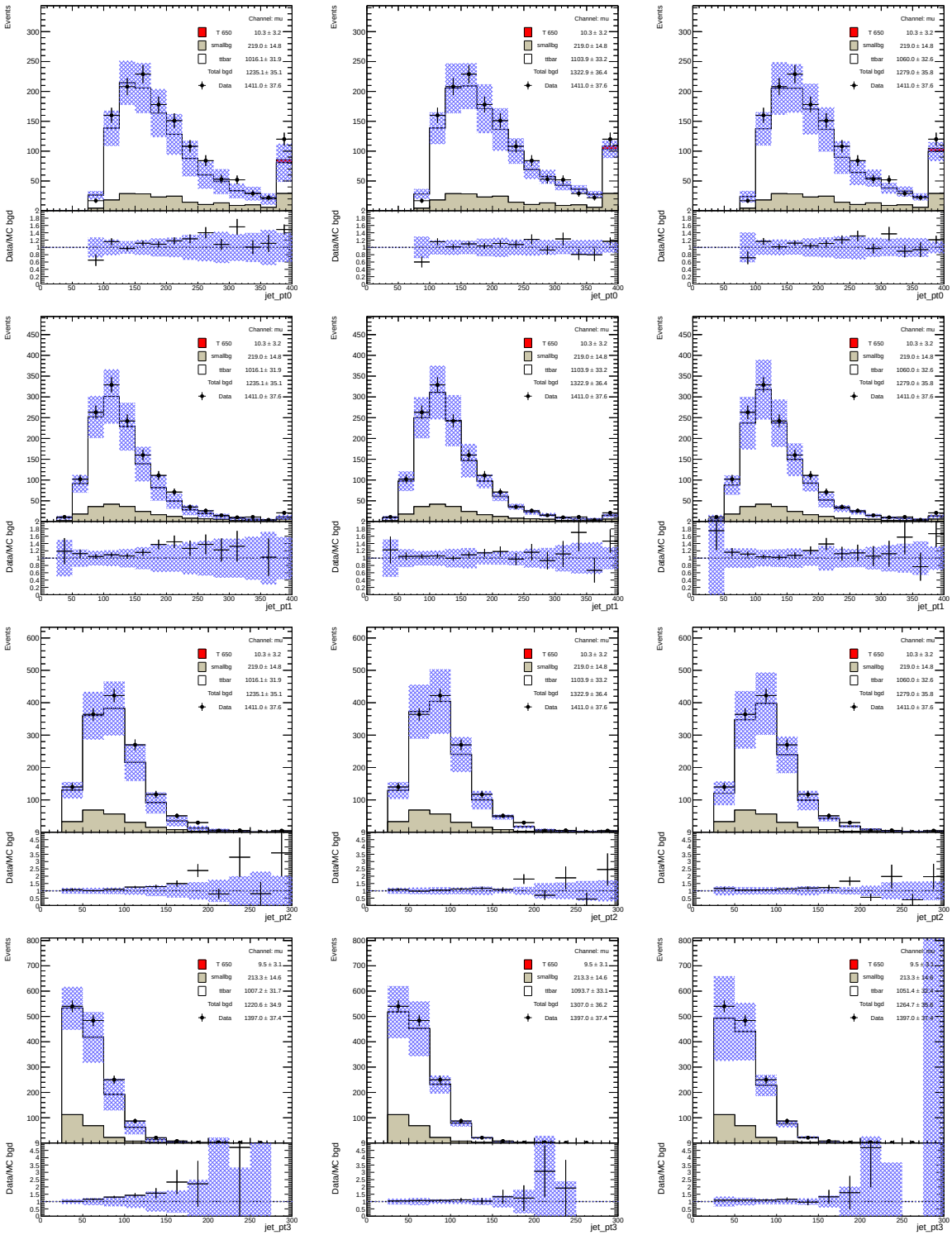


Figure C.41: Data/MC comparison of the p_{T} of the four leading jets after the W_{had}-selection for three samples generated using different generators: MC@NLO (left column), Powheg-Pythia (middle column), Powheg-Jimmy (right column). The blue shaded area represents the 1σ systematic error bands on the expected event yields.

C.6 Correlation Plots for the Default Training with T 650 GeV

Thiqs section contains two dimensional distributions showing the correlation between the m_T discriminant, the NN input variables as well as the H_{T}-variable components, and the NN output distribution (Section C.6.1). The correlations of the NN input variables to the m_T discriminant are shown in Section C.6.2. All plots are presented for the signal (left), combined backgrounds (middle) and data (right) samples, after the W_{had}-Selection. The average values of the variable in question over five bins of the MLP variable are overlaid in the plots (dark red histogram).

The reconstructed T-quark mass, the discriminant of the analysis, shows a moderate linear correlation to the MLP output. The p_{T}-like variables (Figure C.42) are more correlated to the NN output than the angular variables (Figure C.43). Among the variables which compose the H_{T} variable, the p_{T} of the lepton shows the weakest correlation to the NN output.

The correlations of the NN input variables and the reconstructed T-quark mass are shown in Figure C.6.2.

C.6.1 Input Variables versus the NN output, W_{had}-Selection

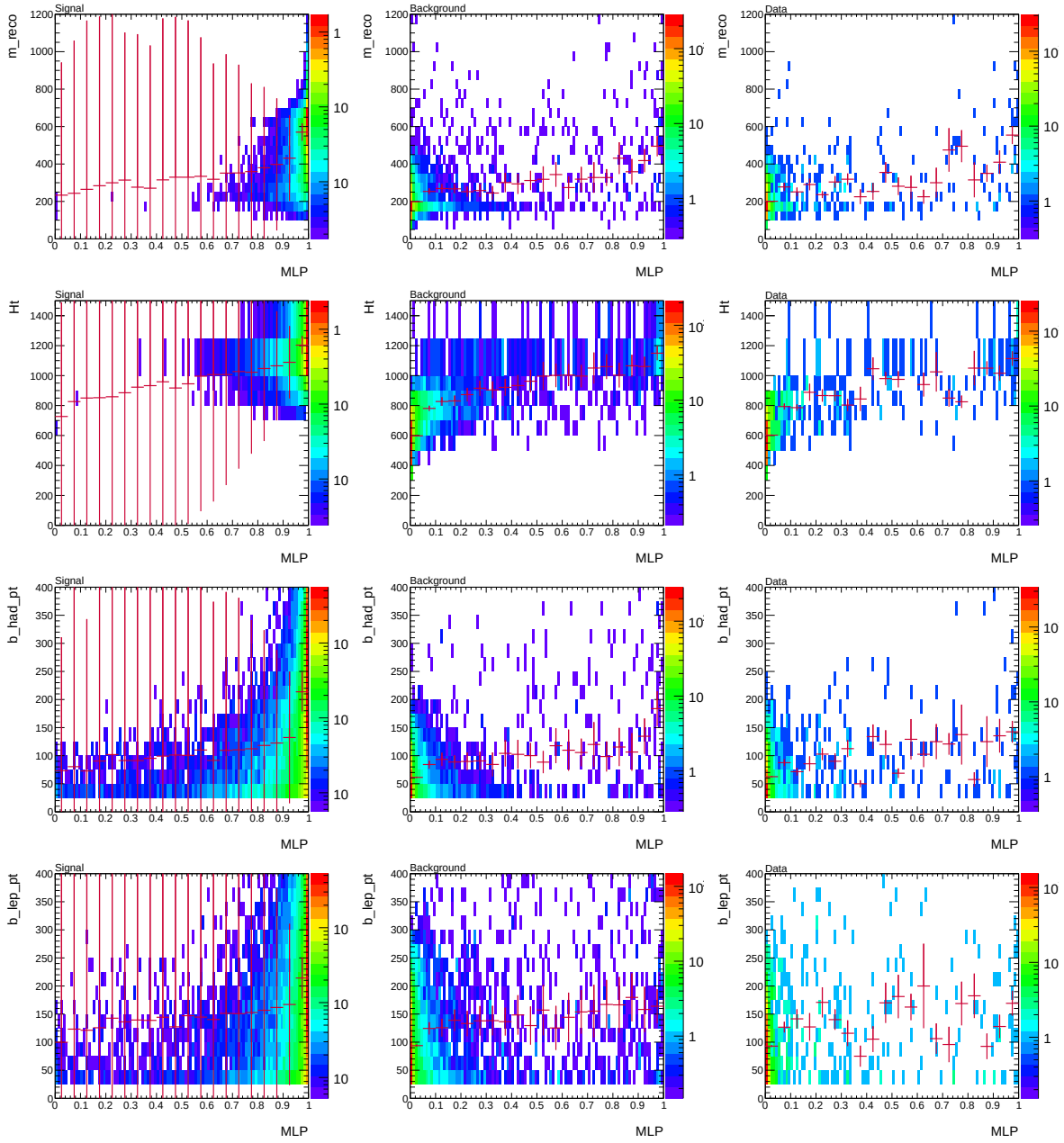


Figure C.42: Scatter plots of the M_{reco}, H_{T}, $p_{\mathrm{T}}^{b_{\mathrm{had}}}$ and $p_{\mathrm{T}}^{b_{\mathrm{lep}}}$ variables versus the NN response (MLP) arranged in this order from the top to the bottom of the figure. The plots are presented for the signal (left), the background (middle) and the data (right) samples after the W_{had}-Selection.

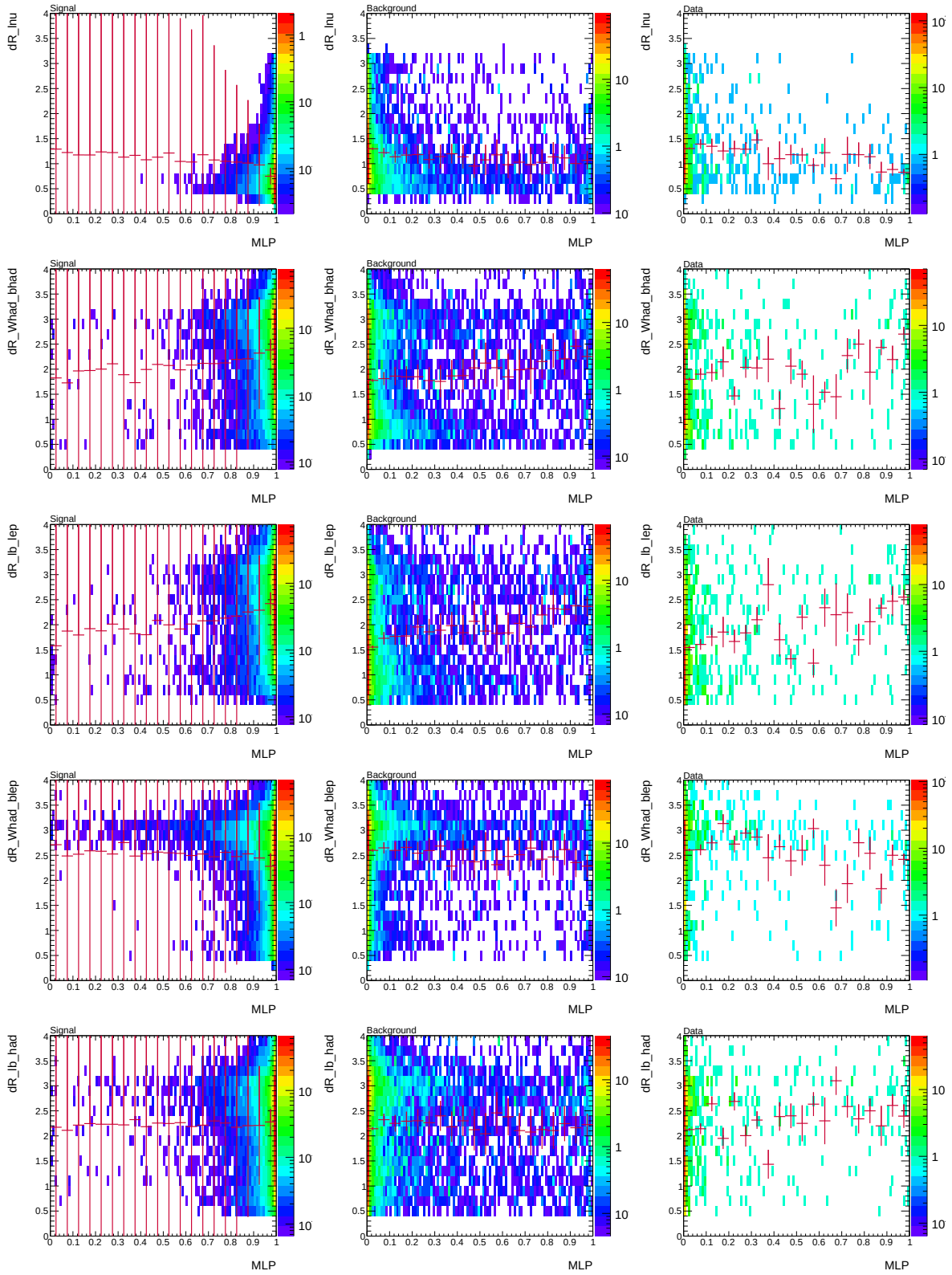


Figure C.43: Scatter plots of the $\Delta R_{l,\nu}$, $\Delta R_{W_{\mathrm{had}},b_{\mathrm{had}}}$, $\Delta R_{l,b_{\mathrm{lep}}}$, $\Delta R_{W_{\mathrm{had}},b_{\mathrm{lep}}}$ and $\Delta R_{l,b_{\mathrm{had}}}$ variables versus the NN response (MLP) arranged in this order from the top to the bottom of the figure. The plots are presented for the signal (left), the background (middle) and the data (right) samples after the W_{had}-Selection.

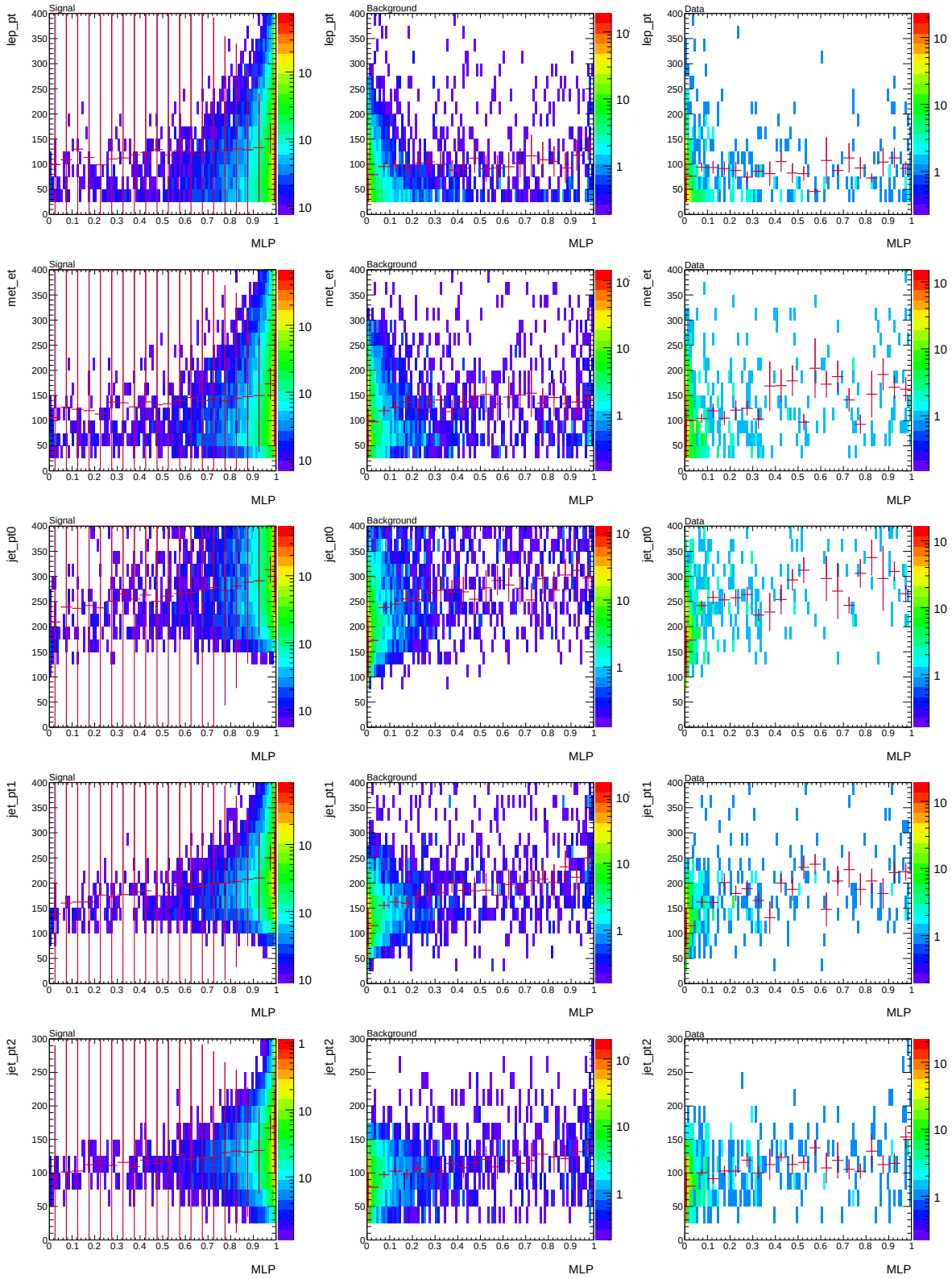


Figure C.44: Scatter plots of the p_{T} of the lepton, $E_{\mathrm{T}}^{\mathrm{miss}}$, as well as the p_{T} of three hardest jets versus the NN response (MLP) arranged in this order from the top to the bottom of the figure. The plots are presented for the signal (left), the background (middle) and the data (right) samples after the W_{had}-Selection.

