Measurement of the Top Quark Pair Production Cross Section in the Muon-Electron Decay Channel at $\sqrt{s} = 7$ TeV with the CMS Experiment

Dissertation
zur Erlangung des Doktorgrades
des Fachbereichs Physik
der Universität Hamburg

vorgelegt von
Diplom-Physiker Markus Marienfeld
aus Hameln

Hamburg
2011

Gutachter der Dissertation: Prof. Dr. Joachim Mnich
PD Dr. Achim Geiser

Gutachter der Disputation: Prof. Dr. Joachim Mnich
Prof. Dr. Peter Schleper

Datum der Disputation: 11. Juli 2011

Vorsitzender des Prüfungsauschusses: Prof. Dr. Dieter Horns
Vorsitzender des Promotionsauschusses: Prof. Dr. Peter H. Hauschildt

Leiterin des Fachbereichs Physik: Prof. Dr. Daniela Pfannkuche
Dekan der MIN-Fakultät: Prof. Dr. Heinrich Graener

Für meinen Großvater Alfred

Aspire to climb as high as you can dream.

Abstract

The start of proton-proton collisions at the LHC inaugurates a new era in high-energy physics. It enables the possibility of discoveries at the high-energy frontier and also allows for studies of known Standard Model processes with unrivalled precision. Top quark pairs are produced at high rates and allow for precision measurements of the properties of the top quark with high statistics.
The measurement of the top quark pair production cross section in proton-proton collisions at $\sqrt{s} = 7$ TeV is presented using the dileptonic decay channel with a muon-electron pair in the final state. The data sample, which is used in this analysis, corresponds the complete 2010 data taking period with an integrated luminosity of 35.9 pb^{-1}. Top quark pair candidate events are selected in a cut-based event selection. Based on 59 observed muon-electron events in the final state event sample, the top quark pair production cross section is measured to be $\sigma_{\mathrm{t\bar{t}}} = (156 \pm 25(\mathrm{stat.}) \pm 14(\mathrm{sys.}))$ pb. Furthermore, a kinematic event reconstruction is applied, which is complementary to the use of b-tagging techniques, and validates the top quark-like topology of the selected events.
First results from the measurement of differential cross sections based on the data from the complete 2010 data taking period are presented. For the first time in the CMS collaboration, the cross section of the production of top quark pairs is measured differentially as a function of the kinematic observables of the final state objects, such as the transverse momentum p_{T} of the leptons and the invariant mass of the lepton pair. Based on the solution of the kinematic event reconstruction, the cross section is also calculated differentially as a function of the kinematic properties of the reconstructed top-antitop quark pair. First results from the measurement of differential cross sections as a function of the kinematics of the final state leptons are presented, using the data recorded in the first part of the 2011 data taking period.

Kurzfassung

Der Beginn von Proton-Proton-Kollisionen am LHC hat eine neue Ära in der Hochenergiephysik eingeleitet. Es ermöglicht Entdeckungen an der Grenze der Hochenergiephysik, sowie Messungen von bereits bekannten Standardmodellprozessen mit bisher unerreichter Präzision. Top-Quark-Paare werden mit hoher Rate erzeugt und ermöglichen präzise Messungen der Eigenschaften des Top-Quarks mit hoher Statistik.

Es wird die Messung des Wirkungsquerschnitts von Top-Quark-Paaren in Proton-Proton-Kollisionen bei $\sqrt{s} = 7$ TeV präsentiert. Dabei wird der dileptonische Zerfallskanal in ein Myon-Elektron-Paar im Endzustand verwendet. In dieser Analyse wird ein Datensatz verwendet, welcher der vollständigen Datennahmeperiode in 2010 mit einer intergrierten Luminosität von 35.9 pb^{-1} entspricht. Kandidatenereignisse zerfallener Top-Quark-Paare werden in einer schnitt-basierten Analyse selektiert. Der gemessene Wirkungsquerschnitt für die Produktion von Top-Quark-Paaren beträgt $\sigma_{\mathrm{t\bar{t}}} = (156 \pm 25(\mathrm{stat.}) \pm 14(\mathrm{sys.}))$ pb, welcher auf 59 selektierten Signalereignissen im Endzustand mit einem Myon-Elektron-Paar basiert. Des Weiteren wird eine kinematische Ereignisrekonstruktion angewendet, welche komplementär zu der Anwendung von b-tagging Algorithmen behandelt wird, und die Top-Quark-artige Topologie der selektierten Ereignisse verifiziert.

Es werden erste Resultate der Messung differentieller Wirkungsquerschnitte basierend auf den Daten der vollständigen Datennahmeperiode in 2010 präsentiert. Zum ersten mal wird in der CMS-Kollaboration der Wirkungsquerschnitt der Produktion von Top-Quark-Paaren differentiell als Funktion der kinematischen Observablen der Objekte im Endzustand gemessen. Die Lösung der kinematischen Ereignisrekonstruktion wird verwendet, um den Wirkungsquerschnitt differentiell als Funktion der rekonstruierten Eigenschaften des Top-Antitop-Quark-Paares zu berechnen. Es werden ebenfalls erste Resultate der Messungen des differentiellen Wirkungsquerschnitts basierend auf den ersten Daten der Datennahmepriode in 2011 als Funktionen der kinematischen Observablen der Leptonen im Endzustand gezeigt.

Contents

Chapter 1

Introduction

With the start of proton-proton collisions at the LHC, a new era of high-energy physics has been inaugurated. It enables the possibility of discoveries at the high-energy frontier and also allows for studies of rare Standard Model processes with unrivalled precision, such as the production of top quark pairs. Top quark pairs are produced at high rates and allow for precision measurements of the properties of the top quark with high statistics. The measurement of the top quark pair production cross section in proton-proton collisions at $\sqrt{s} = 7$ TeV is presented using the dileptonic decay channel into a muon-electron pair in the final state. The data sample, which is used in this analysis, corresponds to the complete 2010 data taking period with an integrated luminosity of 35.9 pb^{-1}. Top quark pair candidate events are selected in a robust, cut-based event selection on an electron- or muon-triggered data sample by requiring one isolated, high-energetic electron and muon of opposite charge, as well as at least two high-energetic jets.

The inclusive top quark pair production cross section is determined from the number of selected signal events and is corrected for the finite experimental efficiencies, such as reconstruction efficiency and detector acceptance. Several correction factors are applied to the total selection efficiency, which is derived from simulated top quark pair events. The single-muon trigger efficiency is measured directly from the data using the tag&probe method. The remaining selection efficiencies are also corrected for the values, which have also been derived from the tag&probe method. An estimation of the remaining background contributions from processes with two real, unlike-flavoured leptons is obtained from simulation, as well as processes with at least one misidentified muon or electron. In order to validate the top quark-like topology of the selected events, a kinematic reconstruction algorithm is applied to the final event sample. The method is complementary to the use of b jet identification methods. As an alternative, also the efficiency for requiring at least one b-tagged jet is investigated on the selected signal sample after requiring at least two high-p_T jets on top of the dilepton selection. Two different b-tagging discriminators are studied and the corresponding b-tagging efficiencies are calculated.

In chapter 2, a brief introduction to the theoretical foundation of the Standard Model of particle physics is given. It is followed by a description of the production and decay mechanisms of top quark pairs, as well as an overview of the properties of the top quark, which is given in chapter 3. Chapter 4 comprises a description of the experimental setup, containing an introduction to the Large Hadron Collider and the Compact Muon Solenoid. The simulation of events produced in proton-proton collisions and the event reconstruction, which is performed both on simulated events, as well as on the real detector output, is described in chapter 5. The correct theoretical understanding of the signal production process and all relevant background processes is a key ingredient to the interpretation of the result of the measurement. The selection of signal candidate events is performed by applying several sequential selection steps on the reconstructed objects, which is described in chapter 6.

The measurement of the **inclusive top quark pair production cross section** is described in chapter 7. Also a discussion of all relevant experimental and theoretical uncertainties is given there. In addition to the standard selection, the cross section is also calculated for the number of events, in which a kinematic solution for the reconstructed top quark pair events is found. A distribution for the single top quark mass is presented, which is found by the kinematic reconstruction of the top quark pair system. Additionally, the cross section is measured in events, which have at least one jet identified as b jet by two different b-tagging algorithms. Also the efficiency of the requirement of b-tagging is studied on the top quark signal-enriched event sample.

The first results from the measurement of **differential cross sections** based on the data from the complete 2010 data taking periods are presented. For the first time in CMS, the cross section of the production of top quark pairs is measured differentially as a function of the kinematic observables of the final state objects, such as the transverse momentum p_{T} of the muon and the electron, as well as the invariant mass of the lepton pair. Based on the solution of the kinematic event reconstruction, the cross section is also calculated differentially as a function of the kinematic properties of the reconstructed top-antitop quark pair. Finally, an outlook of the results from the analysis of the data recorded in the first part of 2011 is given in chapter 8. First results from the measurement of differential cross sections are presented.

Chapter 2

Standard Model of Particle Physics

I am a physicist. I have a working knowledge of the entire universe and everything it contains.
Sheldon, Big Bang Theory

2.1 Overview

The Standard Model (SM) of particle physics is a complete description of the fundamental building blocks of matter (elementary particles) and the interactions between them. Mathematically, it is described by a set of renormalisable, relativistic quantum field theories. The fundamental interactions (forces) are electromagnetism and weak interaction, which are described by the unified electroweak theory, as well as the strong interaction and gravitation. Since gravitation has not yet been successfully formulated in terms of a quantum field theory, it is not part of the SM. The electroweak and strong interactions are based on fundamental symmetries in nature, which are denoted as *gauge symmetries*. These gauge interactions are represented by bosonic quantum fields (gauge bosons) of spin 1. The elementary particles are represented by fermionic quantum fields of spin $\frac{1}{2}$ (fermions). They are separated into leptons and quarks. Leptons are only affected by electroweak interactions, while quarks also participate in strong interactions. Furthermore, quarks are the constituents of the proton and neutron, which constitute the atomic nucleus. A complete overview of the current content of particles and interactions of the SM, as well as a selection of their properties, is given in table 2.1.

2.2 Quarks and Leptons

The fundamental buildings blocks of the matter in the known universe can be described by fermionic quantum fields. There are two different types of fermions: the **quarks** and the **leptons**. They are grouped into *families*, which each contain a pair of a charged lepton and a neutrino, as well as a pair of an *up*- and *down*-type quark:

$$\text{Leptons:} \begin{pmatrix} e \\ \nu_e \end{pmatrix} \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix} \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix} \quad \text{Quarks:} \begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix} . \tag{2.1}$$

There are three copies, referred to as **generations**, of one fermion family. The fermion families are identical, apart from the masses of the fermions. The fundamental fermion

Fermion Generation	I	II	III
Leptons	m = 0.511 MeV Q = -1 s = $\frac{1}{2}$ e^- Electron	m = 105.7 MeV Q = -1 s = $\frac{1}{2}$ μ^- Muon	m = 1776.8 MeV Q = -1 s = $\frac{1}{2}$ τ^- Tau
	m < 2.2 eV Q = 0 s = $\frac{1}{2}$ ν_e Electron neutrino	m < 0.17 MeV Q = 0 s = $\frac{1}{2}$ ν_μ Muon neutrino	m < 18.2 MeV Q = 0 s = $\frac{1}{2}$ ν_τ Tau neutrino
Quarks	m = 2.5 MeV Q = $\frac{2}{3}$ s = $\frac{1}{2}$ u Up	m = 1.27 GeV Q = $\frac{2}{3}$ s = $\frac{1}{2}$ c Charm	m = 173.1 GeV Q = $\frac{2}{3}$ s = $\frac{1}{2}$ t Top
	m = 4.9 MeV Q = -$\frac{1}{3}$ s = $\frac{1}{2}$ d Down	m = 101 MeV Q = -$\frac{1}{3}$ s = $\frac{1}{2}$ s Strange	m = 4.2 GeV Q = -$\frac{1}{3}$ s = $\frac{1}{2}$ b Bottom
Gauge Group	$U(1)_Y$	$SU(2)_L$	$SU(3)_C$
Gauge Bosons	massless Q = 0 s = 1 γ Photon	m = 91.2 GeV Q = 0 s = 1 Z^0 Z^0 boson	massless Q = 0 s = 1 g Gluon
		m = 80.4 GeV Q = ±1 s = 1 $W^\pm$ $W^\pm$ boson	

Table 2.1: Overview of the current content of elementary particles and interactions of the Standard Model (SM). For each particle, three selected properties are given: the mass m, electric charge Q and the spin s.

fields are represented by Dirac spinor fields $\psi(x)$. An overview of the three fermion generations is given in table 2.1. For each fermion, there also exists an **antifermion** of identical mass and opposite quantum numbers.

Quarks cannot be observed as free particles in experiments. They are bound into states of two and three quarks, called *mesons* and *baryons*. This effect will be discussed in further detail in the context of quantum chromodynamics in section 2.3.3. As an example, the up- and down-type quarks of the first fermion generation (called *up* and *down* quark) constitute the proton and the neutron, which are the building blocks of the atomic nucleus. Together with the electron, the constituents of the first fermion generation form all stable matter.