C.6.2 Input Variables versus the M_{reco} discriminant, W_{had}-Selection

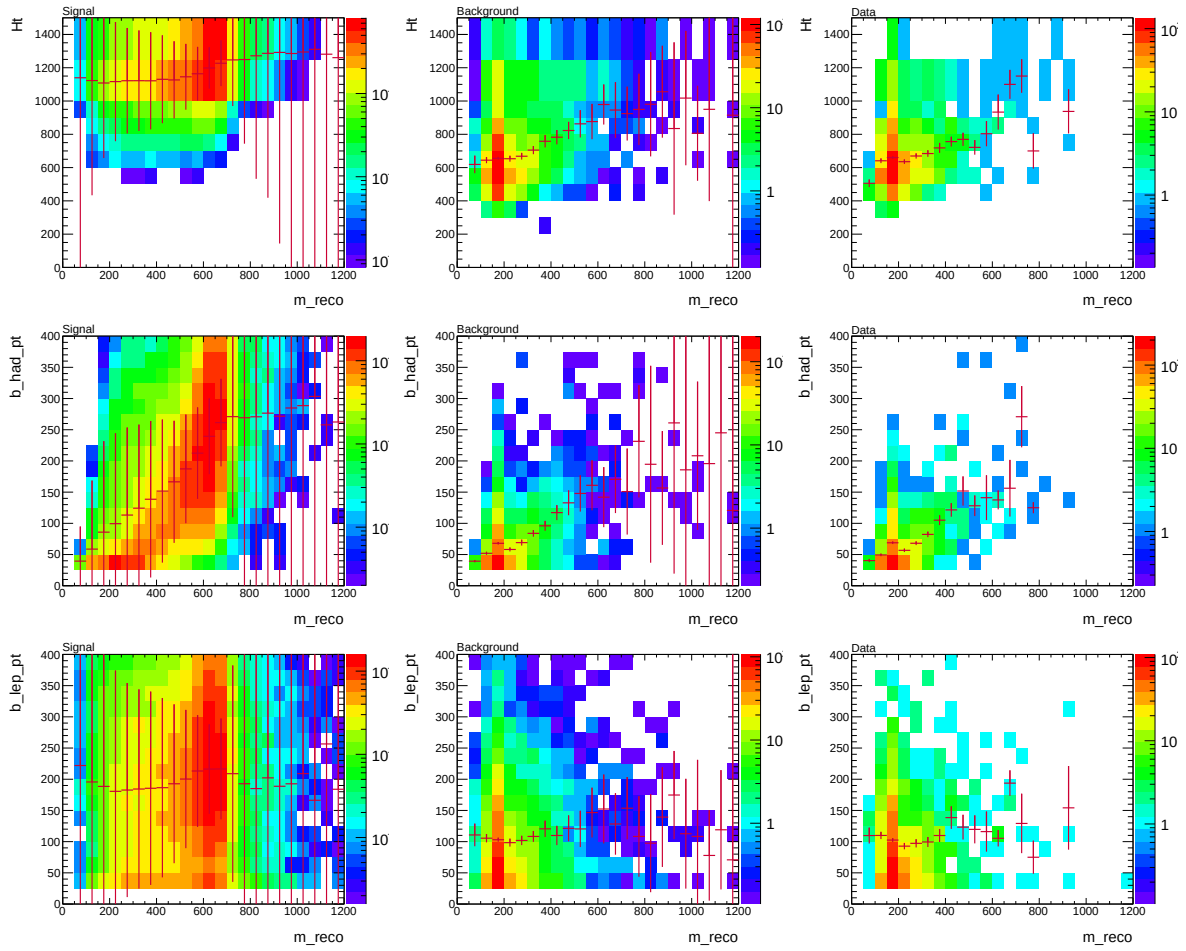


Figure C.45: Scatter plots of the H_{T}, $p_{\mathrm{T}}^{b_{\mathrm{had}}}$ and $p_{\mathrm{T}}^{b_{\mathrm{lep}}}$ variables versus the reconstructed T-quark mass arranged in this order from the top to the bottom of the figure. The plots are presented for the signal (left), the background (middle) and the data (right) samples after the W_{had}-Selection.

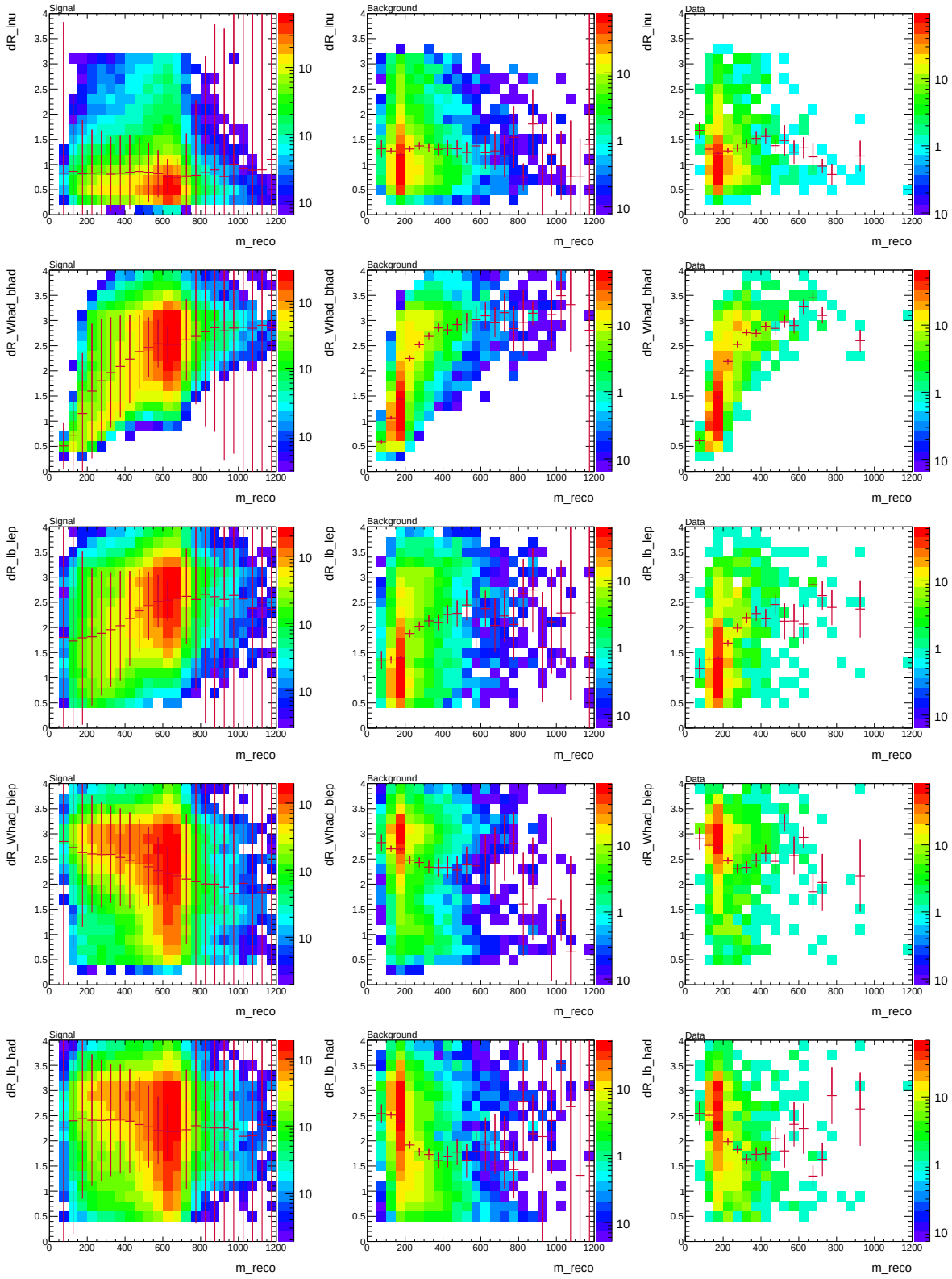


Figure C.46: Scatter plots of the $\Delta R_{l,\nu}$, $\Delta R_{W_{\mathrm{had}},b_{\mathrm{had}}}$, $\Delta R_{l,b_{\mathrm{lep}}}$, $\Delta R_{W_{\mathrm{had}},b_{\mathrm{lep}}}$ and $\Delta R_{l,b_{\mathrm{had}}}$ variables versus the reconstructed T-quark mass arranged in this order from the top to the bottom of the figure. The plots are presented for the signal (left), the background (middle) and the data (right) samples after the W_{had}-Selection.

C.6.3 Input Variables and Jet Multiplicity Correlations, NN>0.9965 Selection

The plots presented in this section show those correlations of the variables which reveal the goodness of the reconstruction of the selected data events. The events with a very large space angle between a W-boson and the accompanying b-jet while at the same time a small space angle between a W-boson and a b-jet from the opposite side of the decay are likely to not be well reconstructed.

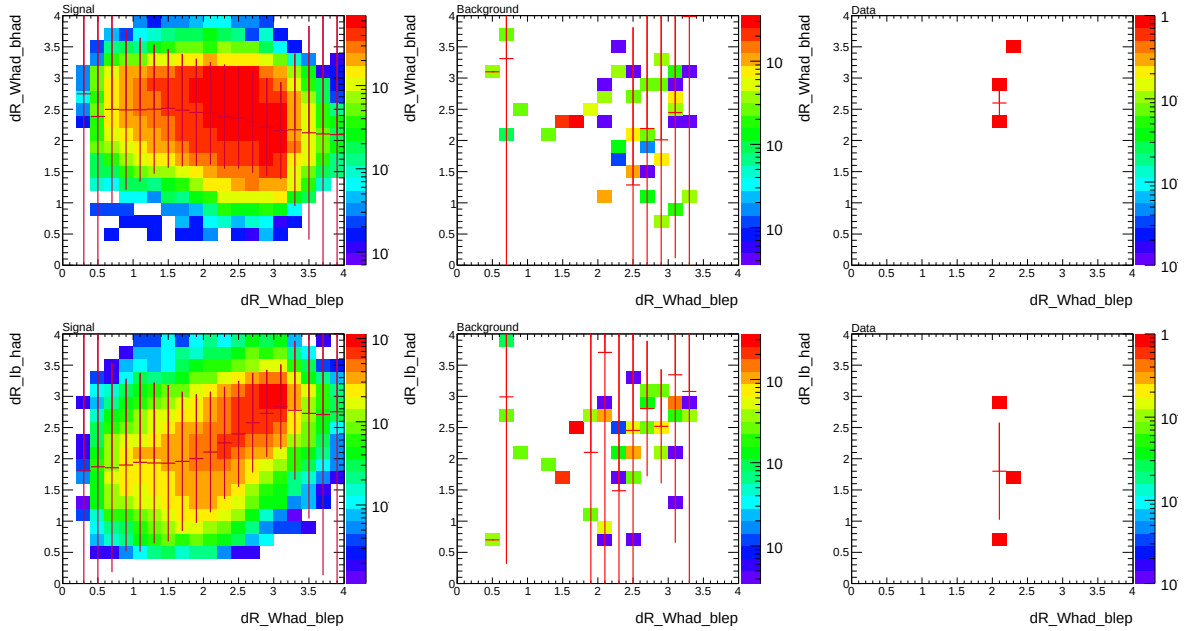


Figure C.47: $\Delta R(W_{\text{had}}, b_{\text{had}})$ and $\Delta R(l, b_{\text{had}})$ versus $\Delta R(W_{\text{had}}, b_{\text{lep}})$ for the electron channel.

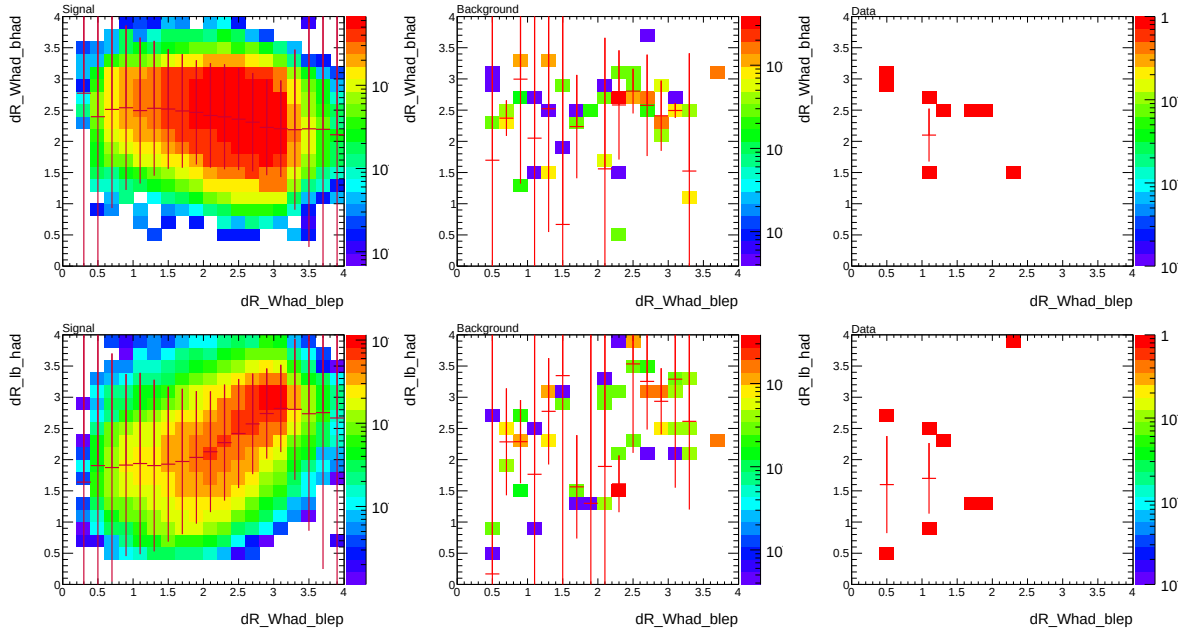


Figure C.48: $\Delta R(W_{\text{had}}, b_{\text{had}})$ and $\Delta R(l, b_{\text{had}})$ versus $\Delta R(W_{\text{had}}, b_{\text{lep}})$ for the muon channel.

C.7 Data Events

In this section all data events passing final selections are printed out in two tables. Both tables contain the event identification information such as run number, event number, the lepton channel, as well as the LOOSE and TIGHT selection flags and the value of the NN response. In addition, Table C.2 contains values of the NN input variables, while Table C.3 contains the p_{T} of the lepton, missing transverse energy and some information about the reconstructed W-bosons.

It is interesting to note that the events selected with the NN selection have notably higher H_{T} than those passing the TIGHT selection. Nevertheless, only two of the events selected by the NN selection are passing the TIGHT selection. Furthermore, two events selected by the NN selection do not pass neither LOOSE nor TIGHT selection. This clearly shows that the NN selection is able to recover the events which are rejected by some of the cuts.

TIGHT Selection															
Run Nr	Event Nr	lep	ch	loose	tight	NN	M_{reco}	H_{T}	$p_{\text{T}}^{b_{\text{had}}}$	$p_{\text{T}}^{b_{\text{lep}}}$	$\Delta R^{lb_{\text{had}}}$	$\Delta R^{lb_{\text{lep}}}$	$\Delta R^{W_{\text{had}}b_{\text{had}}}$	$\Delta R^{W_{\text{had}}b_{\text{lep}}}$	$\Delta R^{l\nu}$
191517	33216587	e	2	1	1	0.9976765	651.8	1251.6	101.1	273.3	1.78	2.96	3.47	2.21	0.89
191190	81583717	e	1	1	1	0.9936956	393.3	1026.1	64.6	291.6	3.39	1.75	2.38	2.78	0.86
190975	44113990	e	2	1	1	0.9813492	454.4	916.0	94.2	266.3	2.72	2.86	3.73	3.09	0.85
187815	45639092	e	2	1	1	0.9824427	418.2	878.1	140.3	207.9	1.92	1.99	2.98	2.71	0.78
183407	26029238	e	2	1	1	0.9765413	647.0	862.9	201.4	105.6	3.44	2.64	2.74	2.48	0.52
180225	123447880	e	2	1	1	0.9690316	741.7	1029.0	187.4	239.4	1.80	2.05	3.51	3.36	0.87
191513	16197594	μ	2	1	1	0.8124023	349.6	1073.7	76.4	274.3	1.76	1.50	2.98	2.73	1.17
190933	103160045	μ	2	1	1	0.9885802	550.6	1124.3	86.4	466.9	3.04	2.67	3.12	3.01	1.13
190872	78991132	μ	2	1	1	0.8977577	315.3	934.2	78.7	186.0	2.47	2.47	2.13	2.10	1.12
190644	14177447	μ	2	1	1	0.9990669	331.2	1123.4	186.9	321.6	3.98	2.23	1.51	2.36	0.94
190611	6404553	μ	1	1	1	0.9684797	587.5	1240.3	131.1	171.9	1.58	2.52	2.43	2.70	1.33
190618	71119922	μ	2	1	1	0.6589344	460.8	810.3	220.3	76.6	3.10	1.40	2.48	2.98	0.80
183127	1492109	μ	2	1	1	0.9783925	578.0	1063.5	279.0	63.7	2.82	2.10	2.91	2.92	0.72
NN>0.9965 Selection															
Run Nr	Event Nr	lep	ch	loose	tight	NN	M_{reco}	H_{T}	$p_{\text{T}}^{b_{\text{had}}}$	$p_{\text{T}}^{b_{\text{lep}}}$	$\Delta R^{lb_{\text{had}}}$	$\Delta R^{lb_{\text{lep}}}$	$\Delta R^{W_{\text{had}}b_{\text{had}}}$	$\Delta R^{W_{\text{had}}b_{\text{lep}}}$	$\Delta R^{l\nu}$
191426	42779435	e	1	0	0	0.9987248	636.5	2650.6	82.9	38.2	2.91	4.23	2.86	2.16	0.73
191517	33216587	e	2	1	1	0.9976765	651.8	1251.6	101.1	273.3	1.78	2.96	3.47	2.21	0.891
189425	13984521	e	2	1	0	0.9997664	853.8	1497.7	415.8	502.3	0.70	2.49	2.28	2.05	0.70
190644	14177447	μ	2	1	1	0.9990669	331.2	1123.4	186.9	321.6	3.98	2.22	1.50	2.35	0.936
190236	27781298	μ	2	1	0	0.9978179	928.0	1354.2	293.9	204.1	2.71	2.85	2.86	0.449	0.727
189184	9180423	μ	2	1	0	0.9991104	580.7	1125.0	382.9	370.3	1.26	2.88	2.43	1.61	1.35
186456	72270691	μ	2	1	0	0.9965881	683.3	1339.4	336.2	312.7	0.517	2.71	3.12	0.574	0.543
183003	128326601	μ	2	0	0	0.9988620	680.1	1165.8	312.7	292.9	1.30	2.05	2.52	1.82	1.41
182486	125455652	μ	1	1	0	0.9974486	949.3	2091.2	271.8	292.1	0.803	2.86	2.64	1.09	0.599
182486	117473436	μ	2	1	0	0.9974595	448.5	1291.4	349.6	100.7	2.40	1.85	1.59	1.01	0.363
182284	70436527	μ	2	1	0	0.9994190	635.6	1457.8	184.6	100.9.52	2.34	3.39	2.45	1.27	0.659

Table C.2: Data events passing the TIGHT and the NN>0.9965 selections printed out. All p_{T}-like variables are expressed in GeV. The column "ch" denotes the analysis channel: 1 stands for the boosted and 2 for the resolved W_{had}.

TIGHT Selection															
Run Nr	Event Nr	lep	ch	tight	NN	N_{pv}	p_T^{lep}	E_T^{miss}	$p_T^{j_{lead}}$	$p_T^{j_{sub}}$	$m^{W_{had}}$	$p_T^{W_{had}}$	$\Delta R^{W_{had}}$	$m^{W_{lep}}$	$p_T^{W_{lep}}$
191517	33216587	e	2	1	0.9976765	4	58.6	140.0	469.7	208.9	108.5	495.3	0.78	80.4	18.9
191190	81583717	e	1	1	0.9936956	5	157.7	52.6	408.7	50.9	73.6	408.7	0	80.4	208.2
190975	44113990	e	2	1	0.9813492	11	185.1	79.5	151.6	139.3	70.2	219.7	0.59	99.8	245.1
187815	45639092	e	2	1	0.9824427	2	129.3	81.6	198.8	120.3	87.4	264.6	0.63	80.4	202.5
183407	26029238	e	2	1	0.9765413	5	203.9	122.2	159.9	69.9	72.1	229.7	0.61	80.4	316.0
180225	123447880	e	2	1	0.9690316	3	48.1	182.0	318.1	54.1	89.8	371.6	0.60	80.4	221.0
191513	16197594	μ	2	1	0.8124023	6	22.3	187.8	336.7	176.3	81.5	369.2	0.59	80.4	210.0
190933	103160045	μ	2	1	0.9885802	6	29.2	231.9	262.9	47.0	85.9	309.9	0.55	88.0	245.8
190872	78991132	μ	2	1	0.8977577	6	173.0	46.2	230.2	180.8	102.4	344.7	0.57	94.9	197.5
190644	14177447	μ	2	1	0.9990669	6	79.8	87.1	289.3	158.7	62.2	225.5	0.49	80.4	165.2
190611	6404553	μ	1	1	0.9684797	6	21.6	189.0	465.3	261.2	74.4	465.3	0	80.4	197.3
190618	71119922	μ	2	1	0.6589344	7	214.2	49.5	144.1	105.5	73.6	191.6	0.78	80.4	252.4
183127	1492109	μ	2	1	0.9783925	4	151.8	164.6	217.1	185.6	68.8	247.0	0.76	111.7	296.1
NN>0.9965 Selection															
Run Nr	Event Nr	lep	ch	tight	NN	N_{pv}	p_T^{lep}	E_T^{miss}	$p_T^{j_{lead}}$	$p_T^{j_{sub}}$	$m^{W_{had}}$	$p_T^{W_{had}}$	$\Delta R^{W_{had}}$	$m^{W_{lep}}$	$p_T^{W_{lep}}$
191426	42779435	e		1 0	0.9987248	6	92.6	128.9	1218.8	1089.3	89.8	1089.3	0	80.4	215.5
191517	33216587	e		2 1	0.9976765	4	58.6	140.0	469.7	208.9	108.5	495.3	0.78	80.4	189.3
189425	13984521	e		2 0	0.9997664	7	115.6	111.2	261.9	90.8	89.7	344.3	0.54	80.4	226.2
190644	14177447	μ		2 1	0.9990669	6	79.8	87.1	289.3	158.7	62.2	225.5	0.49	80.4	165.2
190236	27781298	μ		2 0	0.9978179	8	43.0	280.2	304.5	228.6	96.4	397.0	0.50	80.4	319.7
189184	9180423	μ		2 0	0.9991104	4	38.2	92.6	136.4	104.6	80.4	233.2	0.61	80.4	117.9
186456	72270691	μ		2 0	0.9965881	3	89.1	266.8	180.5	154.0	94.2	326.4	0.54	82.7	346.2
183003	128326601	μ		2 0	0.9988620	6	26.0	180.5	182.9	170.8	105.8	346.1	0.56	89.0	186.3
182486	125455652	μ		1 0	0.9974486	2	47.0	386.4	665.4	428.5	78.9	665.4	0	80.4	428.7
182486	117473436	μ		2 0	0.9974595	3	187.3	261.2	231.0	161.6	67.2	264.6	0.64	80.4	444.4
182284	70436527	μ		2 0	0.9994190	3	52.3	274.4	572.5	273.1	76.4	603.9	0.49	80.4	326.6

Table C.3: Data events passing the TIGHT and the NN>0.9965 selections printed out. All p_T-like variables are expressed in GeV. The column "ch" denotes the analysis channel: 1 stands for the boosted and 2 for the resolved W_{had}.

C.8 Systematic Control Plots for the TIGHT Selection

C.8.1 Signal

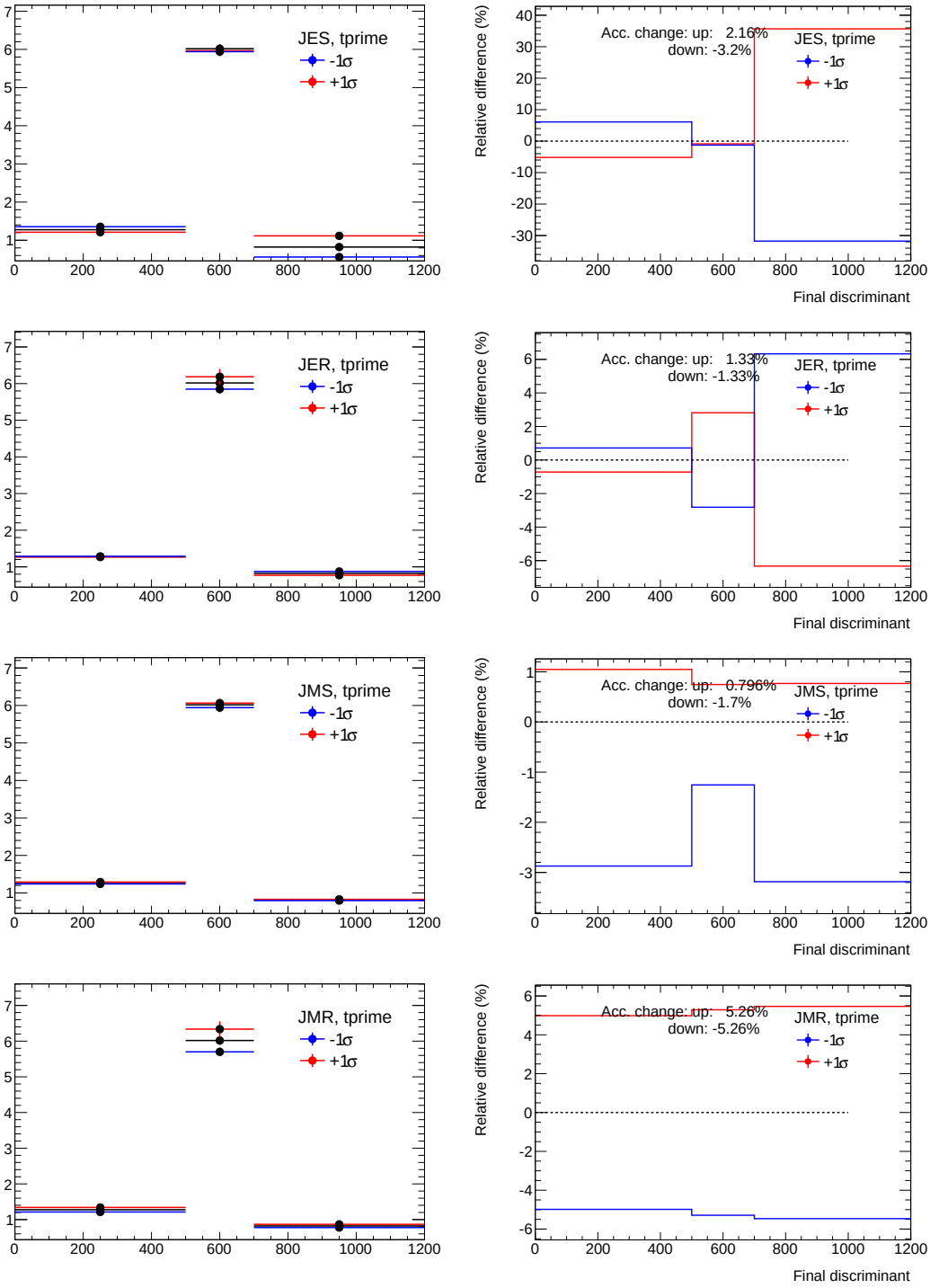


Figure C.49: Absolute (left) and relative (right) bin-wise systematic deviation with respect to the nominal T expectation for the M_{reco} variable in the final TIGHT selection. The systematic sources are from the top to the bottom: jet energy scale, jet energy resolution, jet mass scale, jet mass resolution.

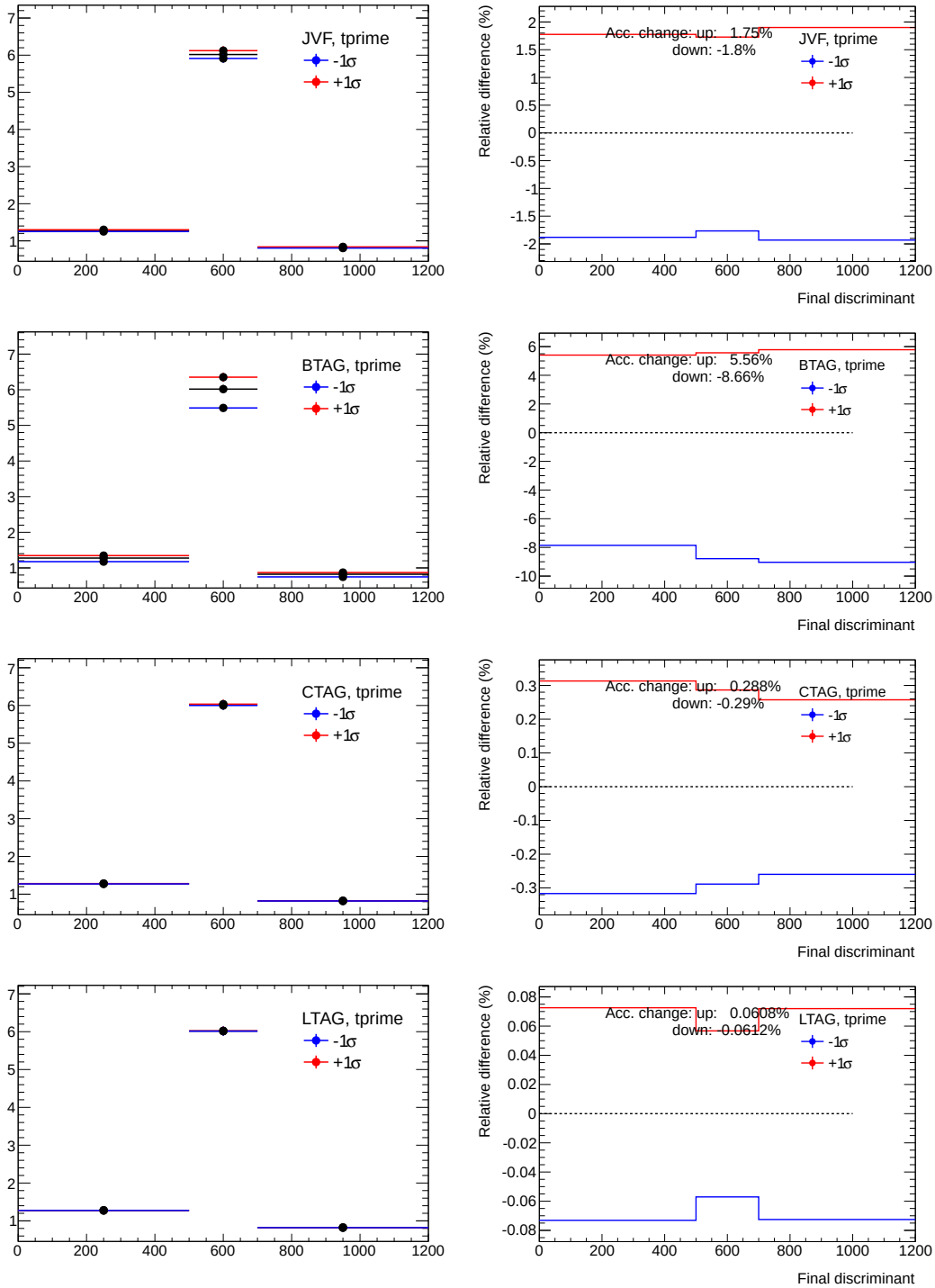


Figure C.50: Absolute (left) and relative (right) bin-wise systematic deviation with respect to the nominal T expectation for the M_{reco} variable in the final TIGHT selection. The systematic sources are from the top to the bottom: jet vertex fraction, b-tag scale factors, c-tag scale factors, mistag scale factors.

C.8.2 $t\bar{t}$ Background

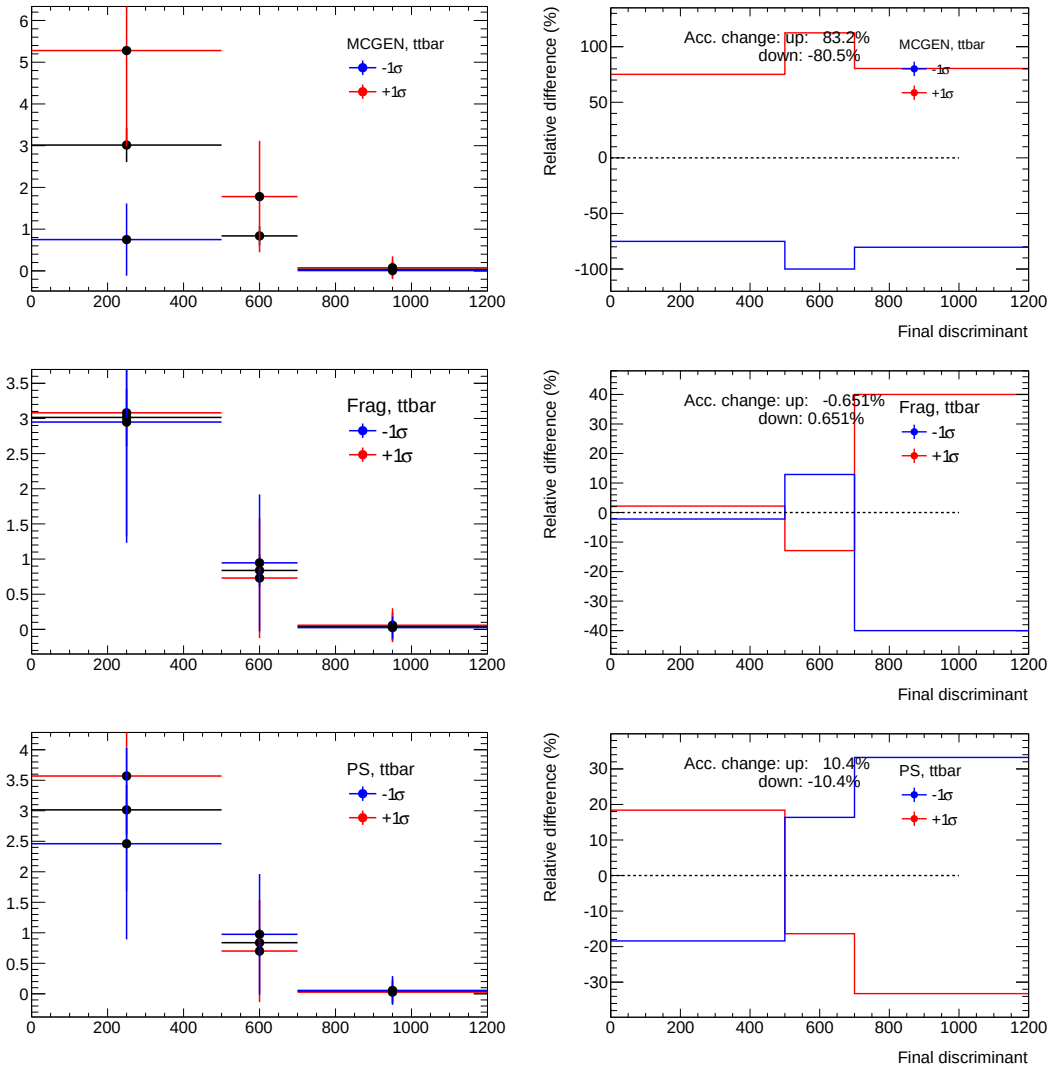


Figure C.51: Absolute (left) and relative (right) bin-wise systematic deviation with respect to the nominal $t\bar{t}$ expectation for the M_{reco} variable in the final TIGHT selection. The systematic sources are from the top to the bottom: NLO event generator (default MC@NLO vs systematic POWHEG interfaced with the same parton shower model in HERWIG), fragmentation model (default HERWIG vs systematic PYTHIA applied to the same parton level generator POWHEG), parton shower model (based on the ACERMC samples with more and less parton shower).