2.3 Interactions and Gauge Invariance

2.3.1 Principle of Gauge Invariance

The underlying principle of the interactions between the elementary particles are fundamental *symmetries* in nature. The equations of motion have to remain identical under a local transformation of the fundamental fields:

$$\psi(x) \to V(x)\psi(x)\,, \tag{2.2}$$

where $V(x)$ is unitary. In that case, the Lagrangian of a given theory is referred to as **locally invariant** under a certain continuous group of symmetries $\mathcal{G}$. The group element $V \in \mathcal{G}$ can be expressed in terms of the basic **generators** of the symmetry group as an infinitesimal transformation:

$$V(x) = e^{i\alpha^{\mathrm{a}}(x)T^{\mathrm{a}}} \approx 1 + i\alpha^{\mathrm{a}}(x)T^{\mathrm{a}} + \mathcal{O}(\alpha^2)\,, \tag{2.3}$$

where T^{a} are the generators of the symmetry group $\mathcal{G}$. The generators fulfil the following commutation relation, known as **Lie algebra**:

$$[T^{\mathrm{a}}, T^{\mathrm{b}}] = i f^{\mathrm{abc}} T^{\mathrm{c}}\,, \tag{2.4}$$

where the numbers f^{abc} are called *structure constants*. The corresponding symmetry group of the Lie algebra is called *Lie group*. In the course of the discussion of the fundamental gauge interactions, it is shown below, how the gauge fields are derived from the fundamental symmetry groups. A brief overview of the gauge symmetries of the SM is given below, before they are discussed in further detail in the following subsections.

Gauge Symmetries of the Standard Model

All fundamental interactions between elementary particle fields in the SM are described by local gauge symmetries. They give rise to the **gauge bosons**, which mediate the interactions between the particles. The number of gauge bosons is derived from the number of generators T^{a} of the corresponding symmetry group $\mathcal{G}$. The **gauge group of the Standard Model** is composed of the following three subgroups:

$$U(1) \;\times\; SU(2) \;\times\; SU(3)\;. \tag{2.5}$$

U(1) describes the **electromagnetic interaction**. It has exactly one generator and the corresponding gauge boson is referred to as the *photon* γ, which is massless and electrically neutral. The symmetry transformation of U(1) is a rotation by a complex phase $\mathrm{e}^{i\alpha(x)}$. U(1) is the only Abelian (commutative) gauge group of the SM. A complete theoretical description of the quantum field theory of electromagnetic interactions, called *quantum electrodynamics* (QED), is described in [1] and [2] and will be commented in more detail in the next section.
SU(2) describes the **weak interaction** via a *non-Abelian* gauge theory, called *Yang-Mills theory*. Unlike in the case of U(1), the gauge bosons of the weak interaction, called the W^+, W^- and Z^0 bosons, carry field charge and interact with each other. Self-coupling is a major feature of non-Abelian (non-commutative) gauge fields. The W and Z bosons

are found to be massive, which requires the introduction of the mechanism of *spontaneous symmetry breaking*, which will also be discussed in the next section.
SU(3) describes the **strong interaction**. Only the quarks interact via the strong coupling. Like in the case of the weak interaction, the gauge bosons of SU(3) (called *gluons*) interact with each other. There are eight different types of gluons, which are considered to be massless, unlike in the case of the SU(2) gauge bosons. There are three field charges of the strong interaction, which are referred to as *colours*: *red*, *green* and *blue*. The coupling of gluons to each other is an important consequence of the non-Abelian nature of the fundamental gauge symmetry and gives rise to effects like the *confinement* of the quarks and the scale dependence of the *strong coupling constant* α_{s}.

2.3.2 Unified Electromagnetic and Weak Interaction

Quantum Electrodynamics (QED)

The mathematically simplest gauge group of the Standard Model U(1) gives rise to the theoretically best understood quantum field theory realised in nature: the quantum field theory of electrodynamics, called *quantum electrodynamics* (QED). U(1) has exactly one generator and therefore one gauge boson field exists. The **Lagrangian of QED** is given by equation 2.6:

$$\mathcal{L}_{\mathrm{QED}} \;=\; i\bar{\psi}(x)\gamma^{\mu}D_{\mu}(x)\psi(x) - m\bar{\psi}(x)\psi(x) - \frac{1}{4}\left(F_{\mu\nu}\right)^{2} \;, \tag{2.6}$$

where $D_{\mu}(x) \equiv \partial_{\mu} + ieA_{\mu}(x)$ is the *gauge-covariant derivative*, which represents the interaction between charged particles. The fermionic matter fields $\psi(x)$, which carry the electric charge e, couple to the gauge field $A_{\mu}(x)$, referred to as the **photon field**. If the definition of the covariant derivative $D_{\mu}(x)$ is compared to equation 2.3, it follows, how the photon field $A_{\mu}(x)$ is derived from the generators T^{a} of the symmetry transformation of the gauge group: $A_{\mu}(x) \propto \partial_{\mu}(\alpha^{\mathrm{a}}(x)T^{\mathrm{a}})$. The gauge field A_{μ} itself can be interpreted as the manifestation of the fundamental symmetry U(1). The electromagnetic field tensor $F_{\mu\nu}$ represents the kinematic terms of the electromagnetic vector field and therefore represents the relativistic form of **Maxwells equations**:

$$F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu} \quad \Longrightarrow \quad \epsilon^{\mu\nu\rho\sigma}\partial_{\nu}F_{\rho\sigma} = 0 \quad \wedge \quad \partial_{\mu}F^{\mu\nu} = ej^{\nu} \;, \tag{2.7}$$

where $j^{\nu} = (\rho, \overrightarrow{j})$ contains the *charge density* ρ and the *electric current* $\overrightarrow{j}$, which are the sources of electromagnetic field.

Glashow-Weinberg-Salam Theory of Electroweak Interactions

The Glashow-Weinberg-Salam theory (*GWS* theory) describes both the electromagnetic and the weak interaction as a $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ gauge theory, which is called the *electroweak* interaction. The theory is based on the non-Abelian $\mathrm{SU}(2)_L$ symmetry of *weak-isospin* $\overrightarrow{T}$ and the $\mathrm{U}(1)_Y$ symmetry of *hypercharge* Y. Therein, $\mathrm{SU}(2)_L$ denotes the coupling of charged, weak currents to the "*left-handed*" component of fermion fields, which is described below. In this description, all electroweak gauge bosons are generated massless, like in the theory of QED. Experimentally, the gauge bosons of the weak interaction, the

$W^\pm$ and Z^0 bosons, are found to be massive [3]. Breaking the $\mathrm{SU}(2)_L$ symmetry by a *complex scalar field* $\phi(x)$ leads to a system, which contains massive gauge bosons, which is referred to as spontaneous symmetry breaking or the **Higgs mechanism**, which is explained in more detail in the next section. The Lagrangian of a Yang-Mills theory with self-coupling vector boson fields is given by the following equation:

$$\mathcal{L}_{\text{Yang-Mills}} = \overline{\psi}(x)(i\gamma^\mu D_\mu(x))\psi(x) - \frac{1}{4}(F^{\mathrm{a}}_{\mu\nu})^2 - m\overline{\psi}(x)\psi(x)\,. \tag{2.8}$$

Its structure is similar to the QED Lagrangian of $\mathrm{U}(1)_Y$. In case of $\mathrm{SU}(2)_L$, there are three fundamental vector fields W^1_μ, W^2_μ and W^3_μ. For $\mathrm{U}(1)_Y$, there is one fundamental vector field B^0_μ. In case of an **unified gauge symmetry** of electroweak interaction, the covariant derivative is defined as:

$$\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y \implies D^{\mathrm{a}}_\mu \equiv \partial_\mu - igW^{\mathrm{a}}_\mu T^{\mathrm{a}} - ig'YB^0_\mu\ . \tag{2.9}$$

T^{a} represent the generators of $\mathrm{SU}(2)_L$, g and g' are the coupling constants of electroweak interactions and Y is the $\mathrm{U}(1)_Y$ hypercharge. The four vector fields $\mathrm{W}^{\mathrm{a}}_\mu$ and B^0_μ are combined into the form of the known, **massive SM vector fields** $\mathrm{W}^\pm_\mu$ and Z^0_μ, as well as the massless vector field A_μ of the photon, orthogonal to Z^0_μ:

$$\begin{array}{lclllcl}
W^\pm_\mu & = & \frac{1}{\sqrt{2}} & \left(W^1_\mu \mp iW^2_\mu\right) & \text{with mass} & m_{W^\pm} & = 80.4\ \text{GeV} \\
Z^0_\mu & = & \frac{1}{\sqrt{g^2+g'^2}} & \left(gW^3_\mu - g'B^0_\mu\right) & \text{with mass} & m_{Z^0} & = 91.2\ \text{GeV} \\
A_\mu & = & \frac{1}{\sqrt{g^2+g'^2}} & \left(g'W^3_\mu + gB^0_\mu\right) & \text{with mass} & m_\gamma & = 0\ .
\end{array}$$

The mixing of the fundamental fields W^3_μ and B^0_μ is formulated in terms of a change of basis, introducing the *weak-mixing angle* θ_{W}:

$$\begin{pmatrix} W^3 \\ B \end{pmatrix} \to \begin{pmatrix} Z^0 \\ A \end{pmatrix} = \begin{pmatrix} \cos(\theta_{\mathrm{W}}) & -\sin(\theta_{\mathrm{W}}) \\ \sin(\theta_{\mathrm{W}}) & \cos(\theta_{\mathrm{W}}) \end{pmatrix} \begin{pmatrix} W^3 \\ B \end{pmatrix} \tag{2.10}$$
$$\text{with } \cos^2(\theta_{\mathrm{W}}) = \frac{m_{\mathrm{W}}}{m_{\mathrm{Z}}}\ .$$

From equation 2.10 follows the important relation for the **electron charge** e and the electric charge quantum number Q:

$$g\,\sin(\theta_{\mathrm{W}}) = g'\,\cos(\theta_{\mathrm{W}}) = e \quad\Longleftrightarrow\quad Q = T^3 + Y\,. \tag{2.11}$$

Coupling to Fermions

A characteristic feature of the weak interaction is, that the $\mathrm{W}^\pm$ bosons only couple to the *left-handed* helicity states of the quark and lepton fields. The left-handed fermion fields ψ_{L} are assigned to **isospin doublets** under $\mathrm{SU}(2)_L$ and the *right-handed* fermion fields ψ_{R} are assigned to **singlets**:

$$\psi = \psi_L + \psi_R \quad \Longleftrightarrow \quad \left\{ \begin{array}{lllll} & \psi_L = \frac{1+\gamma_5}{2}\, E^i_{\mathrm{L}}\,, & \frac{1+\gamma_5}{2}\, Q^i_{\mathrm{L}} & & \\ \text{and} & \psi_R = \frac{1-\gamma_5}{2}\, (e_i)_{\mathrm{R}}\,, & \ldots & \frac{1-\gamma_5}{2}\, (u_i)_{\mathrm{R}}\,, & \ldots \end{array} \right\}$$

$$\text{where} \quad E^i_{\mathrm{L}} = \left(\begin{array}{c} \nu_\ell \\ \ell^- \end{array} \right)_{\mathrm{L}} \quad \text{and} \quad Q^i_{\mathrm{L}} = \left(\begin{array}{c} u_i \\ d_i \end{array} \right)_{\mathrm{L}} .$$

Therein, ℓ represents the three different lepton flavours e, μ and τ. There are also three doublets of left-handed quarks:

$$Q^i_{\mathrm{L}} = \left(\begin{array}{c} u^i \\ d^i \end{array} \right)_{\mathrm{L}} = \left(\left(\begin{array}{c} u \\ d \end{array} \right)_{\mathrm{L}} , \left(\begin{array}{c} c \\ s \end{array} \right)_{\mathrm{L}} , \left(\begin{array}{c} t \\ b \end{array} \right)_{\mathrm{L}} \right) . \tag{2.12}$$

The **charged, weak currents** $\mathcal{J}^{\mu\pm}_W$ couple the down-type quark and lepton fields to up-type fields or vice versa, as described in [1]:

$$\mathcal{J}^{\mu+}_W = \frac{1}{\sqrt{2}}(\bar{\nu}_{\mathrm{L}}\gamma_\mu \ell_{\mathrm{L}} + \bar{u}^i_{\mathrm{L}}\gamma_\mu d^i_{\mathrm{L}}) \quad , \quad \mathcal{J}^{\mu-}_W = \frac{1}{\sqrt{2}}(\bar{\ell}_{\mathrm{L}}\gamma_\mu \nu_{\mathrm{L}} + \bar{d}^i_{\mathrm{L}}\gamma_\mu u^i_{\mathrm{L}}) . \tag{2.13}$$

Flavour Mixing

The charged, weak interactions link the three up-type quarks u^i_{L} with an unitary rotation to the down-type quarks d^i_{L}. The rotation is given by the **Cabibbo-Kobayashi-Maskawa mixing matrix** $V_{\mathrm{CKM}} = \left(U^\dagger_u U_d\right)$. Therefore, the charged, weak current $\mathcal{J}^{\mu+}_W$ is described by unitary transformations among the three quark doublets:

$$\mathcal{J}^{+}_W = \left(\begin{array}{ccc} \bar{u}, & \bar{c}, & \bar{t} \end{array} \right) V_{\mathrm{CKM}} \left(\begin{array}{c} d \\ s \\ b \end{array} \right) \quad \wedge \quad V_{\mathrm{CKM}} = \left(\begin{array}{ccc} V_{\mathrm{ud}} & V_{\mathrm{us}} & V_{\mathrm{ub}} \\ V_{\mathrm{cd}} & V_{\mathrm{cs}} & V_{\mathrm{cb}} \\ V_{\mathrm{td}} & V_{\mathrm{ts}} & V_{\mathrm{tb}} \end{array} \right) . \tag{2.14}$$

There is also evidence for a mixing occurring in the lepton sector. If neutrinos do not have masses equals to zero, they may oscillate among different types of neutrinos (ν_{e}, ν_μ, ν_τ). A detailed description of neutrino masses and lepton flavour mixing can be found in [2].