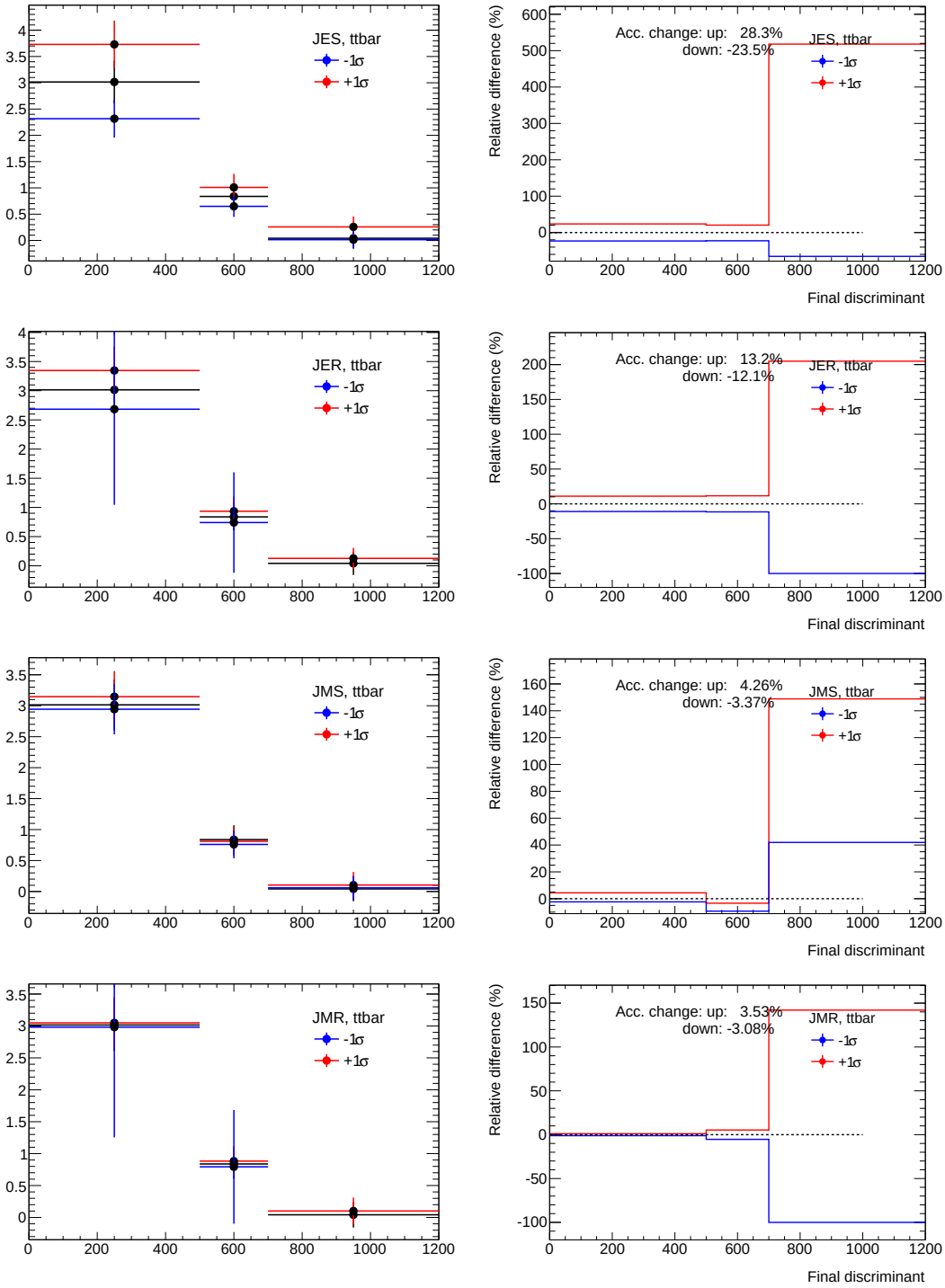


Figure C.52: Absolute (left) and relative (right) bin-wise systematic deviation with respect to the nominal $t\bar{t}$ expectation for the M_{reco} variable in the final TIGHT selection. The systematic sources are from the top to the bottom: jet energy scale, jet energy resolution, jet mass scale, jet mass resolution.

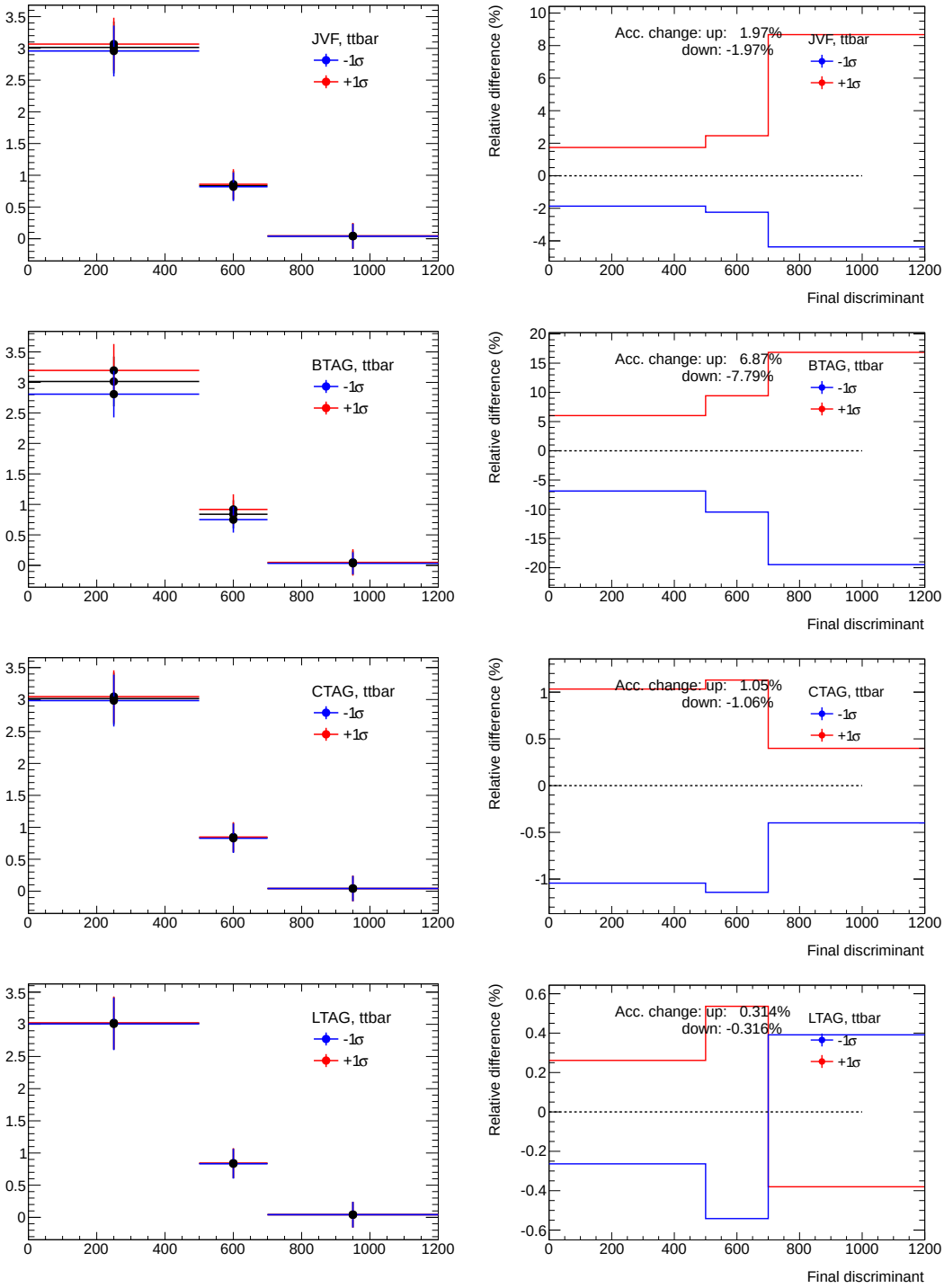


Figure C.53: Absolute (left) and relative (right) bin-wise systematic deviation with respect to the nominal $t\bar{t}$ expectation for the M_{reco} variable in the final TIGHT selection. The systematic sources are from the top to the bottom: jet vertex fraction, b-tag scale factors, c-tag scale factors, mistag scale factors.

C.8.3 Small Backgrounds

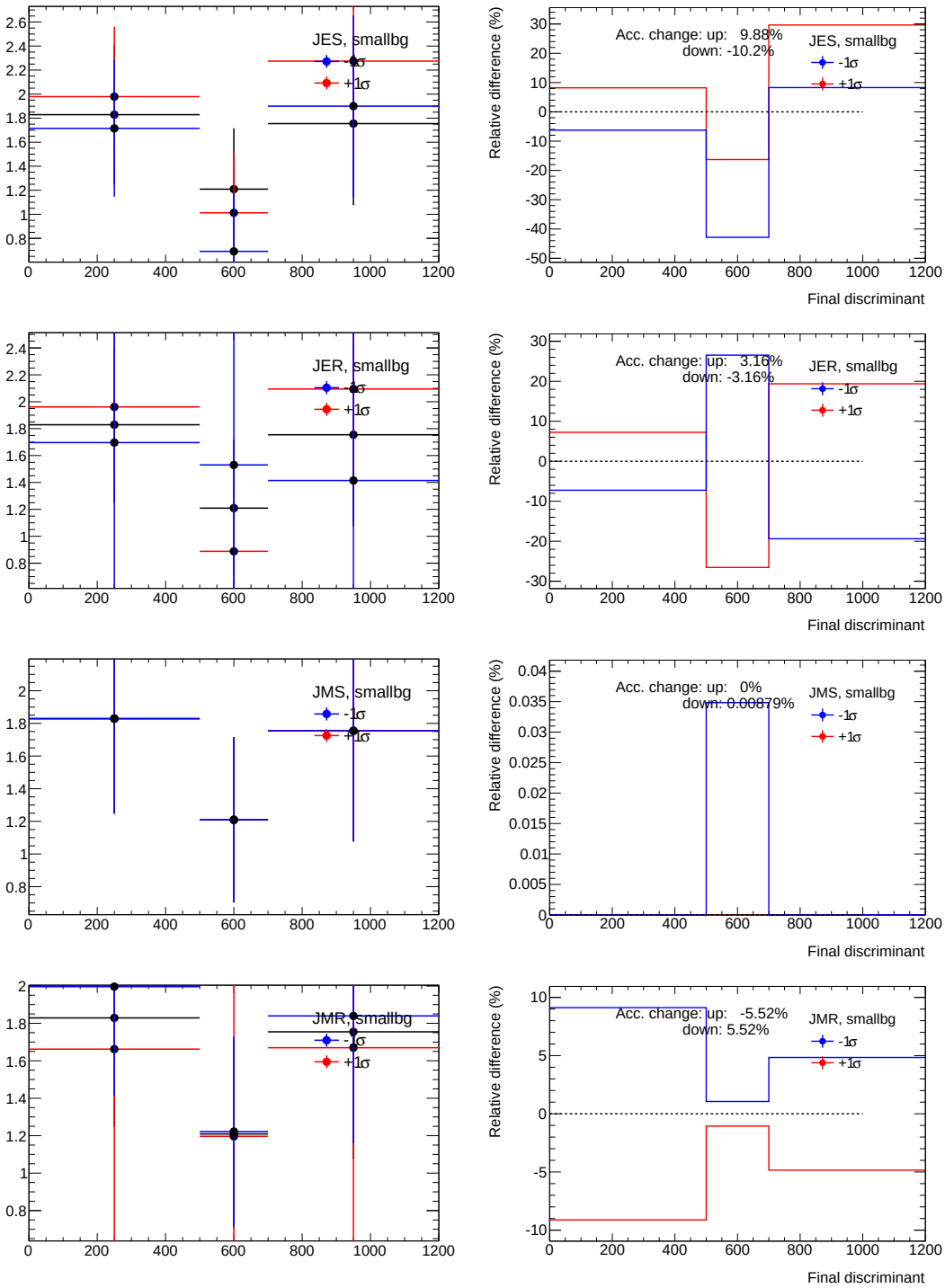


Figure C.54: Absolute (left) and relative (right) bin-wise systematic deviation with respect to the nominal small backgrounds expectation for the M_{reco} variable in the final TIGHT selection. The systematic sources are from the top to the bottom: jet energy scale, jet energy resolution, jet mass scale, jet mass resolution.

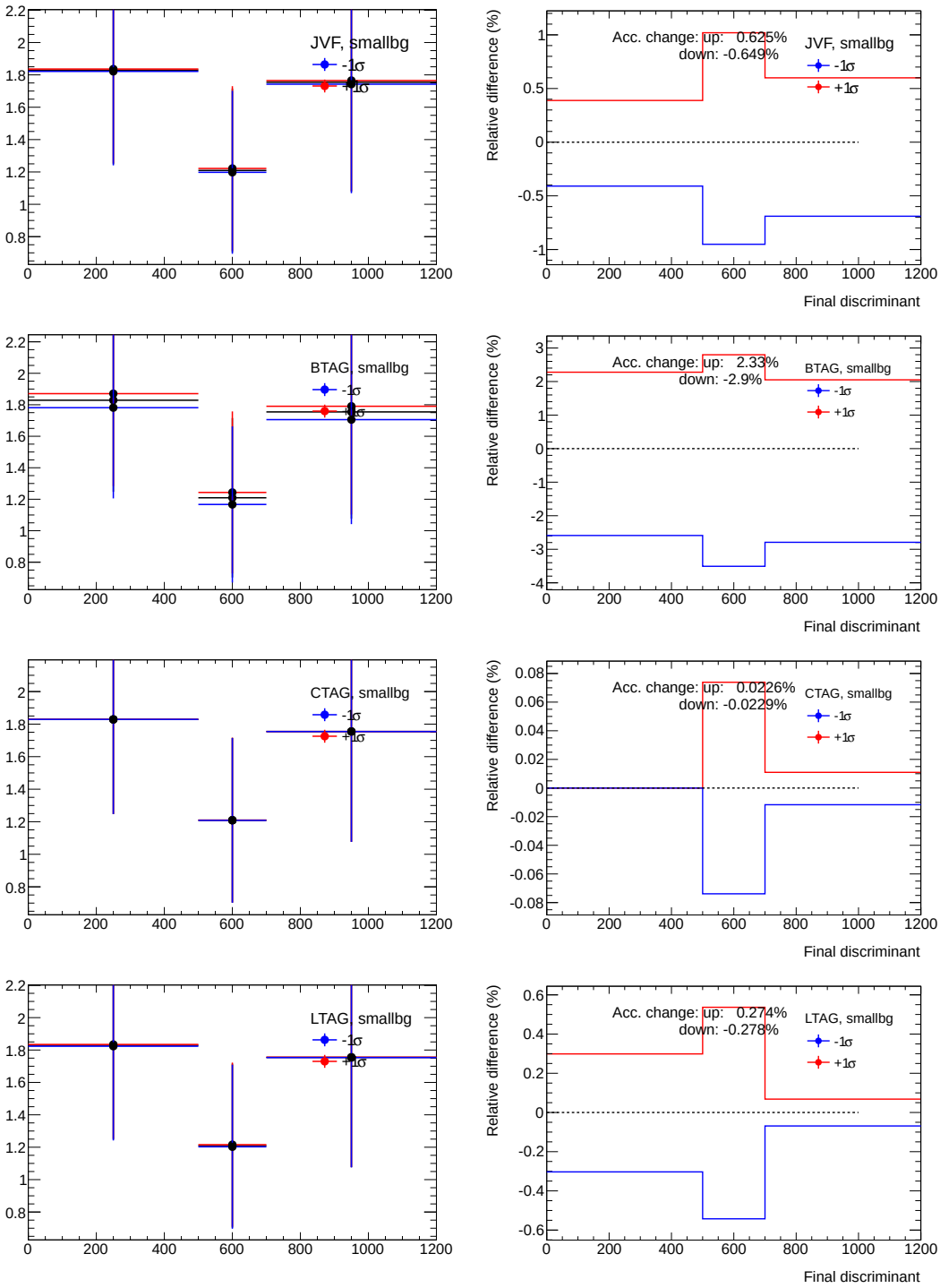


Figure C.55: Absolute (left) and relative (right) bin-wise systematic deviation with respect to the nominal small backgrounds expectation for the M_{reco} variable in the final TIGHT selection. The systematic sources are from the top to the bottom: jet vertex fraction, b-tag scale factors, c-tag scale factors, mistag scale factors.