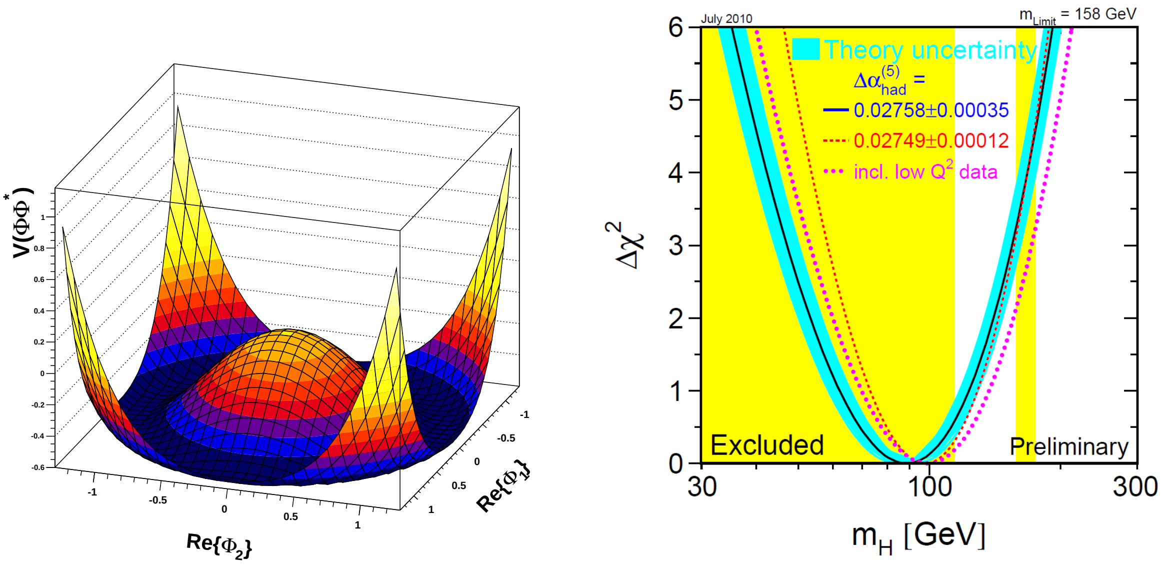
Electroweak Symmetry Breaking and the Higgs Mechanism

In the SM theory of the electroweak interaction, all fields are generated massless. By introducing an additional **complex scalar field** $\phi(x)$, the $\mathrm{W}^\pm$ and Z^0 gauge bosons acquire masses by coupling to the new field with the potential form:

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger\phi)^2 . \tag{2.15}$$

A graphical representation of equation 2.15 is given in figure 2.1 (a). The Lagrangian is symmetric under rotations of $\phi(x)$. If the field takes on a non-zero global value:

$$\langle\phi\rangle = \phi_0 \equiv \pm v = \pm\sqrt{\frac{\mu^2}{\lambda}} , \tag{2.16}$$



(a) (b)

Figure 2.1: (a) Representation of the Higgs potential form $V(\phi)$ given in equation 2.15, with $\phi = (\Phi_1, \Phi_2)$. (b) Standard Model fit of the expectation value for a neutral Higgs boson [4].

referred to as the *vacuum expectation value* v, the symmetry of the Lagrangian is spontaneously broken. The scalar field is reformulated in terms of an expansion around its ground state ϕ_0, which is chosen arbitrarily:

$$\phi(x) \;=\; \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \quad\Longleftrightarrow\quad \mathcal{L}_{\text{Higgs}} \;=\; |D_\mu\phi|^2 + \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2 \ . \qquad (2.17)$$

The scalar spin-0 field $h(x)$ is referred to as the **Higgs field**. The masses of the **gauge bosons** ($m_{\mathrm{W}^\pm}$ and m_{Z^0}) follow from the coupling of the initial weak-isospin triplet W^i_μ and hypercharge singlet B^0_μ vector fields to the Higgs field, represented by the kinetic term of the Lagrangian:

$$D_\mu\phi \;=\; (\partial_\mu - igW^{\mathrm{a}}_\mu T^{\mathrm{a}} - ig'YB^0_\mu)\,\phi \quad\Longrightarrow\quad \left\{ \begin{array}{lcl} m_{\mathrm{W}^\pm} & = & g\,\dfrac{v}{2} \\ m_{\mathrm{Z}^0} & = & \sqrt{g^2+g'^2}\,\dfrac{v}{2} \end{array} \right\} \ . \qquad (2.18)$$

In figure 2.2, the two Feynman graphs of the coupling of the weak gauge bosons to the scalar Higgs field are shown. In the SM, also the masses of the fermions are generated by *ad hoc* coupling to the Higgs field. It is generically introducted by assuming a *Yukawa coupling* between the fermion fields ψ and the scalar field ϕ. The **mass of the fermions** m_f is proportional to the coupling strength λ_f to the Higgs field:

$$m_f \;=\; \lambda_f \frac{v}{\sqrt{2}} \ . \qquad (2.19)$$

Figure 2.3 shows the Feynman graphs for the generation of fermion masses, as well as the **self-coupling** of the massive Higgs boson, which creates its own mass $m_H = \sqrt{2\lambda}\,v$. The couplings of the massive SM particles to the Higgs field via their masses, have to be measured experimentally. Given the measured parameters of the SM, the preferred value of the SM Higgs boson mass is obtained from a fit to electroweak precision data, which is described in further detail in [4]. The result of the fitting procedure, which prefers the **Higgs boson mass** to be between 114 and 158 GeV/c^2, is shown in figure 2.1 (b).

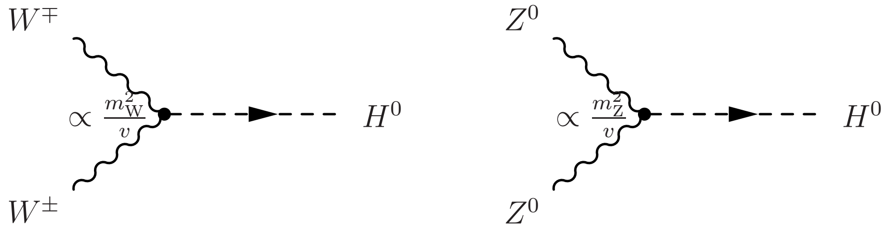


Figure 2.2: (Left) $W^{\pm}$ and (right) Z^0 bosons acquire its mass via coupling to the scalar Higgs field H^0.

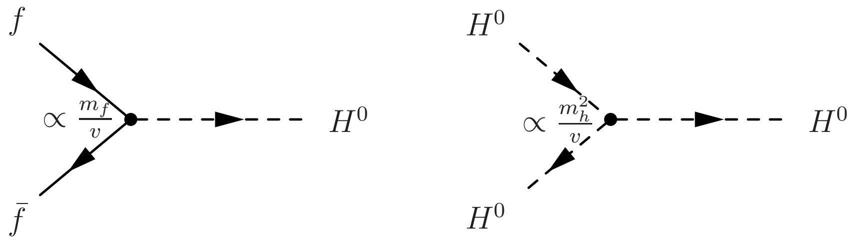


Figure 2.3: (Left) Fermion masses are also generated via coupling to the Higgs field. Therefore, a Yukawa term in the Lagrangian of the Standard Model is introduced. (Right) The Higgs boson H^0 generates its mass via self-coupling. Here, only a trilinear coupling is shown, whereas also quartic coupling is allowed.

2.3.3 Quantum Chromodynamics

The theoretical description of the strong interaction between quarks is given by the theory of *quantum chromodynamics* (QCD). QCD is the gauge theory of strong interactions and its corresponding gauge group is SU(3)$_C$. Besides its fractional electric charge, every quark also carries the charge of the strong force field, called *colour charge*. As mentioned above, there are three flavours of the colour charge: *red* (r), *green* (g) and *blue* (b). Correspondingly, the antiquarks carry the conjugated charges, labled as *anti-red* ($\bar{r}$), *anti-green* ($\bar{g}$) and *anti-blue* ($\bar{b}$).

Interaction of Non-Abelian Gauge Bosons

The self-coupling of the gauge bosons in non-Abelian gauge theories has been introduced in the previous section. A more detailed description of the non-linear couplings of gluons and quarks is given in the following. The Lagrangian of Yang-Mills theories has been given in equation 2.8, where the field tensor $F^{\mathrm{a}}_{\mu\nu}$ is defined as:

$$F^{\mathrm{a}}_{\mu\nu} \;=\; = \partial_\mu A^{\mathrm{a}}_\nu - \partial_\nu A^{\mathrm{a}}_\mu \,+\, g f^{\mathrm{abc}} A^{\mathrm{b}}_\mu A^{\mathrm{c}}_\nu \,. \tag{2.20}$$

The numbers f^{abc} are the *structure constants* of QCD. From the square of the expression given above, the **gluon self-interaction** is derived. The Feynman graphs of the three- and four-gauge boson vertices are shown in figure 2.4.

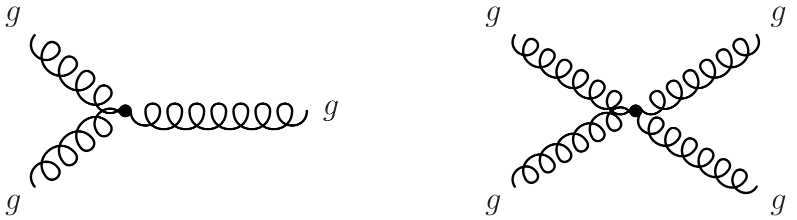


Figure 2.4: The non-linear interactions in QCD give rise to the self-coupling of gluons to each other. (Left) Three- and (right) four-gauge boson vertices are shown.

Strong Coupling Constant

The non-linear behaviour of self-interaction in QCD gives rise to a *scale dependence* of the strong coupling constant α_{s}. In principle, any coupling strength has a momentum or scale dependence. The coupling strength of strong interactions depends on a scale Q (for example the momentum transfer), which is described by the β-function of QCD:

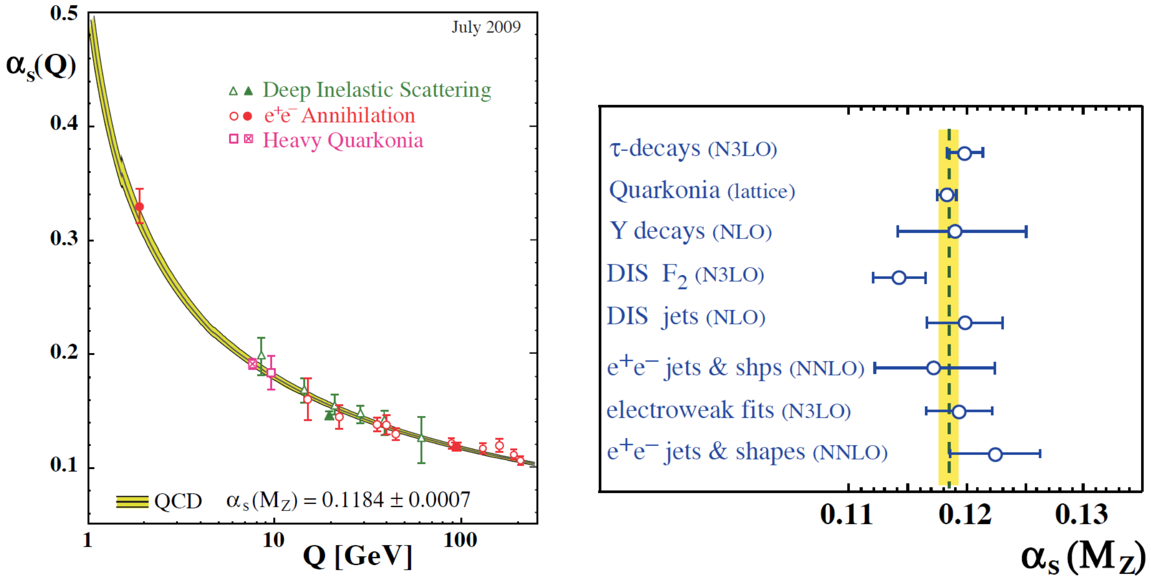
$$\frac{d\alpha_{\mathrm{s}}(Q^2)}{d\ln(Q^2)} \;=\; Q^2\frac{d\alpha_{\mathrm{s}}(Q^2)}{dQ^2} \;=\; \beta\left(\alpha_{\mathrm{s}}(Q^2)\right)\;. \tag{2.21}$$

This equation is referred to as the **renormalisation group equation**. The leading-order solution of equation 2.21 is:

$$\alpha_{\mathrm{s}}(Q) \;=\; \frac{2\pi}{b_0\log(Q/\Lambda_{\mathrm{QCD}})}\;, \tag{2.22}$$

with $b_0 \;=\; 11-\frac{2}{3}n_f$, where n_f is the number of approximately massless quarks. Equation 2.22 shows, that the strong coupling constant α_{s} *decreases* with increasing momentum transfer, which is shown in figure 2.5 (a). A detailed description of measurements of α_{s} is given in [5]. The scale dependence of α_{s} is referred to as "*running coupling*". Therefore, α_{s} becomes small for large momentum transfers and QCD can be treated as a *perturbative*, effective theory (pQCD). Perturbation theory is only valid, if Q is significantly larger than Λ_{QCD}, which is about $\Lambda_{\mathrm{QCD}} \approx 200$ MeV. The value of α_{s} has been measured to high accuracy at the energy scale of $Q = M_{\mathrm{Z}}$, which is also described in [5]. The world average value is $\alpha_{\mathrm{s}}(M_{\mathrm{Z}}) = 0.1184 \pm 0.0007$ and is shown in figure 2.5 (b).

When α_{s} decreases for higher energies, quarks couple less to the gluons. In the limit of a large momentum scale Q_0, quarks can be treated as freely propagating fermions. This fact is referred to as **asymptotic freedom**. On the other hand, at the energy scale of SM processes Q_{SM}, quarks cannot be observed as free particles. The strong potential between two quarks grows, when they are separated from each other. At a certain distance, an additional quark-antiquark pair is created and two colourless (neutral) objects are formed.