C.9 Systematic Control Plots for the NN Selection

C.9.1 Signal

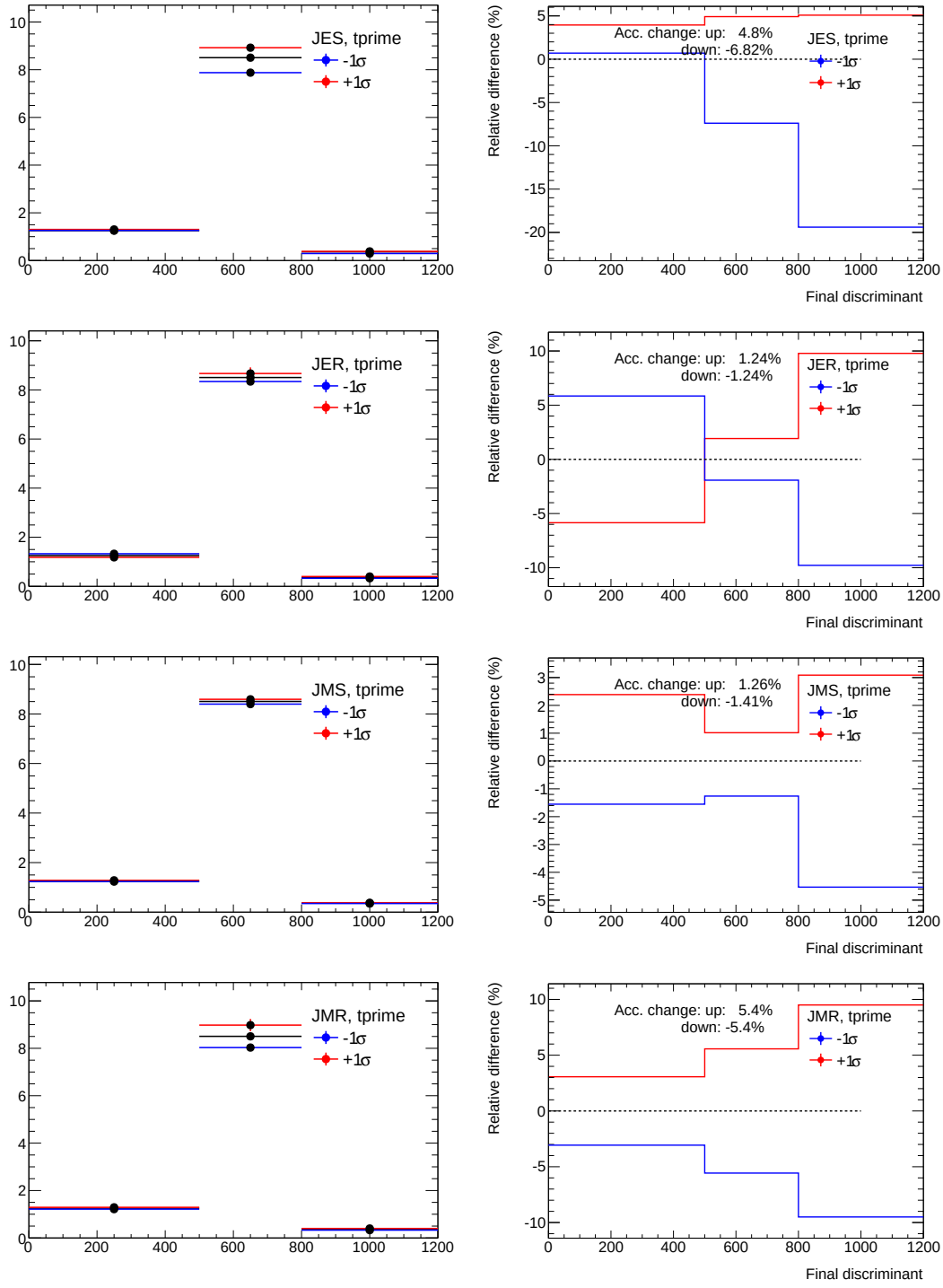


Figure C.56: Absolute (left) and relative (right) bin-wise systematic deviation with respect to the nominal T expectation for the M_{reco} variable in the final NN selection. The systematic sources are from the top to the bottom: jet energy scale, jet energy resolution, jet mass scale, jet mass resolution.

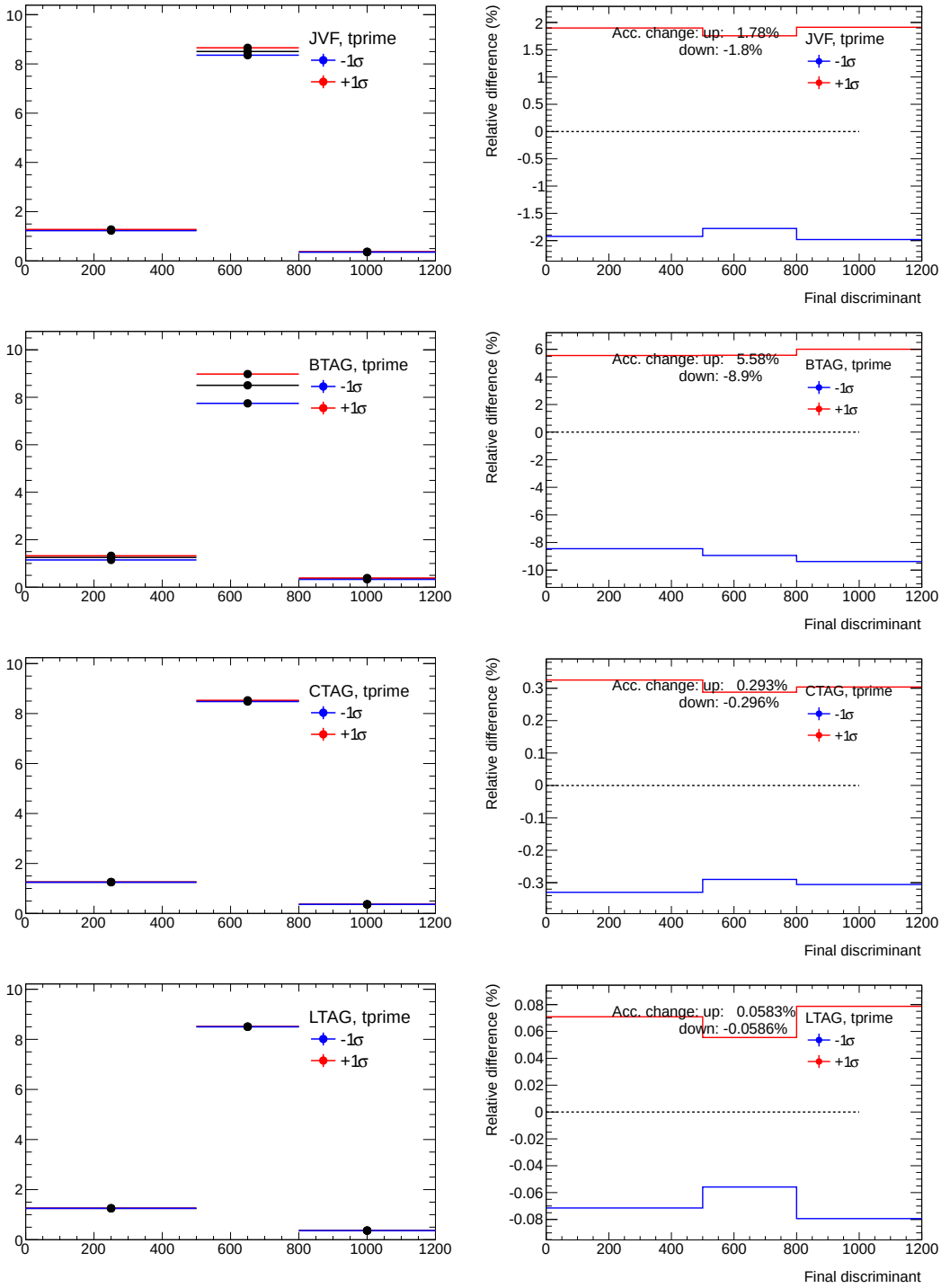


Figure C.57: Absolute (left) and relative (right) bin-wise systematic deviation with respect to the nominal T expectation for the M_{reco} variable in the final NN selection. The systematic sources are from the top to the bottom: jet vertex fraction, b-tag scale factors, c-tag scale factors, mistag scale factors.

C.9.2 $t\bar{t}$ Background

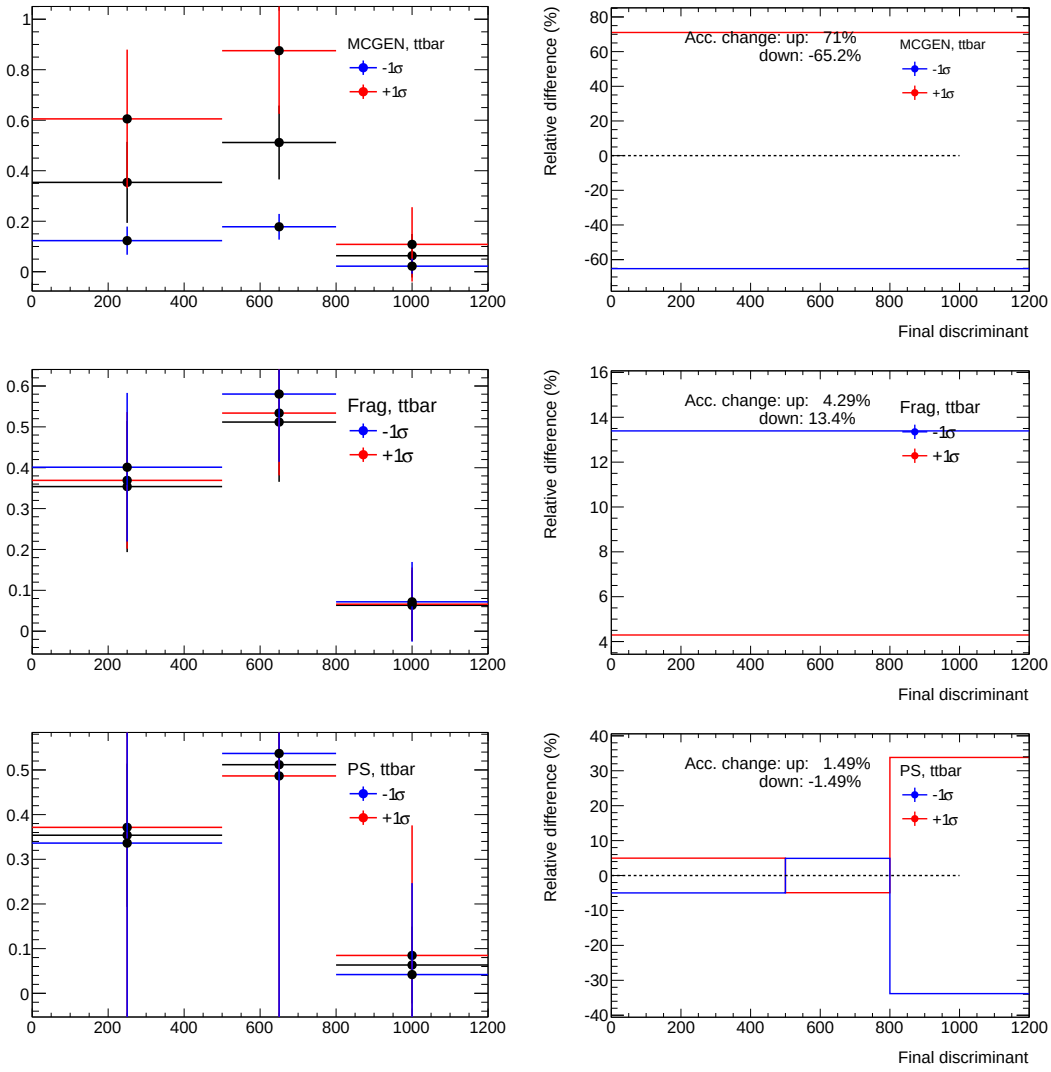


Figure C.58: Absolute (left) and relative (right) bin-wise systematic deviation with respect to the nominal $t\bar{t}$ expectation for the M_{reco} variable in the final NN selection. The systematic sources are from the top to the bottom: NLO event generator (default MC@NLO vs systematic POWHEG interfaced with the same parton shower model in HERWIG), fragmentation model (default HERWIG vs systematic PYTHIA applied to the same parton level generator POWHEG), parton shower model (based on the ACERMC samples with more and less parton shower).

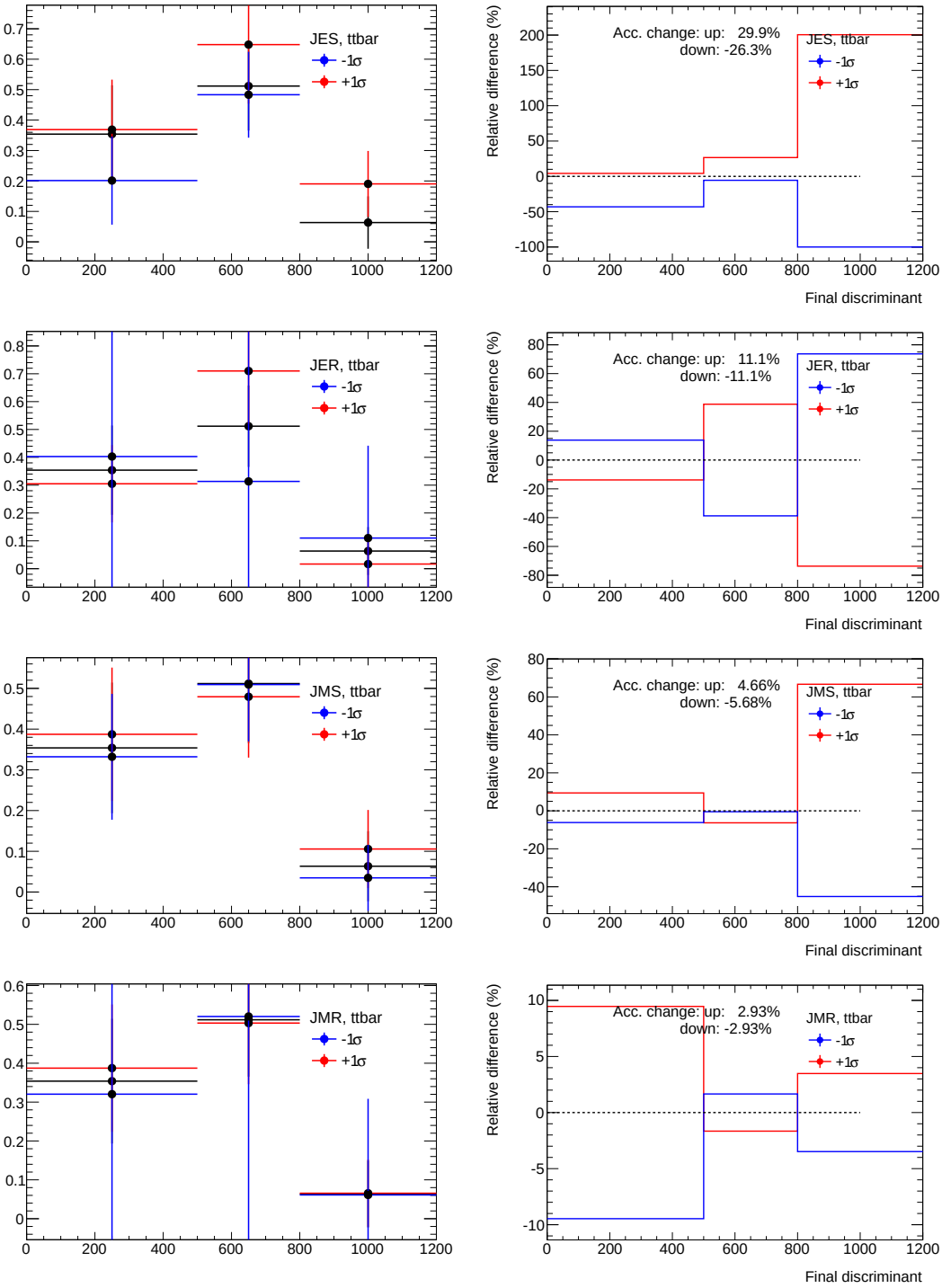


Figure C.59: Absolute (left) and relative (right) bin-wise systematic deviation with respect to the nominal $t\bar{t}$ expectation for the M_{reco} variable in the final NN selection. The systematic sources are from the top to the bottom: jet energy scale, jet energy resolution, jet mass scale, jet mass resolution.

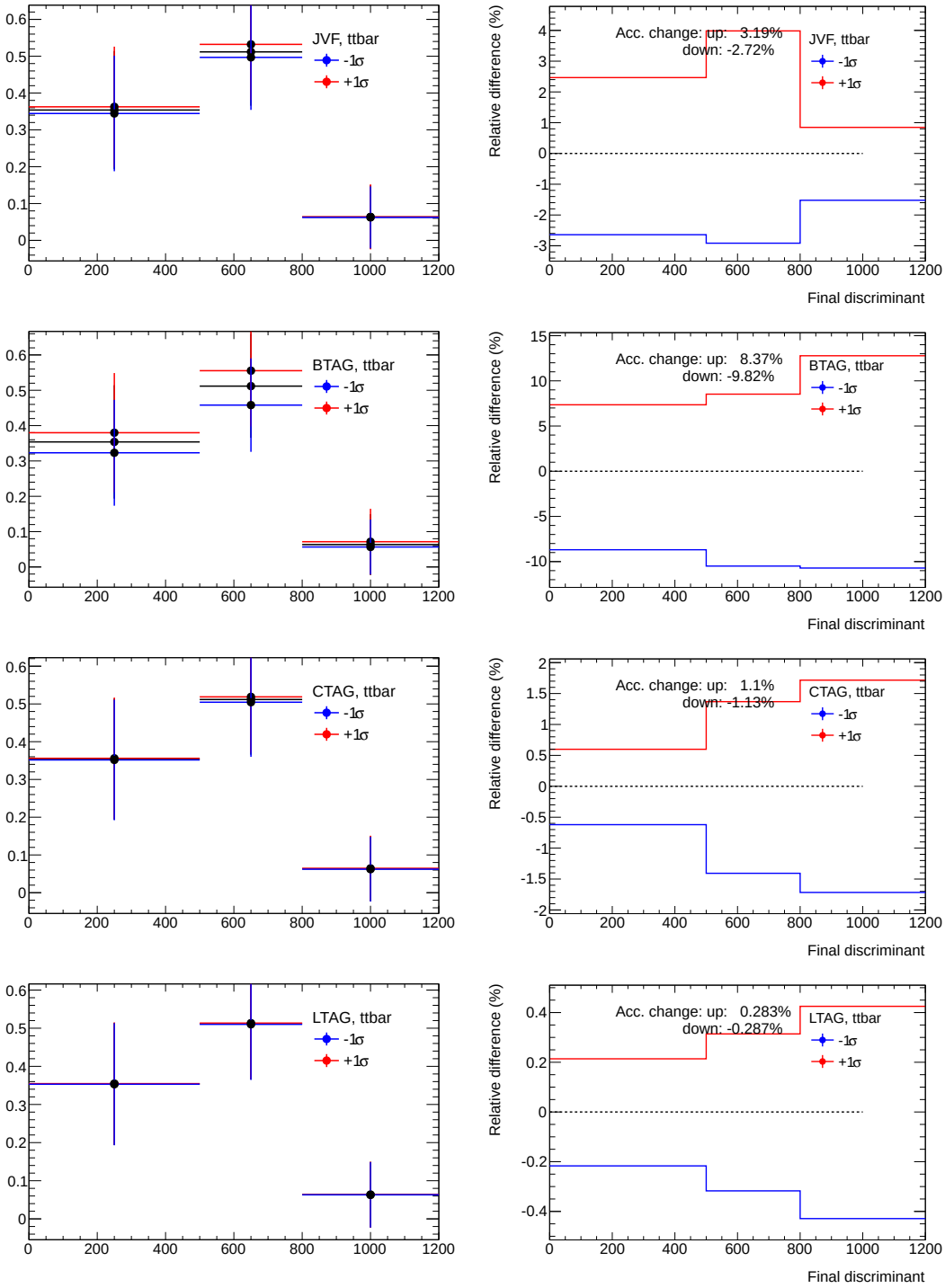


Figure C.60: Absolute (left) and relative (right) bin-wise systematic deviation with respect to the nominal $t\bar{t}$ expectation for the M_{reco} variable in the final NN selection. The systematic sources are from the top to the bottom: jet vertex fraction, b-tag scale factors, c-tag scale factors, mistag scale factors.

C.9.3 Small Backgrounds

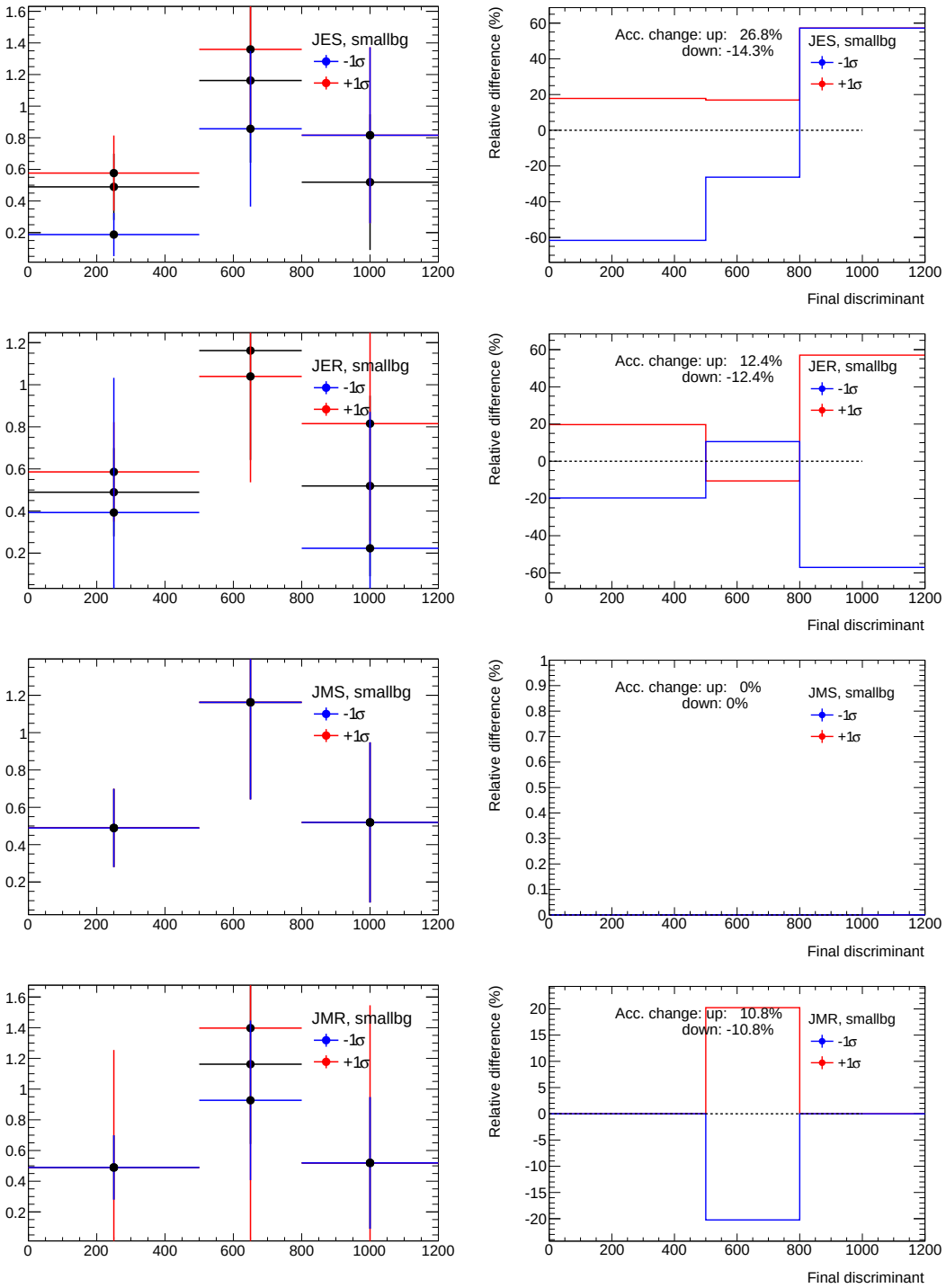


Figure C.61: Absolute (left) and relative (right) bin-wise systematic deviation with respect to the nominal small backgrounds expectation for the M_{reco} variable in the final NN selection. The systematic sources are from the top to the bottom: jet energy scale, jet energy resolution, jet mass scale, jet mass resolution.

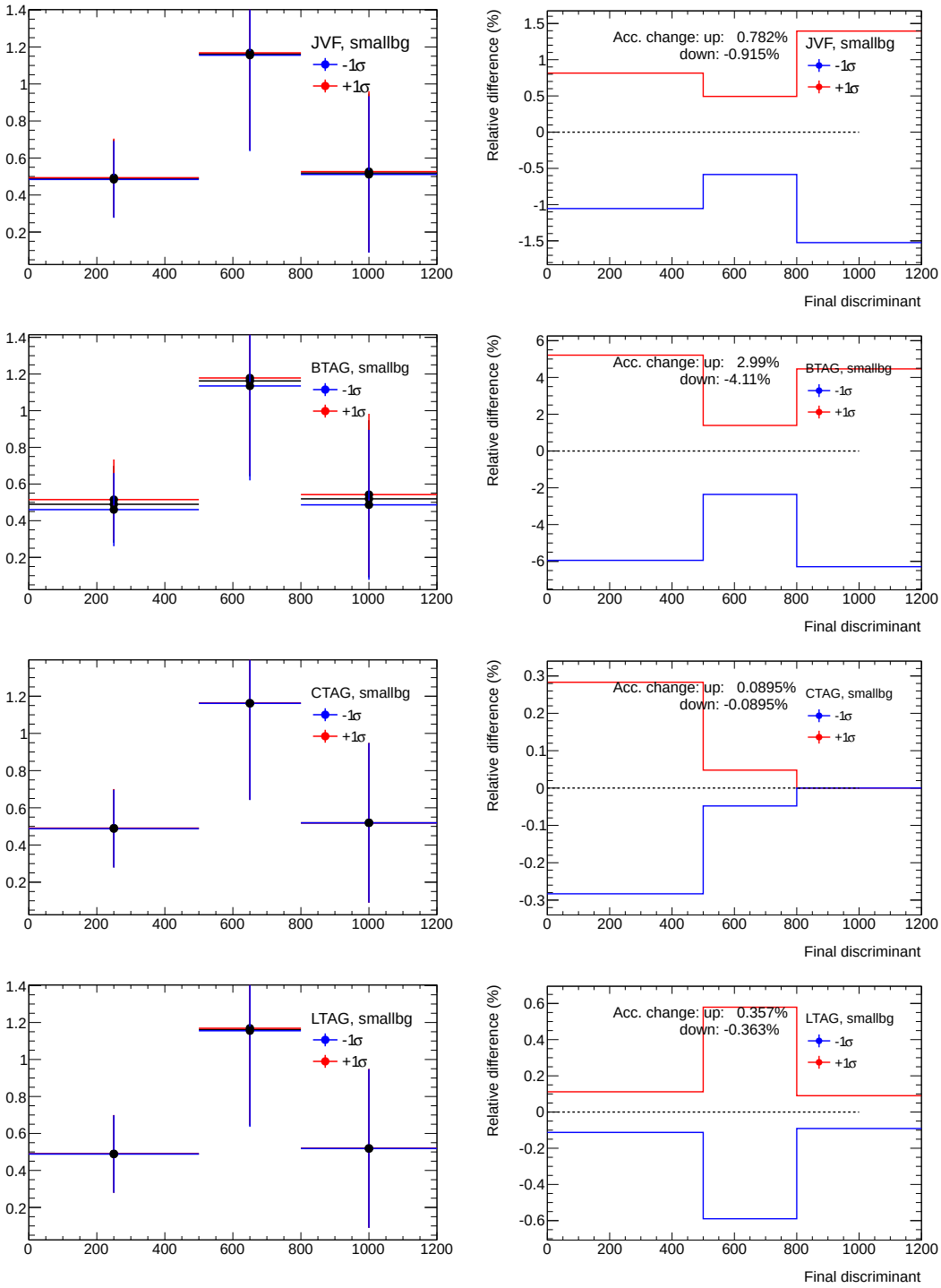


Figure C.62: Absolute (left) and relative (right) bin-wise systematic deviation with respect to the nominal small backgrounds expectation for the M_{reco} variable in the final NN selection. The systematic sources are from the top to the bottom: jet vertex fraction, b-tag scale factors, c-tag scale factors, mistag scale factors.

C.10 Sensitivity Test with a Single Bin

The final result of the analysis reported in this note has been produced by hypothesis testing on the M_{reco} discriminant with the binning (0, 500, 800, 1200) GeV after the NN-based final selection. With that selection, observation overshoots the expectation by the factor of 3.55. This data excess, however, turns out to be compatible with data within $\sim$ 1.6 s.d. This low sensitivity is due to the shape difference between the data and the $s+b$-expectation. However, the question arises how sensitive the analysis would be if the shape difference would be neglected and only acceptance considered. Thus, sensitivity and the exclusion limit have been tested with the same final selection, but a single bin of the M_{reco} discriminant, as in that case a maximal sensitivity is expected.

Figure C.63 shows the M_{reco} template with a single bin. The LLR plot as that explained in Section 9.7.2 shows that the observation is compatible with the b-hypothesis within 2 systematic s.d. and is more compatible to the $s+b$-hypothesis above $\sim$550 GeV. However, the plot showing $1-CLb$ as a function of the T-quark mass demonstrates that this excess is only $\sim$2.2 s.d. away from the b-hypothesis, which is not significant enough to claim an observation. The plot of CLs as a function of the T-mass shows that the observed CLs is within 2 s.d. compatible with the SM expectation. The Expected limit of 700.5 GeV is very similar to that obtained with 3 bins (see Section 9.7.4), while the observed limit has dropped for 60 GeV.

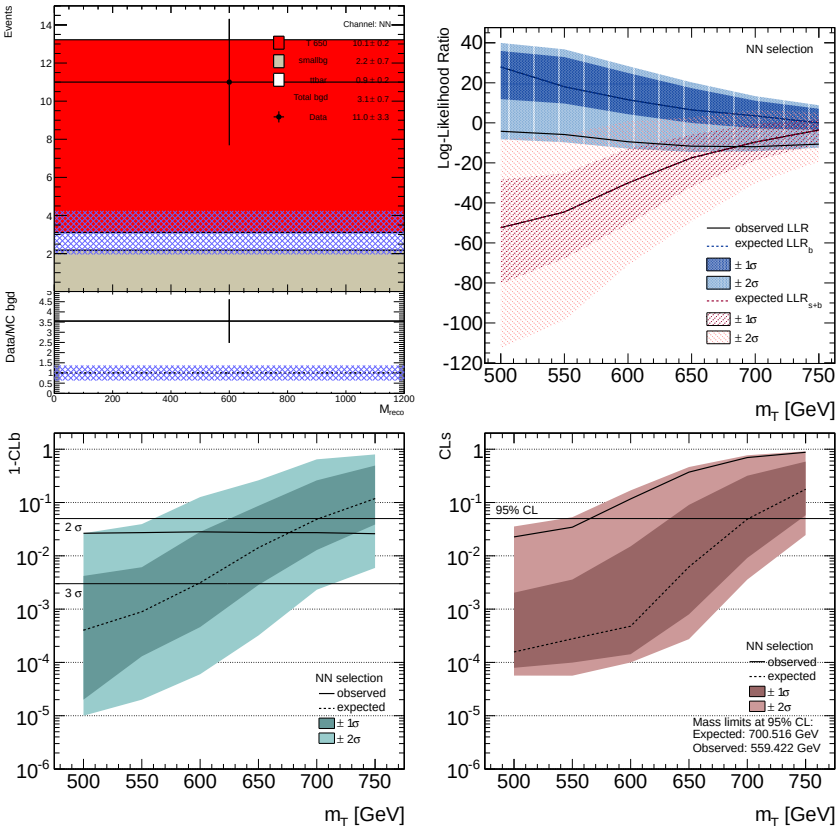


Figure C.63: M_{reco} template with a single bin (top left); the observed and the median expected LLR as well as the 1 and 2 s. d. width for the $s+b$ (red) and b (blue) hypotheses as a function of the T-quark mass (top right); expected and observed CL_b values as a function of the T mass assuming the theoretical cross section for the signal (bottom left); expected and observed CL_b values as a function of the T mass (bottom right).

Bibliography

[1] A. Zannoni, *On the Quantization of the Monoatomic Ideal Gas*, eprint arXiv:cond-mat/9912229 (1999) , cond-mat/9912229.

[2] P. A. Dirac, *On the Theory of quantum mechanics*, Proc.Roy.Soc.Lond. **A112** (1926) 661–677.

[3] S. N. Bose, *Plancks Gesetz und Lichtquantenhypothese*, Zeitschrift fur Physik **26** (1924) 178–181.

[4] Particle Data Group Collaboration, J. Beringer et al., *Review of Particle Physics (RPP)*, Phys.Rev. **D86** (2012) 010001.

[5] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys.Lett. **B716** (2012) 1–29, arXiv:1207.7214 [hep-ex].

[6] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys.Lett. **B716** (2012) 30–61, arXiv:1207.7235 [hep-ex].

[7] F. Halzen and A. D. Martin, *Quarks & Leptons: An Introduction in Modern Particle Physics*. Wiley, 1984.

[8] E. Noether, *Invariant Variation Problems*, Gott.Nachr. **1918** (1918) 235–257, arXiv:physics/0503066 [physics].

[9] F. Stöckli, *Effects of QCD Radiation on Higgs Production at the Large Hadron Collider*. PhD thesis, ETH Zuerich, 2008.

[10] C. N. Yang and R. L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, Phys. Rev. **96** (1954) 191–195.

[11] J. Ellis, *Higgs Physics*, arXiv:1312.5672 [hep-ph].

[12] J. C. Maxwell, *A treatise on electricity and magnetism*. Dover Publications, 1873.

[13] W. Pauli, *Pauli letter collection: letter to Lise Meitner*, Typed copy.

[14] C. Wu et al., *Experimental Test of Parity Conservation in Beta Decay*, Phys.Rev. **105** (1957) 1413–1414.

[15] F. L. Wilson, *Fermi's Theory of Beta Decay*, American Journal of Physics **36** (1968) 1150–1160.

[16] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys.Lett. **8** (1964) 214–215.

[17] G. Zweig, *An SU(3) model for strong interaction symmetry and its breaking. Version 1*, CERN-TH-401 (1964) .

[18] G. Zweig, *An SU(3) model for strong interaction symmetry and its breaking. Version 2*, CERN-TH-412 (1964) 22–101.

[19] J. Davis, Raymond, D. S. Harmer, and K. C. Hoffman, *Search for neutrinos from the sun*, Phys.Rev.Lett. **20** (1968) 1205–1209.

[20] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531–532.

[21] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog.Theor.Phys. **49** (1973) 652–657.

[22] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π Decay of the K20 Meson*, Phys. Rev. Lett. **13** (1964) 138–140.

[23] K. Nishijima, *Charge Independence Theory of V Particles*, Progress of Theoretical Physics **13** (1955) no. 3, 285–304.

[24] M. Gell-Mann, *The interpretation of the new particles as displaced charge multiplets*, Nuovo Cim. **4** (1956) no. S2, 848–866.

[25] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys.Rev.Lett. **13** (1964) 321–323.

[26] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys.Lett. **12** (1964) 132–133.

[27] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys.Rev.Lett. **13** (1964) 508–509.

[28] G. Guralnik, C. Hagen, and T. Kibble, *Global Conservation Laws and Massless Particles*, Phys.Rev.Lett. **13** (1964) 585–587.

[29] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys.Rev. **145** (1966) 1156–1163.

[30] T. Kibble, *Symmetry breaking in nonAbelian gauge theories*, Phys.Rev. **155** (1967) 1554–1561.