(a) (b)

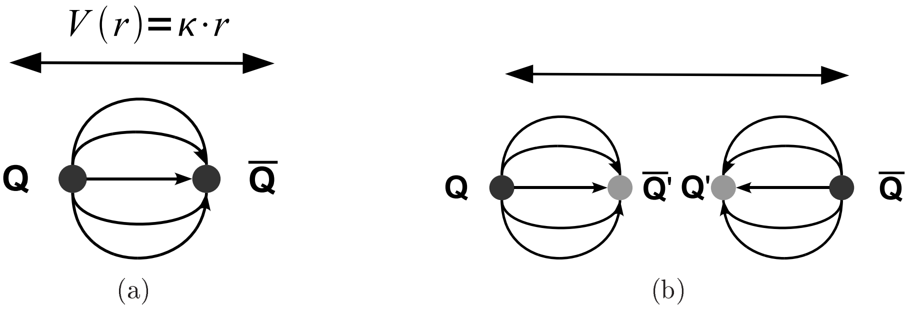
Figure 2.5: (a) The strong coupling constant α_s decreases with increasing momentum transfer Q. The scale dependence of α_s is called "*running coupling*". (b) Results of precision measurements of α_s at the energy scale of $Q = M_Z$. The world average value is $\alpha_s(M_Z) = 0.1184 \pm 0.0007$ [5].

In contrast to the asymptotic freedom at large energy scales, this low scale behaviour is referred to as **colour confinement**.

Hadronisation

The transformation process, when coloured quarks or gluons are confined into mesons and baryons is called *hadronisation*. There are different, phenomenological models, which describe the process of hadronisation, because this effect occurs at low energies, where it cannot be treated in perturbation theory. Therefore, an effective theory, like the **string fragmentation model**, is applied:

The attractive strong force between two coloured particles increases with growing distance between them. In the string fragmentation model, at short distances, this is described approximately by a linear potential of the form $V(r) = \kappa \cdot \mathrm{r}$. The linear confinement gives rise to a *colour dipole field* between charge and anticharge, which is shown in figure 2.6 (a). The *string constant* is approximately $\kappa \approx 1\,\frac{\mathrm{GeV}}{\mathrm{fm}}$, which can be interpreted as the average energy density of a string. If the potential energy, which is stored in the strings, reaches a certain value, the string breaks up and a new quark-antiquark pair is produced. Together with the two initial quarks, two new quark-antiquark systems (mesons) are formed, which is illustrated in figure 2.6 (b). If the string is broken up by a diquark-antidiquark pair, the string breakup gives rise the creation of a baryon system. In the **Lund string model**, the string breakup is described mathematically by quantum mechanical tunnelling, which is used for the simulation of hadronisation processes, which is described in [6].



(a) (b)

Figure 2.6: (a) In the string fragmentation model, at short distances, the strong coupling between quarks and gluons is described approximately by a linear string potential potential of the form $V(r) = \kappa \cdot \mathrm{r}$. (b) If the potential energy, which is stored in the strings, reaches a certain value, the string breaks up and a new quark-antiquark pair is produced. Together with the two initial quarks, two new diquark systems (mesons) are formed.

Particle Phenomenology

Quarks are bound into colourless states of two or three quarks. All strongly interacting particles, which are composed of quarks, are called *hadrons*. Their wavefunctions must be invariant under $\mathrm{SU}(3)_C$ symmetry transformations. Colourless states are either composed of a quark-antiquark pair ($\bar{q}q'$), which constitutes the class of *scalar* (spin-0) or *vector* (spin-1) mesons. On the other hand, the class of hadrons composed of three quarks ($qq'q''$) is referred to as *baryons*. Baryons are the building blocks of stable matter, the atomic nucleus, which is composed of protons (uud) and neutrons (udd).

2.4 Structure of the Proton

In the quark-parton model the proton only consists of three *valance quarks*: two up quarks and one down quark (uud). Furthermore, there is a large number of gluons inside the proton, which can split into quark-antiquark pairs. Quark pairs, which were produced from gluon splitting, are called *sea quarks*. In order to understand the production mechanisms of the top quark in proton-proton collisions, the structure of the proton will be discussed in further detail.

Parton Distribution Functions

The quark and gluon content of the proton is described by the *parton distribution functions* (PDFs). The function $f_i^{\mathrm{A}}(x)$ specifies the probability for finding a parton of type i inside the particle A, which carries the momentum fraction x of the total momentum of A. The cross section of a fermion-antifermion pair production process in proton-proton collisions $p_1 p_2 \to f\bar{f}$ is given in equation 2.23:

$$\sigma_{p_1 p_2 \to f\bar{f}} \;=\; \int dx_1 \int dx_2 \; f_a^{p_1}(x_1) \; f_b^{p_2}(x_2) \; \hat{\sigma}_{ab\to f\bar{f}} \;, \tag{2.23}$$

where $\hat{\sigma}_{ab\to f\bar{f}}$ is the cross section for the hard scattering process of two partons a and b inside the protons, which is calculable in perturbative QCD [6]. The PDFs cannot be derived from first principles, but have to be measured in particle collisions. Using the quark-parton model, the PDFs are linked to the measured **structure functions** $F_1^p(x)$ and $F_2^p(x)$ of the proton [7]. $F_1^p(x)$ and $F_2^p(x)$ are given by the following equations:

$$F_1^p(x) = \frac{1}{2}\sum_i e_i^2 f_i^p(x) \quad , \tag{2.24}$$

$$F_2^p(x) = 2xF_1^p(x) = x\sum_i e_i^2 f_i^p(x) \ . \tag{2.25}$$

The different quark flavours are labelled by the index i. In proton-proton collisions at $\sqrt{s} = 7$ TeV, the up, down and strange quarks (u, d and s) are referred to as *light* quarks, whereas the charm and bottom (c and b) quarks are considered as *heavy* quarks. The electrical charge of a quark is denoted by e_i and is given in table 2.1. The currently most precise measurements of the proton structure have been derived from deep inelastic electron-proton scattering at HERA, described in [8]. In figure 2.7, the results obtained from a fitting procedure at a fixed energy scale of $Q^2 = 10$ GeV^2 of the combined measurements of the H1 and ZEUS experiments are shown. For large values of the momentum fraction x, the dominant contributions stem from up and down valance quarks. For low values of x, the gluon PDFs are dominating the composition of the proton structure.

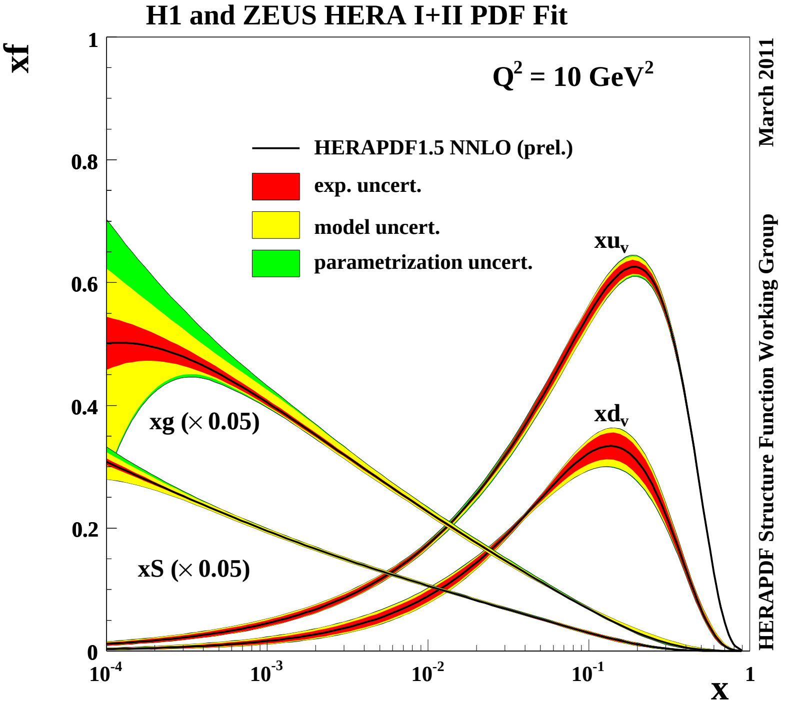


Figure 2.7: Parton distribution functions (PDFs) measured in deep inelastic proton-electron scattering at HERA. The distributions are obtained from a combined fitting procedure of the measurements of the H1 and ZEUS experiments, referred to as *HERA-PDF1.5* [8].

Parton Evolution

The PDFs are measured from deep inelastic electron-proton scattering at a specific momentum transfer $Q_0^2 = 10$ GeV^2, shown in figure 2.7. The hard interaction process of two partons depends on x_1 and x_2 of the two partons, as well as on the *energy scale* Q of the process. In order to perform calculations at higher scales, such as for proton-proton collisions at $\sqrt{s} = 7$ TeV, the structure functions of the proton have to be extrapolated towards higher energy scales. If the Q-dependent structure functions $f_i(x) = f_i(x, Q_0^2)$ are known at a specific scale Q_0, they can be calculated at any energy scale Q, using the **DGLAP equations** (Dokshitzer-Gribov-Lipatov-Altarelli-Parisi), described in [9] and [10], which are given in equation 2.26 at leading-order of α_s:

$$Q^2 \frac{d}{dQ^2} f_i(x, Q^2) \quad = \quad \frac{\alpha_\mathrm{s}(Q^2)}{2\pi} \sum_k \int_x^1 \frac{dz}{z} P_{ik}(z) \, f_k(\frac{x}{z}, Q^2) \ , \tag{2.26}$$

where $P_{ik}(z)$ are the *splitting kernels*, which describe the branching of one parton i into two, where the daughter parton k has the momentum fraction z of the mother parton.

Chapter 3

Physics of the Top Quark

The top quark was discovered in 1995 at the Tevatron in proton-antiproton collisions at $\sqrt{s} = 1.8$ TeV [11]. It is the heaviest elementary particle of the Standard Model. The mass of the top quark was measured to be (172.0 ± 1.6) GeV/c^2 [3]. Compared to the other five quarks, most of the properties of the top quark, like the spin or the electric charge, have not yet been measured with high accuracy. All measurements strongly favour the top quark to be the weak-isospin partner of the bottom quark, with $\mathrm{T}_3 = \frac{1}{2}$ and $\mathrm{Q} = \frac{2}{3}$. Hence, the top and bottom quark constitute the third generation of the quarks. The top quark has also been observed at the LHC within the first data taking period in 2010 in proton-proton collisions $\sqrt{s} = 7$ TeV, which is the subject of this thesis.

3.1 Production and Decay of Top Quarks

Production of Top Quarks

The top quark has been observed in proton-antiproton collisions at the Tevatron and in proton-proton collisions at the LHC at different centre-of-mass energies. An overview of the expected production cross sections of several Standard Model processes, as well as selected scenarios of Higgs boson production, as a function of the centre-of-mass energy, is shown in figure 3.1. In particular, the production cross section of $\mathrm{t\bar{t}}$ quark pairs increases significantly with increasing centre-of-mass energy. Top quark pairs are produced in inelastic scattering processes between two partons of the proton. Due to the high gluon content of the proton, gluon-gluon fusion processes are the dominant production process at the LHC, as described in [13]. The theoretical expectation of $\mathrm{t\bar{t}}$ quark pair production from gluon-gluon fusion processes at the design value of the LHC centre-of-mass energy of 14 TeV is about 90%. The remaining fraction of approximately 10% is expected to be produced in quark-antiquark annihilations. The leading-order production processes of $\mathrm{t\bar{t}}$ quark pairs are shown in figure 3.2.

The top quark pair production cross section is calculated using the Feynman rules for the leading-order processes, given by the following equation [14]:

$$\begin{array}{ccccccccc} q(p_1) & + & \bar{q}(p_2) & \rightarrow & t(p_3) & + & \bar{t}(p_4) & , \\ g(p_1) & + & g(p_2) & \rightarrow & t(p_3) & + & \bar{t}(p_4) & . \end{array} \tag{3.1}$$

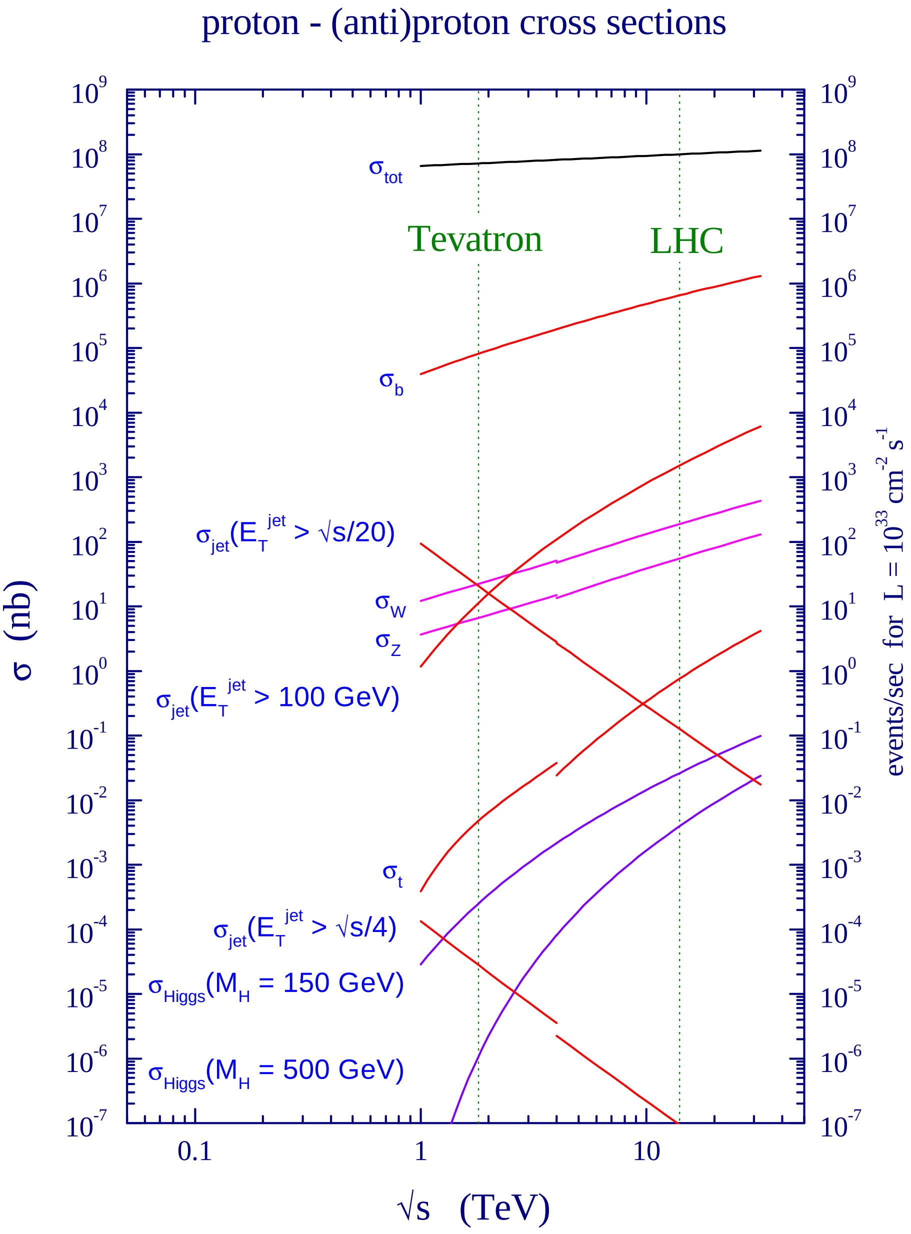


Figure 3.1: Production cross sections of several Standard Model and new physics processes, such as Higgs boson production, as a function of the centre-of-mass energy [12]. The expected cross sections are calculated for proton-antiproton collision for the Tevatron energy range, and for proton-proton collisions for energies feasible at the LHC.