[31] LHCb Collaboration, *Observation of the resonant character of the $Z(4430)^-$ state*, Phys.Rev.Lett. **112** (2014) 222002, arXiv:1404.1903 [hep-ex].

[32] J. C. Ward, *An Identity in Quantum Electrodynamics*, Phys. Rev. **78** (1950) 182–182.

[33] O. Greenberg, *The Parton model*, arXiv:0805.2588 [hep-ph].

[34] M. Bustamante, L. Cieri, and J. Ellis, *Beyond the Standard Model for Montaneros*, arXiv:0911.4409 [hep-ph].

[35] B. Kayser, *Neutrino physics*, eConf **C040802** (2004) L004, arXiv:hep-ph/0506165 [hep-ph].

[36] R. Mohapatra et al., *Theory of neutrinos: A White paper*, Rept.Prog.Phys. **70** (2007) 1757–1867, arXiv:hep-ph/0510213 [hep-ph].

[37] M. Veltman, *Limit on Mass Differences in the Weinberg Model*, Nucl.Phys. **B123** (1977) 89.

[38] M. Baak et al., *Updated Status of the Global Electroweak Fit and Constraints on New Physics*, Eur.Phys.J. **C72** (2012) 2003, arXiv:1107.0975 [hep-ph].

[39] ATLAS Collaboration, *Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels with the ATLAS detector using 25 fb^{-1} of pp collision data*, arXiv:1406.3827 [hep-ex].

[40] CMS Collaboration, *Measurement of the properties of a Higgs boson in the four-lepton final state*, Phys.Rev. **D89** (2014) 092007, arXiv:1312.5353 [hep-ex].

[41] ATLAS Collaboration, *Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC*, Phys.Lett. **B726** (2013) 88–119, arXiv:1307.1427 [hep-ex].

[42] ATLAS Collaboration, *Updated coupling measurements of the Higgs boson with the ATLAS detector using up to 25 fb^{-1} of proton-proton collision data*, ATLAS-CONF-2014-009, CERN, Geneva, 2014.

[43] CMS Collaboration, *Measurement of Higgs boson production and properties in the WW decay channel with leptonic final states*, JHEP **1401** (2014) 096, arXiv:1312.1129 [hep-ex].

[44] ATLAS Collaboration, *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, Phys.Lett. **B726** (2013) 120–144, arXiv:1307.1432 [hep-ex].

[45] ATLAS Collaboration, *Evidence for Higgs Boson Decays to the $\tau^+\tau^-$ Final State with the ATLAS Detector*, ATLAS-CONF-2013-108, CERN, Geneva, 2013.

[46] CMS Collaboration, *Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons*, JHEP **1405** (2014) 104, arXiv:1401.5041 [hep-ex].

[47] CMS Collaboration, *Evidence for the direct decay of the 125 GeV Higgs boson to fermions*, Nature Phys. **10** (2014) , arXiv:1401.6527 [hep-ex].

[48] M. Baak et al., *The Electroweak Fit of the Standard Model after the Discovery of a New Boson at the LHC*, Eur.Phys.J. **C72** (2012) 2205, arXiv:1209.2716 [hep-ph].

[49] J. H. Schwarz and N. Seiberg, *String theory, supersymmetry, unification, and all that*, Rev.Mod.Phys. **71** (1999) S112–S120, arXiv:hep-th/9803179 [hep-th].

[50] C. Sivaram, *What is special about the Planck mass?*, arXiv:0707.0058 [gr-qc].

[51] C. Quigg, *Unanswered Questions in the Electroweak Theory*, Ann.Rev.Nucl.Part.Sci. **59** (2009) 505–555, arXiv:0905.3187 [hep-ph].

[52] WMAP Collaboration, *Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results*, Astrophys.J.Suppl. **208** (2013) 20, arXiv:1212.5225 [astro-ph.CO].

[53] A. Sakharov, *Violation of CP Invariance, c Asymmetry, and Baryon Asymmetry of the Universe*, Pisma Zh.Eksp.Teor.Fiz. **5** (1967) 32–35.

[54] *Proceedings, Workshop on Exploring QCD from the infrared regime to heavy flavour scales at B-factories, the LHC and a Linear Collider (LC13)*, Nuovo Cim. **C37** (2014) no. 02, pp.1–207.

[55] A. Riotto, *Theories of baryogenesis*, arXiv:hep-ph/9807454 [hep-ph].

[56] M. Schmaltz and D. Tucker-Smith, *Little Higgs review*, Ann.Rev.Nucl.Part.Sci. **55** (2005) 229–270, arXiv:hep-ph/0502182 [hep-ph].

[57] F. Jegerlehner, *What is triggering the Higgs mechanism and inflation?*, Nuovo Cim. **C037** (2014) no. 02, 157–163.

[58] M. Schmaltz, *Physics beyond the standard model (theory): Introducing the little Higgs*, Nucl.Phys.Proc.Suppl. **117** (2003) 40–49, arXiv:hep-ph/0210415 [hep-ph].

[59] G. F. Giudice, *Naturally Speaking: The Naturalness Criterion and Physics at the LHC*, arXiv:0801.2562 [hep-ph].

[60] G. Degrassi et al., *Higgs mass and vacuum stability in the Standard Model at NNLO*, JHEP **1208** (2012) 098, arXiv:1205.6497 [hep-ph].

[61] D. Buttazzo et al., *Investigating the near-criticality of the Higgs boson*, JHEP **1312** (2013) 089, arXiv:1307.3536.

[62] A. Spencer-Smith, *Higgs Vacuum Stability in a Mass-Dependent Renormalisation Scheme*, arXiv:1405.1975 [hep-ph].

[63] Tevatron Electroweak Working Group, CDF, D0 Collaboration, *Combination of CDF and DO results on the mass of the top quark using up to 8.7 fb^{-1} at the Tevatron*, arXiv:1305.3929 [hep-ex].

[64] CMS Collaboration, *Determination of the top-quark pole mass and strong coupling constant from the ttbar production cross section in pp collisions at sqrt(s) = 7 TeV*, Phys.Lett. **B728** (2014) 496–517, arXiv:1307.1907 [hep-ex].

[65] K. K. Boddy and S. M. Carroll, *Can the Higgs Boson Save Us From the Menace of the Boltzmann Brains?*, arXiv:1308.4686 [hep-ph].

[66] CDF Collaboration, *Search for a Heavy Top-Like Quark in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV*, Phys.Rev.Lett. **107** (2011) 261801, arXiv:1107.3875 [hep-ex].

[67] CMS Collaboration, *Search for heavy, top-like quark pair production in the dilepton final state in pp collisions at $\sqrt{s} = 7$ TeV*, Phys.Lett. **B716** (2012) 103–121, arXiv:1203.5410 [hep-ex].

[68] ALEPH, DELPHI, L3, OPAL, SLD, LEP Electroweak Working Group, SLD Electroweak Group, SLD Heavy Flavour Group Collaboration, *Precision electroweak measurements on the Z resonance*, Phys.Rept. **427** (2006) 257–454, arXiv:hep-ex/0509008 [hep-ex].

[69] B. Holdom et al., *Four Statements about the Fourth Generation*, PMC Phys. **A3** (2009) 4, arXiv:0904.4698 [hep-ph].

[70] CDF Collaboration, *Evidence for a Mass Dependent Forward-Backward Asymmetry in Top Quark Pair Production*, Phys.Rev. **D83** (2011) 112003, arXiv:1101.0034 [hep-ex].

[71] D0 Collaboration, *Measurement of the Forward-Backward Asymmetry in Top-Antitop Quark Production in Proton-Antiproton Collisions*, arXiv:1110.2062 [hep-ex].

[72] B. Bhattacherjee, S. S. Biswal, and D. Ghosh, *Top quark forward-backward asymmetry at Tevatron and its implications at the LHC*, Phys.Rev. **D83** (2011) 091501, arXiv:1102.0545 [hep-ph].

[73] O. Eberhardt et al., *Impact of a Higgs boson at a mass of 126 GeV on the standard model with three and four fermion generations*, Phys.Rev.Lett. **109** (2012) 241802, arXiv:1209.1101 [hep-ph].

[74] B. Bellazzini, C. Cski, and J. Serra, *Composite Higgses*, arXiv:1401.2457 [hep-ph].

[75] S. Gopalakrishna et al., *LHC Signatures of Warped-space Vectorlike Quarks*, arXiv:1306.2656 [hep-ph].

[76] K. Hamaguchi et al., *Isospin-Violating Dark Matter with Colored Mediators*, JHEP **1405** (2014) 086, arXiv:1403.0324 [hep-ph].

[77] W. Chao and M. J. Ramsey-Musolf, *Electroweak Baryogenesis, Electric Dipole Moments, and Higgs Diphoton Decays*, arXiv:1406.0517 [hep-ph].

[78] M.-L. Xiao and J.-H. Yu, *Stabilizing Electroweak Vacuum in a Vector-like Fermion Model*, Phys.Rev. **D90** (2014) 014007, arXiv:1404.0681 [hep-ph].

[79] D. Marzocca, M. Serone, and J. Shu, *General Composite Higgs Models*, JHEP **1208** (2012) 013, arXiv:1205.0770 [hep-ph].

[80] H. Georgi and A. Pais, *Calculability and Naturalness in Gauge Theories*, Phys.Rev. **D10** (1974) 539.

[81] H. Georgi and A. Pais, *Vacuum Symmetry and the PseudoGoldstone Phenomenon*, Phys.Rev. **D12** (1975) 508.

[82] N. Arkani-Hamed, A. G. Cohen, and H. Georgi, *Electroweak symmetry breaking from dimensional deconstruction*, Phys.Lett. **B513** (2001) 232–240, arXiv:hep-ph/0105239 [hep-ph].

[83] N. Arkani-Hamed, S. Dimopoulos, and G. Dvali, *The Hierarchy problem and new dimensions at a millimeter*, Phys.Lett. **B429** (1998) 263–272, arXiv:hep-ph/9803315 [hep-ph].

[84] L. Randall and R. Sundrum, *A Large mass hierarchy from a small extra dimension*, Phys.Rev.Lett. **83** (1999) 3370–3373, arXiv:hep-ph/9905221 [hep-ph].

[85] K. Agashe et al., *RS1, custodial isospin and precision tests*, JHEP **0308** (2003) 050, arXiv:hep-ph/0308036 [hep-ph].

[86] J. Aguilar-Saavedra et al., *Handbook of vectorlike quarks: Mixing and single production*, Phys.Rev. **D88** (2013) no. 9, 094010, arXiv:1306.0572 [hep-ph].

[87] J. Berger, J. Hubisz, and M. Perelstein, *A Fermionic Top Partner: Naturalness and the LHC*, JHEP **1207** (2012) 016, arXiv:1205.0013 [hep-ph].

[88] J. Aguilar-Saavedra, *Identifying top partners at LHC*, JHEP **0911** (2009) 030, arXiv:0907.3155 [hep-ph].

[89] S. A. Ellis et al., *Survey of vector-like fermion extensions of the Standard Model and their phenomenological implications*, arXiv:1404.4398 [hep-ph].

[90] J. Baglio, A. Djouadi, and R. Godbole, *The apparent excess in the Higgs to di-photon rate at the LHC: New Physics or QCD uncertainties?*, Phys.Lett. **B716** (2012) 203–207, arXiv:1207.1451 [hep-ph].

[91] A. Djouadi, *The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model*, Phys.Rept. **457** (2008) 1–216, arXiv:hep-ph/0503172 [hep-ph].

[92] S. Dawson and E. Furlan, *A Higgs Conundrum with Vector Fermions*, Phys.Rev. **D86** (2012) 015021, arXiv:1205.4733 [hep-ph].

[93] N. Kidonakis and B. D. Pecjak, *Top-quark production and QCD*, Eur.Phys.J. **C72** (2012) 2084, arXiv:1108.6063 [hep-ph].

[94] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[95] O. S. Bruening et al., *LHC Design Report*. CERN, Geneva, 2004.

[96] O. Bruening et al., *LHC Design Report*. CERN, Geneva, 2004.

[97] M. Benedikt et al., *LHC Design Report*. CERN, Geneva, 2004.

[98] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.

[99] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[100] CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004.

[101] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.

[102] LHCf Collaboration, *LHCf experiment: Technical Design Report*. Technical Design Report LHCf. CERN, Geneva, 2006.

[103] MoEDAL Collaboration, *Technical Design Report of the MoEDAL Experiment*, CERN-LHCC-2009-006. MoEDAL-TDR-001, CERN, Geneva, 2009.

[104] TOTEM Collaboration, *Total cross-section, elastic scattering and diffraction dissociation at the Large Hadron Collider at CERN: TOTEM Technical Design Report.* Technical Design Report TOTEM. CERN, Geneva, 2004.

[105] ATLAS, *Luminosity Public Results*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults.

[106] ATLAS Collaboration, *ATLAS detector and physics performance: Technical Design Report, 1.* Technical Design Report ATLAS. CERN, Geneva, 1999.

[107] F. Meloni et al., *Vertexing Performance Data vs MC comparison for LPCC*, ATL-COM-PHYS-2011-1312, CERN, Geneva, 2011.

[108] L. Skinnari, A. Morley, and S. Marti, *Inner Detector Alignment Performance Plots using Z-gt;mumu*, ATL-COM-PHYS-2011-865, CERN, Geneva, 2011.

[109] ATLAS Collaboration, *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data*, Eur.Phys.J. **C74** (2014) 2941, arXiv:1404.2240 [hep-ex].

[110] ATLAS, *Combined Muon Performance Public Plots*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/MuonPerformancePublicPlots.

[111] ATLAS Collaboration, *Alignment of the ATLAS Inner Detector and its Performance in 2012*, ATLAS-CONF-2014-047, CERN, Geneva, Jul, 2014.

[112] ATLAS Collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data*, Eur.Phys.J. **C74** (2014) no. 11, 3130, arXiv:1407.3935 [hep-ex].

[113] ATLAS Collaboration Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur.Phys.J. **C74** (2014) no. 10, 3071, arXiv:1407.5063 [hep-ex].

[114] S. Klous, *Event streaming in the online system*, ATL-DAQ-PROC-2010-017, CERN, Geneva, 2010.

[115] R. Brun, F. Rademakers, and S. Panacek, *ROOT, an object oriented data analysis framework*, Conf.Proc. **C000917** (2000) 11–42.

[116] M. L. Mangano, *QCD and the Physics of Hadronic Collisions*, CUSO Lecture notes, 2011.

[117] J. C. Collins and D. E. Soper, *The Theorems of Perturbative QCD*, Annual Review of Nuclear and Particle Science **37** (1987) no. 1, 383–409.

[118] J. Bjorken, *Asymptotic Sum Rules at Infinite Momentum*, Phys.Rev. **179** (1969) 1547–1553.

[119] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972) 438–450.

[120] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nuclear Physics B **126** (1977) no. 2, 298–318.

[121] Y. Dokshitzer, *Calculation of the Structure Func tions for Deep Inelastic Scattering and ee–Annihila tion by Perturbation Theory in Quantum Chromody namics*, Zh. Eksp. Teor. Fiz **73** (1977) 1216.

[122] J. Pumplin, D. Stump, J. Huston, H. Lai, P. M. Nadolsky, et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **0207** (2002) 012, arXiv:hep-ph/0201195 [hep-ph].

[123] A. Martin, W. Stirling, R. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur.Phys.J. **C63** (2009) 189–285, arXiv:0901.0002 [hep-ph].

[124] D. E. Soper, *Parton distribution functions*, vol. 53. Nucl. Phys. Proc. Suppl. 53 (1997) 69-80, 1997. arXiv:hep-lat/9609018.

[125] G. Watt and R. Thorne, *Study of Monte Carlo approach to experimental uncertainty propagation with MSTW 2008 PDFs*, JHEP **1208** (2012) 052, arXiv:1205.4024 [hep-ph].

[126] B. Webber, *Parton shower Monte Carlo event generators*, **6** (2011) no. 12, 10662. revision 128236.

[127] B. Andersson et al., *Parton Fragmentation and String Dynamics*, Phys.Rept. **97** (1983) 31–145.

[128] B. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*, Nucl.Phys. **B238** (1984) 492.

[129] D. J. Callaway and A. Rahman, *The Microcanonical Ensemble: A New Formulation of Lattice Gauge Theory*, Phys.Rev.Lett. **49** (1982) 613.

[130] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **0605** (2006) 026, arXiv:hep-ph/0603175 [hep-ph].

[131] G. Corcella et al., *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **0101** (2001) 010, arXiv:hep-ph/0011363 [hep-ph].

[132] J. Butterworth, J. R. Forshaw, and M. Seymour, *Multiparton interactions in photoproduction at HERA*, Z.Phys. **C72** (1996) 637–646, arXiv:hep-ph/9601371 [hep-ph].

[133] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **0206** (2002) 029, arXiv:hep-ph/0204244 [hep-ph].

[134] B. P. Kersevan and E. Richter-Was, *The Monte Carlo event generator AcerMC versions 2.0 to 3.8 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1*, Comput.Phys.Commun. **184** (2013) 919–985, arXiv:hep-ph/0405247 [hep-ph].

[135] M. L. Mangano et al., *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **0307** (2003) 001, arXiv:hep-ph/0206293 [hep-ph].

[136] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **0711** (2007) 070, arXiv:0709.2092 [hep-ph].

[137] GEANT4 Collaboration, *GEANT4: A Simulation toolkit*, Nucl.Instrum.Meth. **A506** (2003) 250–303.

[138] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, arXiv/1005.4568.

[139] W. Lukas, *Fast Simulation for ATLAS: Atlfast-II and ISF*, ATL-SOFT-PROC-2012-065, CERN, Geneva, 2012.

[140] S. Fleischmann, *FATRAS - A Novel Fast Track Simulation Engine for the ATLAS Experiment*, ATL-INDET-PROC-2010-010, CERN, Geneva, 2010.