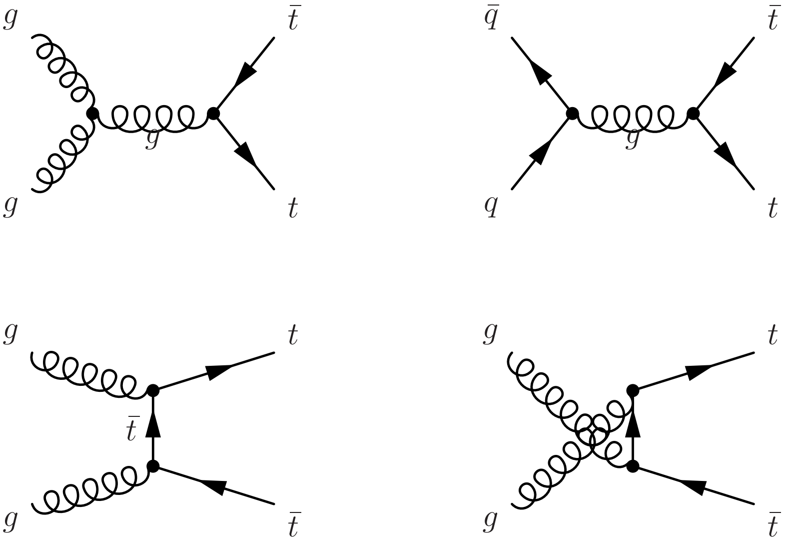


Figure 3.2: Leading-order top quark production processes in proton-proton collisions.

The four-momenta of the partons are defined as $p^\mu = (E, p_x, p_y, p_z)$ and are given in brackets. The leading-order, differential cross section of $\mathrm{t\bar{t}}$ quark pair production is obtained from the invariant matrix elements $\mathcal{M}_{ij}$, given by the following equation:

$$d\hat{\sigma}_{ij} \;=\; \frac{1}{2(p_1+p_2)^2}\,\frac{d^3p_3}{(2\pi)^3 2E_3}\,\frac{d^3p_4}{(2\pi)^3 2E_4}\,(2\pi)^4\,\delta^4(p_1+p_2-p_3-p_4)\,\overline{\sum}|\mathcal{M}_{ij}|^2\ , \qquad (3.2)$$

where the first factor $1/(2(p_1+p_2)^2)$ is the *flux factor* and $\overline{\sum}|\mathcal{M}_{ij}|^2$ is the sum of the square of the matrix elements, averaged over the initial and final state *colour and spin indices*, which are given in [14].

Decay of Top Quarks

The electroweak decay of the top quark is given by the CKM-matrix elements V_{td}, V_{ts} and V_{tb}, which have been defined by equation 2.14. The matrix elements can be most precisely determined by a global fit, that uses all available measurements of the CKM-matrix elements, described in [3]. Assuming unitarity of the CKM-matrix, V_{tb} is expected to be approximately 1.0 and the two matrix elements V_{td} and V_{ts} are negligibly small. Hence, the top quark decays almost to 100% into a bottom quark and a W boson, with results into a branching ratio of $\mathrm{BR}(t \to Wb) \approx 1.0$.

The W boson further decays either hadronically into two quarks $W \to q\bar{q}'$ or leptonically into a lepton-neutrino pair $W \to \ell\,\nu_\ell$. The dominant final states of the hadronic W decay are $u\bar{d}$ and $c\bar{s}$, which each have three coloured degrees of freedom. The other, allowed final states $c\bar{d}$ and $u\bar{s}$ are Cabibbo-suppressed. Therefore, as a leading-order approximation, the hadronically decaying W boson accounts for six final states. In addition to that, the leptonically decaying W boson accounts for three different final states with an electron, a muon or a tau lepton, which results in total into nine final states, each having an equal branching fraction of 1/9. A graphical representation of the fraction of the different decay channels is given in figure 3.3. According to the two different decay channels of the W boson, three different decay channels are assigned to $\mathrm{t\bar{t}}$ quark pairs:

Dileptonic decay ($\frac{1}{9} \approx 11\%$): $\mathrm{b\,W^+\,\bar{b}\,W^-} \;\rightarrow\; \mathrm{b}\,\ell^+\,\nu_\ell\,\mathrm{\bar{b}}\,\ell^-\,\bar{\nu}_\ell$

Semileptonic decay ($\frac{4}{9} \approx 44\%$): $\mathrm{b\,W^+\,\bar{b}\,W^-} \;\rightarrow\; \mathrm{b}\,\ell^+\,\nu_\ell\,\mathrm{\bar{b}\,q\,\bar{q}'} \;/\; \mathrm{b\,q\,\bar{q}'\,\bar{b}}\,\ell^-\,\bar{\nu}_\ell$

Fullhadronic decay ($\frac{4}{9} \approx 44\%$): $\mathrm{b\,W^+\,\bar{b}\,W^-} \;\rightarrow\; \mathrm{b\,q\,\bar{q}'\,\bar{b}\,\bar{q}''\,q'''}$

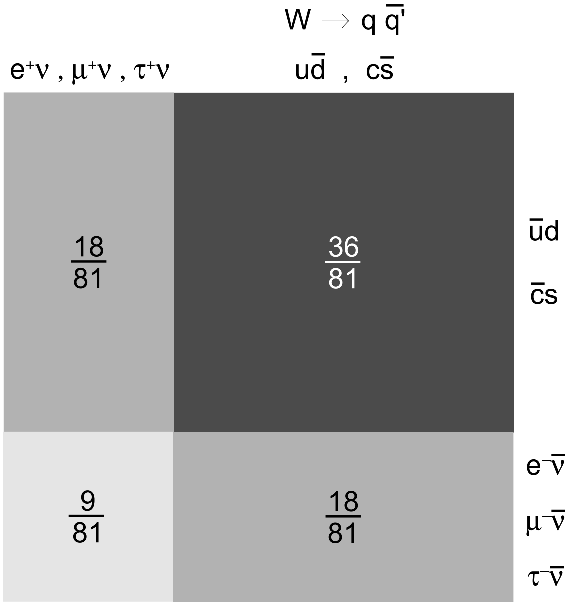


Figure 3.3: Fraction of the different decay modes of the W boson.

The event signature of the **dileptonic decay channel** of the top quark pair consists of two isolated, oppositely-charged leptons with high transverse momentum, two high-energetic b quarks and large missing transverse energy $\not{E}_\mathrm{T}$, due to the neutrinos from the W boson decays, which leave the detector without interaction and can only be detected indirectly. The dileptonic decay of a $\mathrm{t\bar{t}}$ quark pair is given by the following equation:

$$\mathrm{t\,\bar{t}} \;\longrightarrow\; \mathrm{b\,W^+\,\bar{b}\,W^-} \;\longrightarrow\; \mathrm{b} + \ell^+\,\nu_\ell + \mathrm{\bar{b}} + \ell^-\,\bar{\nu}_\ell \;. \qquad (3.3)$$

Figure 3.4 shows the dileptonic decay channel, which is the clearest event signature of $\mathrm{t\bar{t}}$ quark pair decays. The direct branching ratio of $\mathrm{t\bar{t}}$ quark pairs into a muon and an electron is 2.47%, which is enhanced to 3.41%, if leptonic τ decays into the muon-electron final state are also considered. Further corrections to the leading-order approximation of the branching ratio are applied at a later stage in the context of systematic uncertainties in section 7.1.

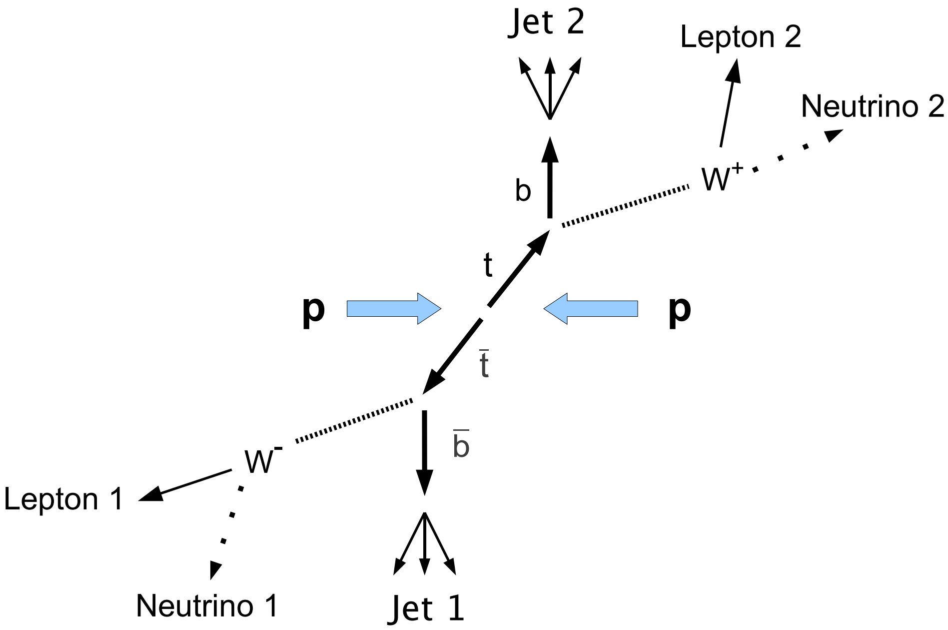


Figure 3.4: The dileptonic decay channel of a $t\bar{t}$ quark pair. The final state consists of two oppositely charged leptons, two b quarks and two neutrinos, which can only be detected indirectly by missing transverse energy.

3.2 Measurements of Top Quark Properties

The properties of $t\bar{t}$ quark pairs have been measured at the Tevatron, such as the mass of the top quark and the production cross section at 1.8 and 1.96 TeV centre-of-mass energy in proton-antiproton collisions. At the LHC, the top quark is produced with a much higher rate compared to the Tevatron. Hence, the LHC is referred to as the first "*top factory*" and allows for precision measurements of the properties of the top quark with high statistics. Due to its large mass, the lifetime of the top quark is shorter than the timescale of hadronisation. Therefore, top quarks decay before they hadronise, which gives an unique opportunity to study the properties of bare quarks.