[141] ATLAS Collaboration, M. Beckingham et al., *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, ATL-PHYS-PUB-2010-013, CERN, Geneva, 2010.

[142] B. Acharya et al., *Object selection and calibration, background estimations and MC samples for the Autumn 2012 Top Quark analyses with 2011 data*, ATL-COM-PHYS-2012-1197, CERN, Geneva, 2012.

[143] ATLAS Collaboration, *Characterization of Interaction-Point Beam Parameters Using the pp Event-Vertex Distribution Reconstructed in the ATLAS Detector at the LHC*, ATLAS-CONF-2010-027, CERN, Geneva, 2010.

[144] T. Cornelissen, , et al., *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, ATL-SOFT-PUB-2007-007, ATL-COM-SOFT-2007-002, CERN, Geneva, 2007.

[145] H. M. Gray, *The Charged Particle Multiplicity at Center of Mass Energies from 900 GeV to 7 TeV measured with the ATLAS Experiment at the Large Hadron Collider*. PhD thesis, Pasadena, USA, California Institute of Technology, Pasadena, USA, 2010. Presented on 09 Nov 2010.

[146] ATLAS Collaboration, *Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC*, New J.Phys. **13** (2011) 053033, arXiv:1012.5104 [hep-ex].

[147] ATLAS Collaboration, *Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s}$ =7 TeV in the ATLAS experiment*, ATLAS-CONF-2010-069, CERN, Geneva, 2010.

[148] W. Waltenberger et al., *Adaptive vertex fitting*, Journal of Physics G: Nuclear and Particle Physics **34** (2007) no. 12, N343.

[149] ATLAS, *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, The European Physical Journal C **72** (2012) no. 3, 1–46.

[150] ATLAS Collaboration, *Electrons Energy and Direction*, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ElectronsEnergyDirection.

[151] ATLAS Collaboration, *IsEM Identification*, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/IsEMIdentification.

[152] J. Hartert and I. Ludwig, *Electron isolation in the ATLAS experiment*, ATL-PHYS-INT-2010-052, CERN, Geneva, 2010.

[153] ATLAS Collaboration, *Muon Reconstruction Performance*, ATLAS-CONF-2010-064, CERN, Geneva, 2010.

[154] ATLAS Collaboration, *Muon reconstruction efficiency and momentum resolution of the ATLAS experiment in proton-proton collisions at $\sqrt{s}$=7 TeV in 2010*, arXiv:1404.4562 [hep-ex].

[155] O. Kortner and E. Moyse, *Pile-up Depedence of the ATLAS Muon Performance*, ATL-COM-PHYS-2011-1640, CERN, Geneva, 2011.

[156] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, Phys.Lett. **B641** (2006) 57–61, arXiv:hep-ph/0512210 [hep-ph].

[157] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **0804** (2008) 063, arXiv:0802.1189 [hep-ph].

[158] ATLAS Collaboration, *Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, arXiv:1406.0076 [hep-ex].

[159] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV*, Eur.Phys.J. **C73** (2013) 2304, arXiv:1112.6426 [hep-ex].

[160] ATLAS Collaboration, *Jet energy scale and its systematic uncertainty in proton-proton collisions at sqrt(s)=7 TeV with ATLAS 2011 data*, ATLAS-CONF-2013-004, CERN, Geneva, 2013.

[161] ATLAS Collaboration, *Performance of Missing Transverse Momentum Reconstruction in Proton-Proton Collisions at 7 TeV with ATLAS*, Eur.Phys.J. **C72** (2012) 1844, arXiv:1108.5602 [hep-ex].

[162] ATLAS Collaboration, *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*, ATLAS-CONF-2011-102, CERN, Geneva, 2011.

[163] *Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data*, ATLAS-CONF-2014-046, CERN, Geneva, 2014.

[164] ATLAS, *B-Tagging Benchmarks*, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/BTaggingBenchmarks.

[165] ATLAS Collaboration, *Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons with 5 fb-1 of Data from the ATLAS Detector*, ATLAS-CONF-2012-043, CERN, Geneva, 2012.

[166] ATLAS Collaboration, *b-jet tagging calibration on c-jets containing D^{*+} mesons*, ATLAS-CONF-2012-039, CERN, Geneva, 2012.

[167] ATLAS Collaboration, *Measurement of the Mistag Rate with 5 fb^{-1} of Data Collected by the ATLAS Detector*, ATLAS-CONF-2012-040, CERN, Geneva, 2012.

[168] J. Erdmann et al., *A likelihood-based reconstruction algorithm for top-quark pairs and the KLFitter framework*, arXiv:1312.5595 [hep-ex].

[169] A. Caldwell, D. Kollar, and K. Kroninger, *BAT: The Bayesian Analysis Toolkit*, Comput.Phys.Commun. **180** (2009) 2197–2209, arXiv:0808.2552 [physics.data-an].

[170] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, *Introducing markov chain monte carlo*, Markov chain Monte Carlo in practice (1996) 1–19.

[171] C. Tsallis and D. A. Stariolo, *Generalized simulated annealing*, Physica A: Statistical Mechanics and its Applications **233** (1996) no. 1, 395–406.

[172] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput.Phys.Commun. **10** (1975) 343–367.

[173] A. Hoecker et al., *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039.

[174] E. Bishop, *A generalization of the Stone-Weierstrass theorem.*, Pacific Journal of Mathematics **11** (1961) no. 3, 777–783.

[175] D. Rumelhart, G. Hintont, and R. Williams, *Learning representations by back-propagating errors*, Nature **323** (1986) no. 6088, 533–536.

[176] C. G. Broyden, *The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations*, IMA Journal of Applied Mathematics **6** (1970) no. 1, 76–90.

[177] R. Fletcher, *A new approach to variable metric algorithms*, The Computer Journal **13** (1970) no. 3, 317–322.

[178] D. Goldfarb, *A family of variable-metric methods derived by variational means*, Mathematics of computation **24** (1970) no. 109, 23–26.

[179] D. F. Shanno, *Conditioning of quasi-Newton methods for function minimization*, Mathematics of computation **24** (1970) no. 111, 647–656.

[180] S. J. Russell and P. Norvig, *Artificial intelligence: a modern approach (International Edition)*, .

[181] T. Junk, *Confidence level computation for combining searches with small statistics*, Nucl.Instrum.Meth. **A434** (1999) 435–443, arXiv:hep-ex/9902006 [hep-ex].

[182] CDF Collaboration, *Search for Standard Model Higgs Boson Production in Association with a W Boson Using a Matrix Element Technique at CDF in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV*, Phys.Rev. **D85** (2012) 072001, arXiv:1112.4358 [hep-ex].

[183] ATLAS Collaboration, *LAr Event Veto Rel17 Twiki*, https://twiki.cern.ch/twiki/bin/viewauth/AtlasComputing/LArEventVetoRel17.

[184] ATLAS, *Standard Model Results*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/StandardModelPublicResults.

[185] CDF Collaboration, *Observation of Single Top Quark Production and Measurement of —Vtb— with CDF*, Phys.Rev. **D82** (2010) 112005, arXiv:1004.1181 [hep-ex].

[186] ATLAS Collaboration, *ATLAS Computing: technical design report.* Technical Design Report ATLAS. CERN, Geneva, 2005.

[187] ATLAS Collaboration, *Luminosity Determination in pp Collisions at sqrt(s) = 7 TeV using the ATLAS Detector in 2011*, ATLAS-CONF-2011-116, CERN, Geneva, 2011.

[188] *ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11*, ATL-PHYS-PUB-2011-009, CERN, Geneva, 2011.

[189] CTEQ Collaboration, *The PYTHIA6 - tune AUET2B-CTEQ6L1 web page*, http://rivet.hepforge.org/tunecmp/PYTHIA6/AUET2B-CTEQ6L1/index.html.

[190] A. Sherstnev and R. Thorne, *Different PDF approximations useful for LO Monte Carlo generators*, arXiv:0807.2132 [hep-ph].

[191] H. Lacker and A. Menzel, *Simultaneous Extraction of the Fermi constant and PMNS matrix elements in the presence of a fourth generation*, JHEP **1007** (2010) 006, arXiv:1003.4532 [hep-ph].

[192] M. Aliev et al., *HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR*, Comput.Phys.Commun. **182** (2011) 1034–1046, arXiv:1007.1327 [hep-ph].

[193] P. M. Nadolsky et al., *Implications of CTEQ global analysis for collider observables*, Phys.Rev. **D78** (2008) 013004, arXiv:0802.0007 [hep-ph].

[194] J. Ferrando and D. Wendland, *Reference $t\bar{t}$ production cross sections for use in ATLAS analyses*, ATL-COM-PHYS-2014-112, CERN, Geneva, 2014.

[195] M. Botje et al., *The PDF4LHC Working Group Interim Recommendations*, arXiv:1101.0538 [hep-ph].

[196] H.-L. Lai et al., *New parton distributions for collider physics*, Phys.Rev. **D82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[197] *New ATLAS event generator tunes to 2010 data*, ATL-PHYS-PUB-2011-008, CERN, Geneva, 2011.

[198] *First tuning of HERWIG/JIMMY to ATLAS data*, ATL-PHYS-PUB-2010-014, CERN, Geneva, 2010.

[199] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, Phys.Rev. **D83** (2011) 091503, arXiv:1103.2792 [hep-ph].

[200] N. Kidonakis, *NNLL resummation for s-channel single top quark production*, Phys.Rev. **D81** (2010) 054028, arXiv:1001.5034 [hep-ph].

[201] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-*, Phys.Rev. **D82** (2010) 054018, arXiv:1005.4451 [hep-ph].

[202] D. Stump et al., *Inclusive jet production, parton distributions, and the search for new physics*, JHEP **0310** (2003) 046, arXiv:hep-ph/0303013 [hep-ph].

[203] ATLAS Collaboration, *Heavy Flavor Overlap Removal Tool*, https://twiki.cern.ch/twiki/bin/view/Main/HforToolD3PD.

[204] J. Butterworth et al., *Single Boson and Diboson Production Cross Sections in pp Collisions at sqrts=7 TeV*, ATL-COM-PHYS-2010-695, CERN, Geneva, 2010.

[205] ATLAS Collaboration, *Search for pair production of a heavy up-type quark decaying to a W boson and a b quark in the lepton+jets channel with the ATLAS detector*, Phys.Rev.Lett. **108** (2012) 261802, arXiv:1202.3076 [hep-ex].

[206] ATLAS Collaboration, *Search for pair production of heavy top-like quarks decaying to a high-pT W boson and a b quark in the lepton plus jets final state at $\sqrt{s}$=7 TeV with the ATLAS detector*, Phys.Lett. **B718** (2013) 1284–1302, arXiv:1210.5468 [hep-ex].

[207] ATLAS Collaboration, *Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC*, Eur.Phys.J. **C71** (2011) 1630, arXiv:1101.2185 [hep-ex].

[208] N. Benekos et al., *Lepton trigger and identification for the Winter 2011 top quark analyses*, ATL-COM-PHYS-2011-123, CERN, Geneva, 2011. Supporting document for Winter 2011 top physics measurements.

[209] H. Bawa et al., *Searches for Quark Contact Interactions and Other New Physics in Dijet Angular Distributions in pp Collisions at sqrt(s) = 7 TeV Measured with the ATLAS Detector at the LHC*, ATL-COM-PHYS-2011-032, CERN, Geneva, 2011.

[210] D0 Collaboration, *High-p_T jets in $\bar{p}p$ collisions at $\sqrt{s} = 630$ GeV and 1800 GeV*, Phys.Rev. **D64** (2001) 032003, arXiv:hep-ex/0012046 [hep-ex].

[211] UA2, Bern-CERN-Copenhagen-Orsay-Pavia-Saclay Collaboration Collaboration, *Measurement of Jet Production Properties at the CERN anti-p p Collider*, Phys.Lett. **B144** (1984) 283.

[212] ATLAS Collaboration, *Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector*, Eur.Phys.J. **C73** (2013) 2306, arXiv:1210.6210 [hep-ex].

[213] ATLAS Collaboration, *JERProvider Twiki*, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JetEnergyResolutionProviderForPhysicsAnalysis.

[214] R. Alon et al., *Backup Note for Jet Mass and Substructure in QCD Jets with the ATLAS Experiment*, ATL-COM-PHYS-2011-1319, CERN, Geneva, 2011.

[215] ATLAS Collaboration, *Calibration Data Interface Twiki*, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/BTaggingCalibrationDataInterface.

[216] ATLAS Collaboration, *Measurement of the top quark-pair production cross section with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV*, Eur.Phys.J. **C71** (2011) 1577, arXiv:1012.1792 [hep-ex].

[217] ATLAS Collaboration, *Search for heavy top-like quarks decaying to a Higgs boson and a top quark in the lepton plus jets final state in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, ATLAS-CONF-2013-018, CERN, Geneva, 2013.

[218] ATLAS Collaboration, *Search for pair production of heavy top-like quarks decaying to a high-p_{T} W boson and a b quark in the lepton plus jets final state in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, ATLAS-CONF-2013-060, CERN, Geneva, 2013.

[219] ATLAS Collaboration, *Search for pair and single production of new heavy quarks that decay to a Z boson and a third generation quark in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, ATLAS-CONF-2014-036, CERN, Geneva, 2014.

[220] ATLAS Collaboration, *Search for anomalous production of events with same-sign dileptons and b jets in 14.3 fb^{-1} of pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, ATLAS-CONF-2013-051, CERN, Geneva, 2013. Not published in the proceedings.

[221] ATLAS Collaboration, *Public Exotics Results*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ExoticsPublicResults.

[222] P. Baldi, P. Sadowski, and D. Whiteson, *Searching for Exotic Particles in High-Energy Physics with Deep Learning*, arXiv:1402.4735 [hep-ph].

[223] M. Feindt, *A neural Bayesian estimator for conditional probability densities*, arXiv preprint physics/0402093 (2004) .

[224] B. P. Roe et al., *Boosted decision trees as an alternative to artificial neural networks for particle identification*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **543** (2005) no. 2, 577–584.

Acknowledgements

At the end of this long journey, I owe a big thank you to all these people who's support made it happen and become a memorable part of my life.

First of all, I am deeply grateful to my supervisor, Prof. Allan Clark, for giving me the opportunity to carry out my doctoral education at the University of Geneva, and for following my work since its beginning until its completion. I am especially thankful for his kind support in the two last years of my PhD, when he remained my only supervisor on spot (despite his retirement!), while I was completing the core of my thesis work. Thank you, Allan.

The person who definitely left the strongest mark on my PhD education, by steering its directions and following every step closely, and yet leaving me space to explore myself and harvest the joy of "discovering" something new, is my then postdoc, Prof. Alison Lister. I want to sincerely thank her for guiding my way through this long journey with her competent decisions and valuable discussions, sharpening my exactness by capturing any loose corner, and above all, for boosting my moral in the difficult moments by her positive attitude! And after all, for her sincere and enthusiastic personal involvement, with an always present dose of humour, which turned a hard work into a fun thing to do. Thank you so much!

I'd further like to thank my partial supervisors: Jose (aka Hombre), for supervising my qualification work (in addition to helping with any kind of annoying computing problems!), and Will, for providing the useful computing support as well as interesting discussions throughout my work. On the analysis side, I would especially like to thank my analysis coordinator Marc, for following and supporting my work, and providing very useful inputs. I also owe a very special thank you to my friend Ennio, for the priceless theoretical explanations and answers to my numerous questions, especially those during the preparation of my defence, and for doing all of that in the language understandable for an experimentalist.

Finally, I'd like to thank my co-supervisor, Prof. Xin WU, and the members of my defence committee, Prof. Giora Mikenberg and Prof. Giuseppe Iacobucci, for reviewing my thesis, as well as for their useful feedback and enjoyable discussions.

I also owe a big thank you to people who worked hard to ensure a smooth working environment to me and everyone else: Szymon and Yann, who took care of our computing facilities, and Catherine, who took care od all administrative matters. Thanks a lot! Also, big thanks to Jose, Thijs and Valerio, who pre-reviewd some chapters of my thesis. And of course, people without who's support my french resumée would have been completely unreadable: Moritz, Andrée and Gaetano. Thanks!

And after all, I have to express my deepest gratitude to my only "supervisor" who wasn't formally entitled to be it, but still was (and continues to be) the most stringent scrutiniser

and the most passionate consultant of my work: my dear friend Valerio, who was always there to provide his unconditioned support (both in physics and in the life), and bring in an avalanche of good ideas, even on his own initiative, when he wasn't asked for it, but sensed that I was struggling with some problem. I really wouldn't have made it without you!

My PhD wouldn't have been as complete and enriching without all the amazing people I was surrounded with! I really enjoyed the highly competent and very friendly environment within the ATLAS Geneva group. I learned a lot about physics in our group meetings, and about life in our casual moments. Above all, I really felt home among my dear colleagues, with whom I shared a lot of memorable moments through all these years, be it in our lunch breaks, after-meeting coffee breaks, parties, barbecue or skiing outings, or just on the corridors. For all that, I would like to say a big thank you to my dear colleagues: Elina, Jose, Francesca, Clemencia, Gaetano, Sergio, Gabi, Andrée, Andrew, Gauthier, Kilian, Eleonora, Kati, Francesco, Attilio, Johanna, Akshay, Chris, Reina, Philippe, Caterina, Ian, Paul, Lucian, Antonello and Javier. And of course, an important part of my well being and feeling home at CERN is definitely thank to my current and ex compatriots: Pedja, Jelena, Nina, Lili, Judita, Adriana, Sanja, Vladimir, Marija, Nenad and the others, who provided that warm taste of homeland. Thank you all!

And in the end, I owe an infinite gratitude to the people who's presence in my life gives it all a meaning. Vast thanks to Moritz, for being my truthful company on this journey, and my deepest gratitude to my family: my father Radoslav, my mother Ljiljana, and my sister Marija, for always giving their best to free my way for growing as a professional and, more importantly, as a human. Hvala vam od sveg srca!