Top Quark Pair Cross Section

The top quark pair production cross section has been measured in proton-antiproton collisions by the CDF and DØ experiments at $\sqrt{s} = 1.8$ and 1.96 TeV, described in [16]. The combined result for a top quark mass of 175 GeV/c^2 is (7.3 ± 0.9) pb. The measurements are continued at $\sqrt{s} = 7$ TeV in proton-proton collisions, performed by the CMS and ATLAS experiments at the LHC, based on the data taking period in 2010 with an integrated luminosity of about 3 and 36 pb^{-1}, described in [17] and [15]. Figure 3.5 shows the top quark pair cross section as a function of the centre-of-mass energy in proton-proton and proton-antiproton collisions, which shows the measured values obtained by all four experiments. The measured values at collision energies of 1.8, 1.96 and 7 TeV agree with the expected values obtained from next-to-leading order calculations. The next-to-leading order prediction for the $t\bar{t}$ quark pair cross section at $\sqrt{s} = 7$ TeV is $(157.5^{+23.2}_{-24.4})$

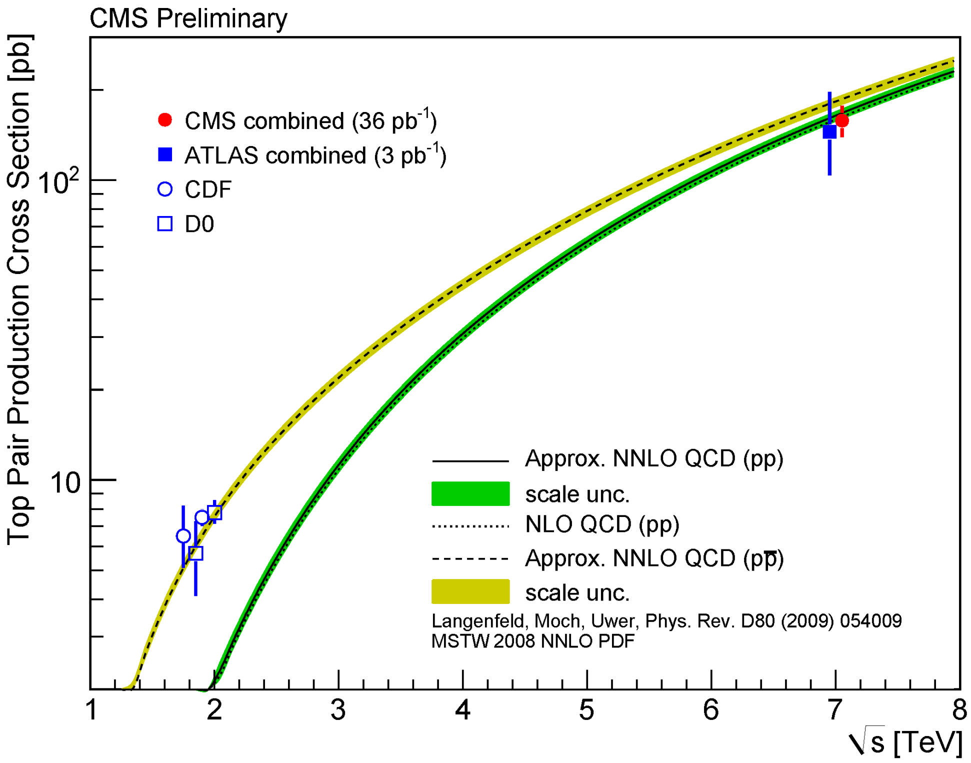


Figure 3.5: Top quark pair cross sections as a function of the centre-of-mass energy in proton-proton and proton-antiproton collisions. The measured values at collision energies of 1.8, 1.96 and 7 TeV are compared to the theoretical predictions obtained from next-to-leading order calculations [15].

pb, given in [18], based on a top quark mass of 172.5 GeV/c^2. The measurement performed by the top quark working group of the CMS collaboration, which is described in [19], is compared to the measurement presented in this thesis in section 7.5. Furthermore, first results from the measurement of differential cross sections based on the complete 2010 and the first part of the 2011 data taking periods are presented in section 7.6 and chapter 8.

Mass of the Top Quark

The mass of the top quark is a fundamental parameter of the Standard Model. Since the masses of the particles are free parameters, they cannot be calculated from theory, but they have to be measured experimentally. The top quark is the heaviest particle of the Standard Model. In combination with the mass of the W boson, it plays an important role on constraining the mass of the Higgs boson via quantum loop corrections, described in [4]. In figure 3.6, two quantum loop processes are shown, which include the top quark and the W boson. Based on the observed values of the top quark and W boson masses, limits on the mass of the Higgs boson can be calculated, which is shown in figure 3.7 (b).

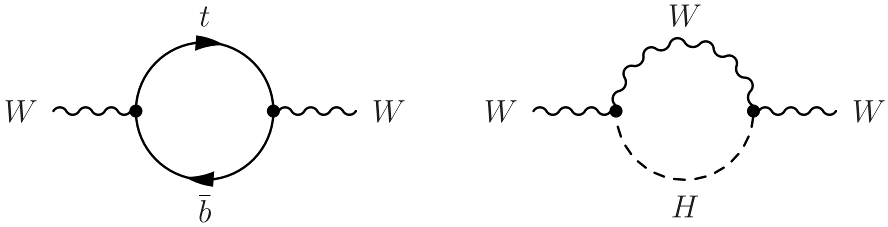
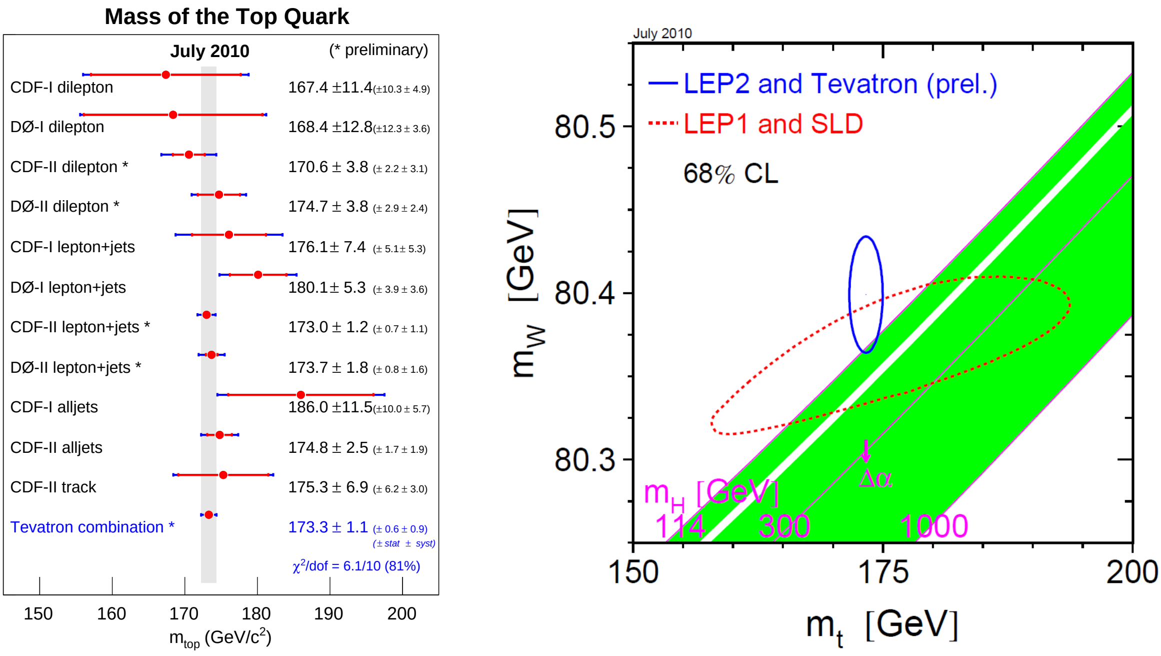


Figure 3.6: Quantum loop processes, which include the top quark and the W boson. The top quark and W boson masses play a key role by constraining the mass of the Higgs boson via quantum loop corrections.

The most precise measurement of the mass of the top quark has been performed in all final states at the CDF and DØ experiments, described in [20]. Figure 3.7 (a) shows the most recent combination of the measurements of the CDF and DØ collaborations, which result in a top quark mass of (173.3 ± 1.1) GeV/c^2. A first measurement of the top quark mass at the LHC is presented in [19], using only the dileptonic decay channel, which results in a top mass of (175.5 ± 6.5) GeV/c^2.



(a) (b)

Figure 3.7: (a) Combination of the measurements of the top quark mass by the CDF and DØ collaborations [20]. (b) Limits on the mass of the Higgs boson, which are calculated using quantum loop processes, which include the top quark and the W boson [4].

Chapter 4

LHC and CMS Experiment

Erstens kommt es anders und zweitens als man denkt.
Wilhelm Busch

4.1 Large Hadron Collider

The Large Hadron Collider (LHC) is a superconducting accelerator experiment, hosted at the European Organisation for Nuclear Research (CERN, French: *Organisation Européenne pour la Recherche Nucléaire*), Geneva (Switzerland). It is a **proton-proton collider** aiming for the discovery of the Higgs boson and physics beyond the Standard Model with a centre-of-mass energy of up to 14 TeV [21]. It was built between 2000 and 2008 into the existing tunnel complex of the Large Electron Positron (LEP) collider. After eight years of construction, the LHC had first beams injected in summer 2008. After a major technical accident in September 2008, the LHC provided first collisions in November 2009. Protons are accelerated in two separate rings of 26.7 km circumference, which are hosted in the same tunnel complex. In 2010, protons were accelerated up to energies of 3.5 TeV, resulting in a centre-of-mass energy of $\sqrt{s} = 7$ TeV. The LHC is currently the worlds most high-energetic proton accelerator, with a design luminosity is $\mathcal{L} = 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ and a design energy of the protons of 7 TeV. In 2010, a peak luminosity of $\mathcal{L} = 2 \cdot 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ was achieved.

The LHC comprises 1232 dipole magnets [22], which provide a magnetic dipole field of 8.33 T at an energy of 7 TeV per proton beam. The superconducting dipole magnets are cooled with super-fluid helium down to temperatures of 1.9 K. Each proton beam is separated into bunches of $N_{\mathrm{p}} = 1.15 \times 10^{11}$ protons per bunch, which have a spacing of 25 ns. This corresponds to a collision frequency of 40 MHz. The design value for the number of bunches per beam is $n_{\mathrm{b}} = 2808$. Figure 4.1 gives an overview over the LHC accelerator complex [23]. Before the proton bunches are injected into the main accelerator LHC, they are accelerated in several pre-accelerators. Proton bunches of 26 GeV and the correct spacing of 25 ns are formed in the Proton Synchrotron (PS). Afterwards, beams are accelerated to 450 GeV in the Super Proton Synchrotron (SPS), before they are transferred to the LHC.

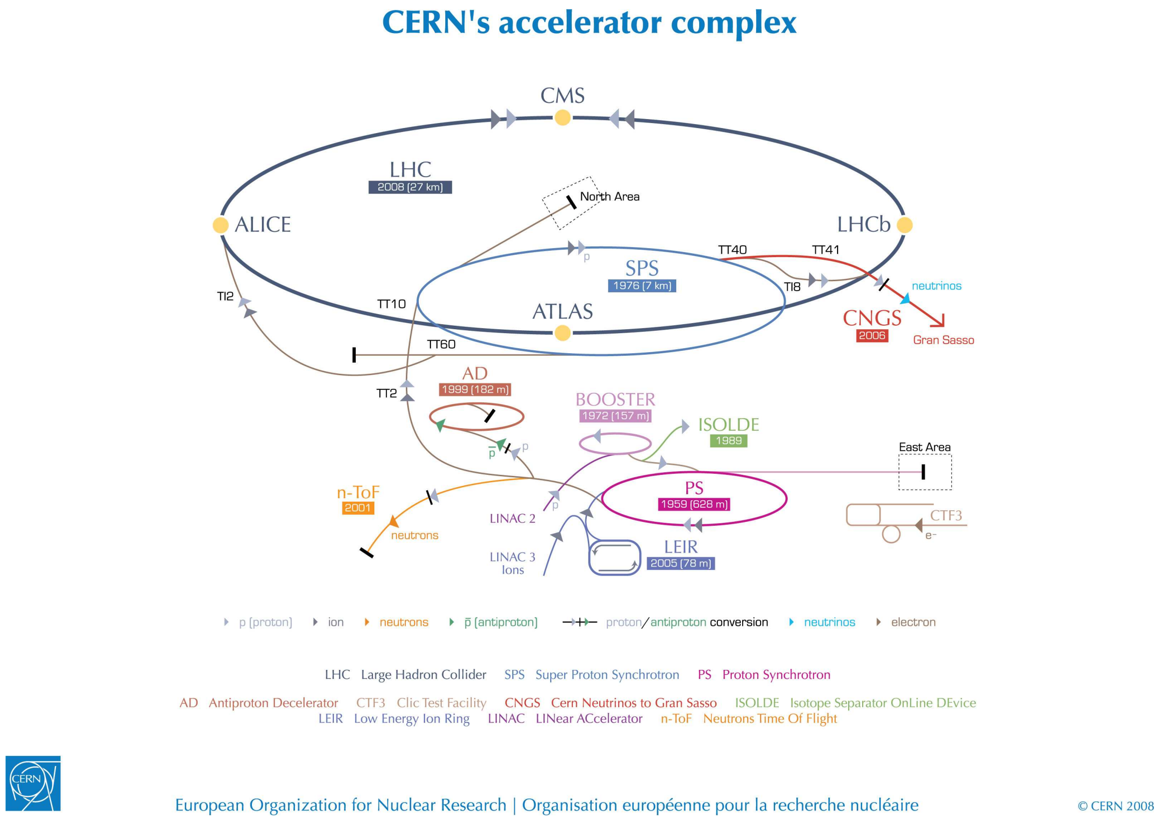


Figure 4.1: The Large Hadron Collider (LHC) accelerator complex, hosted at the European Organisation for Nuclear Research (CERN, French: *Organisation Européenne pour la Recherche Nucléaire*), Geneva (Switzerland). It is a superconducting collider experiment, where protons are collided at centre-of-mass energies of 7 TeV in an accelerator complex of 26.7 km circumference [23].

Luminosity

The **number of events**, generated in proton-proton collisions, is given by:

$$N_{\mathrm{event}} \;=\; L \times \sigma_{\mathrm{event}} \ , \tag{4.1}$$

where $L = \int dt\, \mathcal{L}$ is the *integrated luminosity* and σ_{event} is the cross section for a specific process. The instantaneous luminosity (for simplicity referred to as the *luminosity*) is given by equation 4.2, as described in [21]:

$$\mathcal{L} \;=\; \frac{N_{\mathrm{p}}^2\, n_{\mathrm{b}}\, \gamma\, f_{\mathrm{r}}}{4\pi\, \epsilon_{\mathrm{n}}\, \beta^*}\, F \ , \tag{4.2}$$

where N_{p} is the number of protons per bunch, n_{b} the number of bunches per beam, γ the relativistic gamma factor, f_{r} the revolution frequency, ϵ_{n} the normalised transverse beam emittance, β^* the beta function at the collision point and F the geometric luminosity reduction factor, due to the crossing angle at the interaction point (IP) [21]. The total integrated luminosity L provides an absolute normalisation for most physics analysis, such as the measurement of cross sections. Different methods for the measurement of the luminosity are described below in 4.2.7.

Experiments at the LHC

There are four main experiments at the LHC: ATLAS, CMS, ALICE and LHCb. The **ATLAS** and **CMS** experiments are multi-purpose detectors. They are designed to cover the full spectrum of measurements at the energy frontier and precision measurements of Standard Model processes. Important areas of *new physics searches* are the potential discovery of the Higgs boson, as well as supersymmetric particles and exotic signatures. ATLAS and CMS are complemented by two special-purpose experiments: **LHCb** is particularly designed for the study of b physics at a peak luminosity of $10^{32}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$ and **TOTEM** for the measurement of the elastic proton-proton cross section. LHCb will also allow for performing precise measurements of *CP-violation*, which allow for indirect searches of new physics in rare decays. In addition to the proton beams, the LHC was also operated with heavy-ion beams at the end of 2010. The **ALICE** experiment is a dedicated heavy-ion experiment for the studies of lead-lead collisions. The design energy per nucleon in Pb-Pb collisions is 2.76 TeV, which results in a total energy of 1148 TeV of two colliding lead nuclei.

4.2 Compact Muon Solenoid

A comprehensive description of the Compact Muon Solenoid (CMS) detector is given in [22] and [24]. The CMS detector is a hermetic, 4π multi-purpose detector, shown in figure 4.2. All subdetectors are arranged concentrically around the interaction point (IP). The detector is divided into a central region, called *barrel*, and two *endcap* regions. The main feature of the CMS detector is the 3.8 T magnetic field, created by a **superconducting solenoid**, which allows an excellent momentum resolution of reconstructed objects. Furthermore, it has three different subdetectors for the **detection of muons**, which are located outside the solenoid. In the barrel region, aluminium drift tubes (DT) are used, and cathode strip chambers (CSC) in the endcaps. Both muon systems are complemented by resistive plate chambers (RPC) for improved trigger timing. The sandwich-like arrangement of the muon detectors and the iron return yoke gives rise to the characteristic appearance of the CMS detector. The weight of the solenoid including the iron return yoke is 10 000 t, the total weight of the whole detector is 12 500 t. Its dimensions are a length of 22 m and a diameter of 15 m. Therefore the detector is denoted as "*compact*".

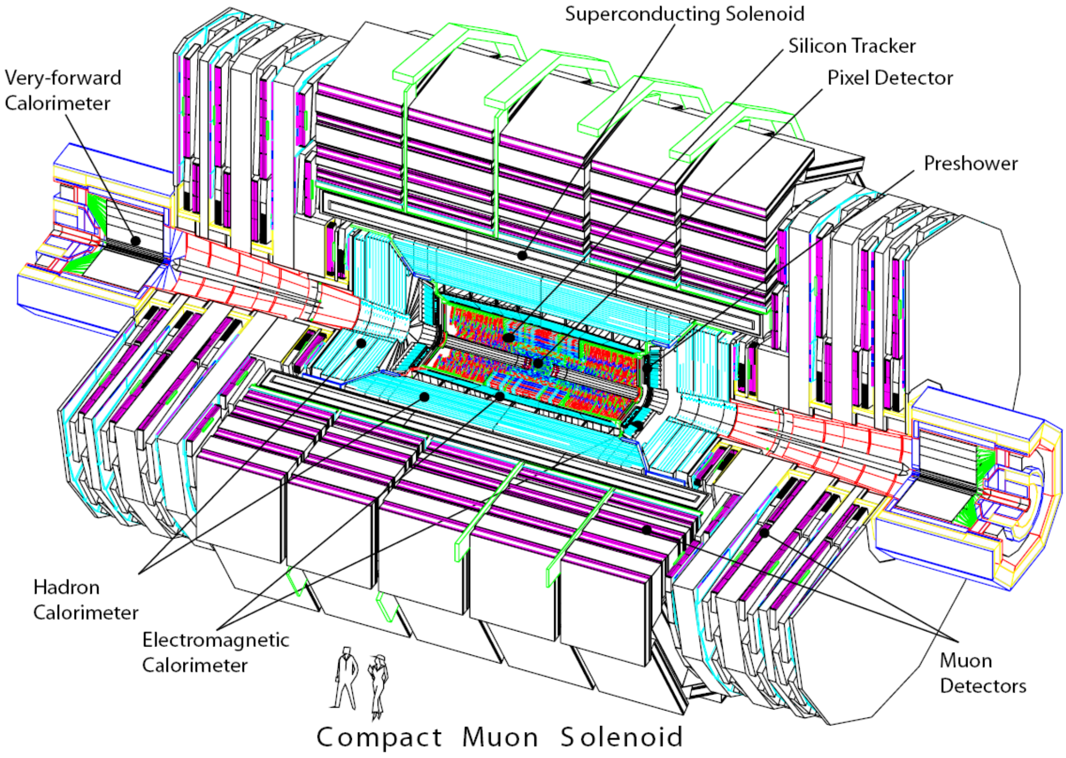


Figure 4.2: Schematic overview of the CMS detector [22]. Its dimensions are a length of 22 m and a diameter of 15 m.

The fully silicon-based *inner tracking system* comprises a 3-layer silicon-pixel and a 10-layer silicon-strip detector, which allow for an excellent spatial resolution close to the IP, as well as a high momentum resolution. The *electromagnetic calorimeter* (ECAL) is composed of lead-tungstate crystals, which provide an excellent electromagnetic energy resolution for the measurement of electrons and photons. It is optimised for a potential discovery of a light, neutral Higgs boson, decaying into a pair of photons ($H^0 \rightarrow \gamma\gamma$). A sampling *hadronic calorimeter* (HCAL) composed of brass absorbers and plastic scintillators completes the calorimetric system, which is hosted inside the solenoid. One feature of the CMS detector is, that the inner tracking system, as well as the complete calorimetry (ECAL and HCAL) are hosted inside the solenoid coil. The calorimetric system is supplemented by an additional hadronic outer (HO) calorimeter, which detects possible tails of hadronic showers, referred to as "*tail-catcher*". The CMS coordinate system uses the **pseudorapidity**, based on the polar angle θ, which is defined as $\eta \equiv -\ln\left(\tan\left(\frac{\theta}{2}\right)\right)$.

4.2.1 Superconducting Magnet

For the precise measurement of the particle momenta, a high magnetic field is needed. The magnetic field is created by a **superconducting solenoid** of 13 m length and 6 m inner diameter. This solenoid is the largest magnet coil ever built and provides a magnetic field of up to 4 T. In the 2010 data taking period, it was operated at 3.8 T. The total energy stored in the coil is 2.6 GJ, which corresponds to a stored energy of 11.6 kJ per kilogramme of cold mass, with a weight of the cold mass of 220 t. The thickness of the cold mass is 3.9 radiation lengths. The **iron yoke**, which returns the magnetic flux, comprises five barrel modules (*wheels*) and two endcaps. The mass of the iron barrel is 6 000 t and each endcap has a mass of 2 000 t, which results in a total weight of 10 000 t of the return yoke, as described in [24].

4.2.2 Inner Tracking System

The inner tracking system, also shortly referred to as the *tracker*, is the innermost subdetector closest to the interaction point, and thus has to cope with a high charged-particle flux. The main task of the inner tracking system is to reconstruct **charged-particle trajectories** and to measure particle momenta, as well as the precise reconstruction of **primary and secondary vertices**. Also in heavy-ion collisions, where the density of charged tracks is much higher, the CMS tracker permits the reconstruction of tracks in a high-density environment, as described in [22]. The reconstruction of charged-particle tracks and the charge identification of reconstructed objects is performed by the inner tracking system. The precise determination of particle momenta and charges is a major requirement of the analysis of top quark decays, as well as for the use of b quark identification algorithms, which rely on the exact reconstruction of displaced vertices.

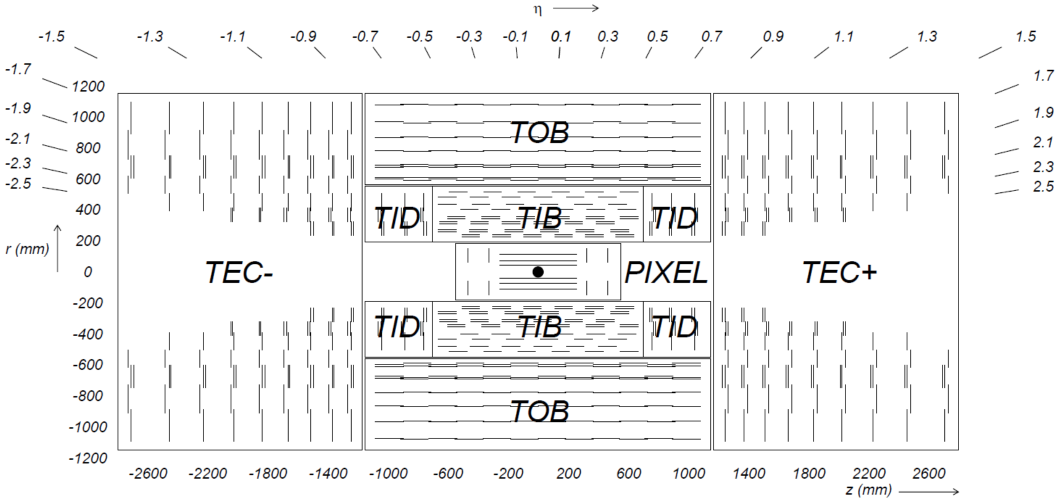


Figure 4.3: Overview of the fully silicon-based inner tracking system [24]. The inner tracking system is divided into two subdetectors: pixel tracker, also referred to as vertex detector, and strip tracker. The geometry is shown, using the pseudorapidity η.

The inner tracking system is divided into two subdetectors: **pixel tracker**, which is also referred to as *vertex detector*, and **strip tracker**. Both subsystems are fully based on silicon sensors. The inner tracking system consists of 66 million silicon pixels and 9.6 million silicon strips. Figure 4.3 shows a schematic view of the CMS inner tracking system, also showing the geometry in terms of the pseudorapidity η, which has been defined above.

Silicon Pixel Detector (Vertex Detector)

In the barrel region, close to the interaction point, three layers of pixel detectors are placed in parallel to the beam pipe at radii of $r = 4.4$, 7.3 and 10.2 cm with a size of a pixel of $100 \times 150\,\mu\text{m}^2$. Additionally, there are two pixel endcap disks in the forward direction transverse to the beam line, which both have two pixel layers at distances of $|z| = 34.5$ and 46.5 cm. The pixel tracker is also referred to as **vertex detector**, because the main task of the inner tracking system is to provide information about the primary interaction point (primary vertex) and displaced interaction points (secondary vertices) from the decay of long-lived unstable particles, like bottom or charm quarks. Due to the high magnetic field, charged particles traversing the silicon sensors are deflected by the *Lorentz force*. In the endcap disks, the blades carrying the pixel modules are rotated by 20° to compensate for the deflection, which leads to a turbine-like geometry of the endcaps, which is shown in figure 4.4. A detailed description of the pixel tracker is given in [22].

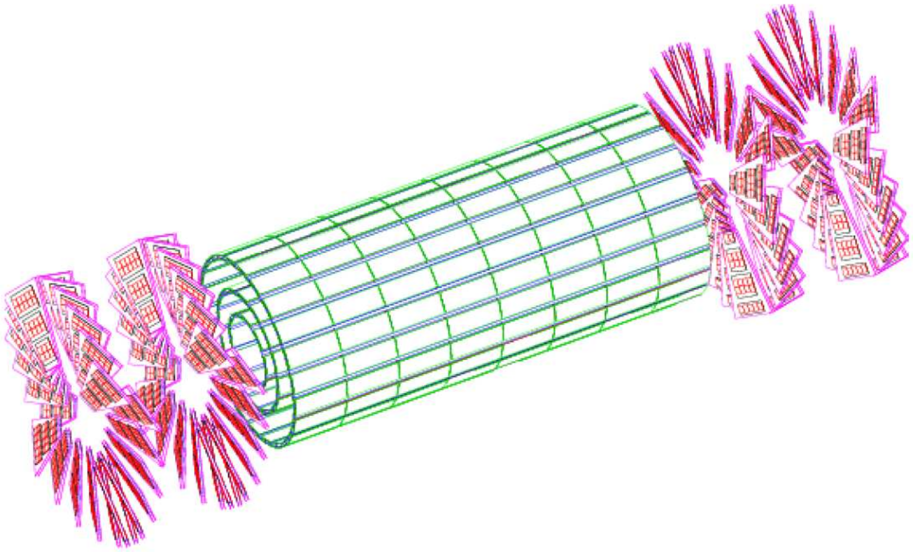


Figure 4.4: The pixel tracker (vertex detector) comprises three layers of pixel detectors placed in parallel to the beam pipe, as well as two pixel endcap disks in the forward direction transverse to the beam line [22]. In the endcap disks, the blades carrying the pixel modules are rotated by 20° to compensate for the deflection of charged particles by the Lorentz force.

Silicon Strip Detector (Tracker)

The silicon strip tracker system provides a coverage of $|\eta| < 2.4$ and consists of almost 15400 modules, which are mounted on a carbon-fibre structure and are operated at a temperature of around -20°C. The strip tracker covers a region between $20 < r < 110$ cm, where the particle flux decreases with respect to the innermost region of the pixel tracker. Therefore silicon microstrip detectors are used. A comprehensive description of the strip tracker is given in [22].

The **tracker inner barrel** (TIB) consists of four layers of strip detectors. In the first two layers of the TIB, the modules are mounted as *stereo* modules with an angle of 100 mrad in order to provide a measurement in $(r - \phi)$ and $(r - z)$ direction. The outermost region of the inner tracking system is referred to as the **tracker outer barrel** (TOB), which covers a radius between $55 < r < 110$ cm. The significantly lower particle flux allows for the use of larger-pitch silicon microstrip detectors with a good signal-to-noise ratio. The TOB comprises six layers of strip detectors, where the first two layers are also mounted as *stereo* modules. In the forward region, there are nine layers of microstrips in each of the two **tracker endcaps** (TEC). Additionally, there are three layers of **tracker inner disks** (TID) on each side, in order to fill the gap in the transition region between TIB and TEC.

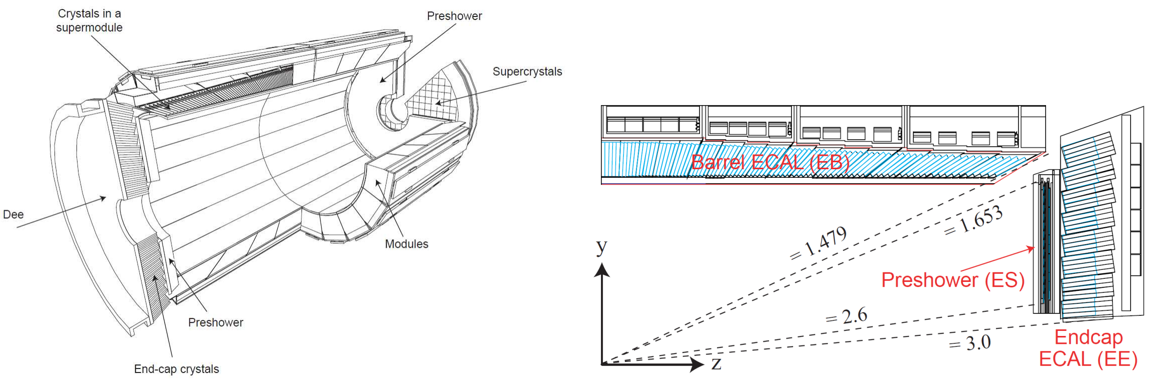
4.2.3 Calorimetry

The calorimetry provides an inclusive energy measurement of particles. Electromagnetically interacting particles, like the electron and the photon, are measured in the **electromagnetic calorimeter** (ECAL). Strongly interacting particles, like the neutron and proton, are observed as jets in the **hadronic calorimeter** (HCAL) and the ECAL. Weakly interacting particles, like neutrinos, escape the detector without depositing any energy. This is measured as an imbalance in the total sum of the transverse energy deposit, called *missing transverse energy* (MET) $\not{E}_{\mathrm{T}}$. A hermetic calorimetric system is needed, in order to precisely measure all energy deposits and to correctly reconstruct the E_{T} imbalance from neutrinos or other escaping particles. A complete describtion of the ECAL and HCAL is given in [24].

Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) is made of lead tungstate ($PbWO_4$). These scintillating crystals are a homogeneous, active material. The ECAL is hermetic and comprises about 75 000 crystals in the barrel and endcap regions. Figure 4.5 (a) shows a schematic view of the ECAL. The choice of $PbWO_4$ crystals provides a high density of $8.3\ \frac{\mathrm{g}}{\mathrm{cm}^3}$ and a short radiation length of $X_0 = 0.89$ cm, which results in a fine granularity and a compact ECAL [24]. The choice of $PbWO_4$ crystals allows for the construction of a compact calorimeter hosted inside the solenoid, which is radiation hard and has an excellent energy resolution.

The **ECAL barrel** (EB) covers a region in pseudorapidity of $0 < |\eta| < 1.479$, shown in figure 4.5 (b). The inner radius of the EB is 129 cm. The crystals of the ECAL have a front face cross-section of $22 \times 22\,\mathrm{mm}^2$ and a length of 230 mm, which corresponds to



(a) (b)

Figure 4.5: (a) Schematic view of the electromagnetic calorimeter (ECAL) [24]. (b) Geometry of the ECAL subdetectors: ECAL endcaps (EE), preshower (ES) and barrel (EB) [22].

nearly 26 radiation lengths [22]. The crystals in the EB are grouped into 36 identical *supermodules*. The **ECAL endcaps** (EE) cover a region in pseudorapidity of $1.479 < |\eta| < 3.0$. Each endcap consists of two semi-circular aluminium plates (*dees*), where 25 crystals are arranged into an array of 5×5 *supercrystals*. They have a front face cross-section of $28.6 \times 28.6\,\mathrm{mm}^2$ and a length of 220 mm, which corresponds to nearly $25\,X_0$. In front of each EE, a **preshower** (ES) device is mounted between $1.653 < |\eta| < 2.6$, which is shown in 4.5 (b). The aim of the preshower detector is to identify neutral pions, which dominantly decay into a pair of photons ($\pi^0 \rightarrow \gamma\gamma$), as described in [24].

Hadronic Calorimeter

The identification of hadronic jets and the measurement of missing transverse energy $\not\!\!E_{\mathrm{T}}$ is an important challenge at hadron colliders. The choice of the magnet and the almost complete inclusion of the calorimetric system into the solenoid determine the design parameters of the hadronic calorimeter (HCAL). The design of the HCAL maximises the material inside the magnet coil in terms of interaction lengths and provides hermeticity for the $\not\!\!E_{\mathrm{T}}$ measurement [22]. An additional layer of scintillators is placed outside the solenoid, referred to as the hadron outer (HO) detector, in order to detect tails of hadronic showers. The *absorber material* is brass, because it has a sufficiently short interaction length and is non-magnetic. The *active medium* consists of plastic scintillator tiles, which are read out with wavelength-shifting fibres. The scintillator plates have a thickness of 3.7 mm, except for the first layer attached to the ECAL barrel, which has a thickness of 9 mm. For the detection of photons, multi-channel hybrid photodiodes (HPD) are used. Figure 4.6 gives an overview of the different HCAL subcomponents.

The **HCAL barrel** (HB) covers a region in pseudorapidity of $|\eta| < 1.4$ and comprises 2 304 towers with a segmentation of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$. The HB has 15 layers of brass plates, each with a thickness of 5 cm. The **hadronic outer** (HO) calorimeter samples the amount of energy from hadronic showers, which may leak through the rear of the HB. Including the HO, the effective thickness of the hadronic calorimeter more than 10

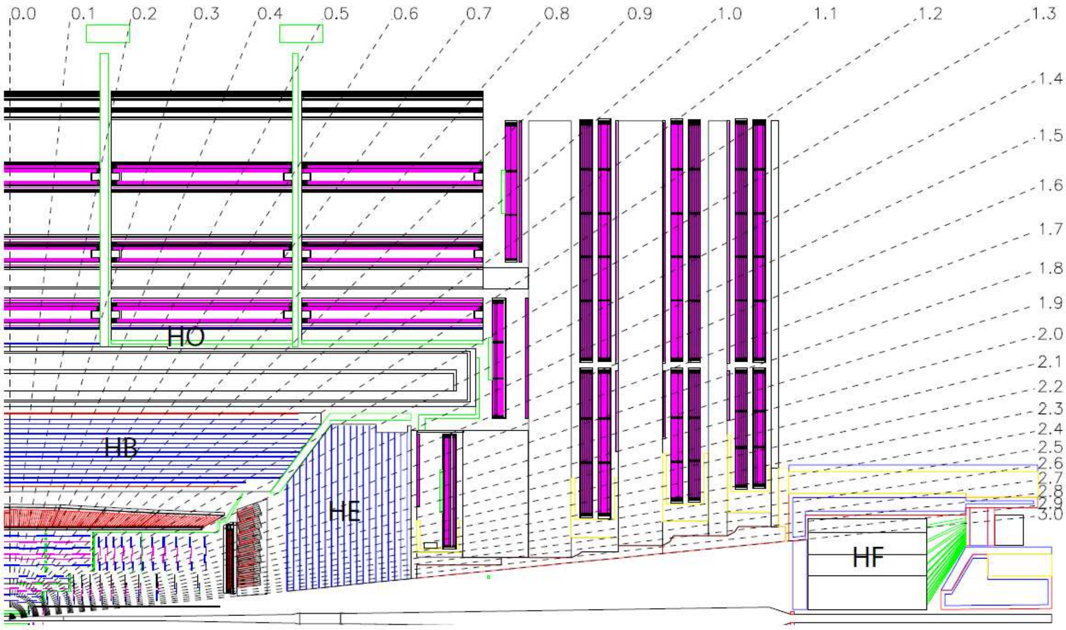


Figure 4.6: The hadronic calorimeter (HCAL) is comprises HCAL barrel (HB) and endcaps (HE), which are hosted inside the magnet coil. An additional layer of scintillators is placed outside the solenoid, referred to as the hadron outer (HO). The hadron forward (HF) detector is used for the online measurement of the instantaneous luminosity in real-time [24].

interaction lengths. The forward region of the HCAL comprises two subdetector systems: the **hadron endcaps** (HE) cover a region of $1.3 < |\eta| < 3.0$ on each side of the detector and contain also 2 304 towers in total. For the **hadron forward** (HF) calorimeter, which covers a region of $2.9 < |\eta| < 5.2$, steel as absorber material and quartz fibres were chosen. The signals, which originate from Cerenkov light emitted in the quartz fibres, are detected by photomultipliers. The HF is located at 11.2 m from the interaction point and is used for the online measurement of the *instantaneous luminosity* in real-time, which is described below in 4.2.7.

4.2.4 Muon System

Muons are identified unambiguously in the muon detectors, as all other particles are stopped in the calorimetry before. Therefore, the reconstruction and identification of muons is an essential requirement of many analysis at hadron colliders. The design and the performance of the muon detectors were one of the driving concepts of the CMS experiment, in order to provide a precise and robust measurement of muons, as described in [24]. Additionally, the charge and the momentum of muons is combined with the measurements from the inner tracking system, based on charged-particle tracks. Muons play a key role in many Standard Model processes, like Drell-Yan production, as well as in the discovery of new physics, for example the decay of a heavy, neutral Higgs boson $H^0 \to Z^0(Z^0)^\star \to \mu^+\mu^-\mu^+\mu^-$.

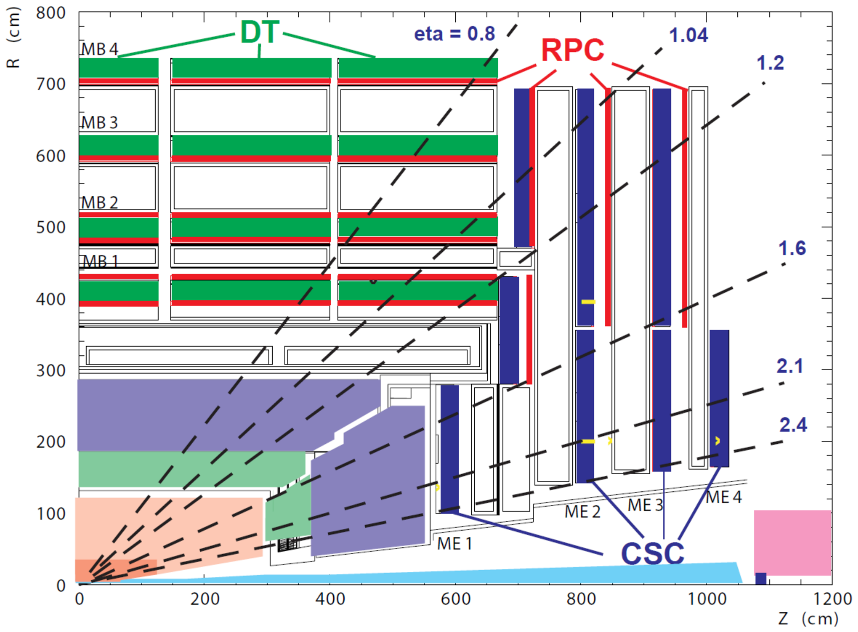


Figure 4.7: The muon system comprises three types of gaseous detectors: drift tube (DT) chambers are used in the barrel region, cathode strip chambers (CSC) in the endcaps, as well as resistive plate chambers (RPC), which are used both in the barrel and endcaps for improved timing resolution [22].

Figure 4.7 gives an overview of the muon system of the CMS detector. There are three types of *gaseous detectors*, which are used for the detection of muons in the CMS experiment: in the barrel region, which covers a range in pseudorapidity of $|\eta| < 1.2$, where a relatively low muon rate is expected and the residual magnetic field is low, **drift tube** (DT) chambers are used [22]. In the two endcaps, which cover a region of $0.9 < |\eta| < 2.4$, much higher muon and background rates are expected and therefore **cathode strip chambers** (CSC) are used. These two detectors are complemented by **resistive plate chambers** (RPC) in both the barrel and endcaps up to a pseudorapidities of $|\eta| < 1.6$. The main feature of the RPCs is their fast response and good timing resolution for operation at high rates. The CSCs/DTs and the RPCs provide two independent sources of information within the first trigger level and form a flexible and robust muon trigger system.

Drift Tubes

Drift tube chambers cover a range of $|\eta| < 1.2$ and comprise 250 chambers, which are arranged in four layers, referred to as *stations*, inside the magnet return yoke at distances of 4.0, 4.9, 5.9 and 7.0 m from the beam axis. The muon stations are arranged in five *wheels*, whereas each wheel is further divided into 12 *sectors*, covering 30° in azimuthal angle. Muons with high transverse momentum cross at least three out of four stations. In the two innermost layers of drift tube chambers, each DT is embedded into two RPCs. In the two outer layers, one RPC is attached to the innermost side of the DT chamber. Therefore, a high-p_{T} muon crosses up to six RPCs and four DTs, resulting into up to 44 measured points as input to the track reconstruction algorithm [22].

Cathode Strip Chambers

The two muon endcaps consist of 468 cathode strip chambers and cover a range between $0.9 < |\eta| < 2.4$. Each CSC has a trapezoidal shape and comprises six gas gaps, which have one plane of radial cathode strips and one plane of anode wires, which are arranged perpendicular to the strips [22]. In order to avoid gaps in the ϕ-plane, CSCs overlap in ϕ. A charged particle traversing a CSC causes ionisation and the subsequent electron avalanche leads to charges drifting to the anode wire and image charges at a group of cathode strips.

Resistive Plate Chambers

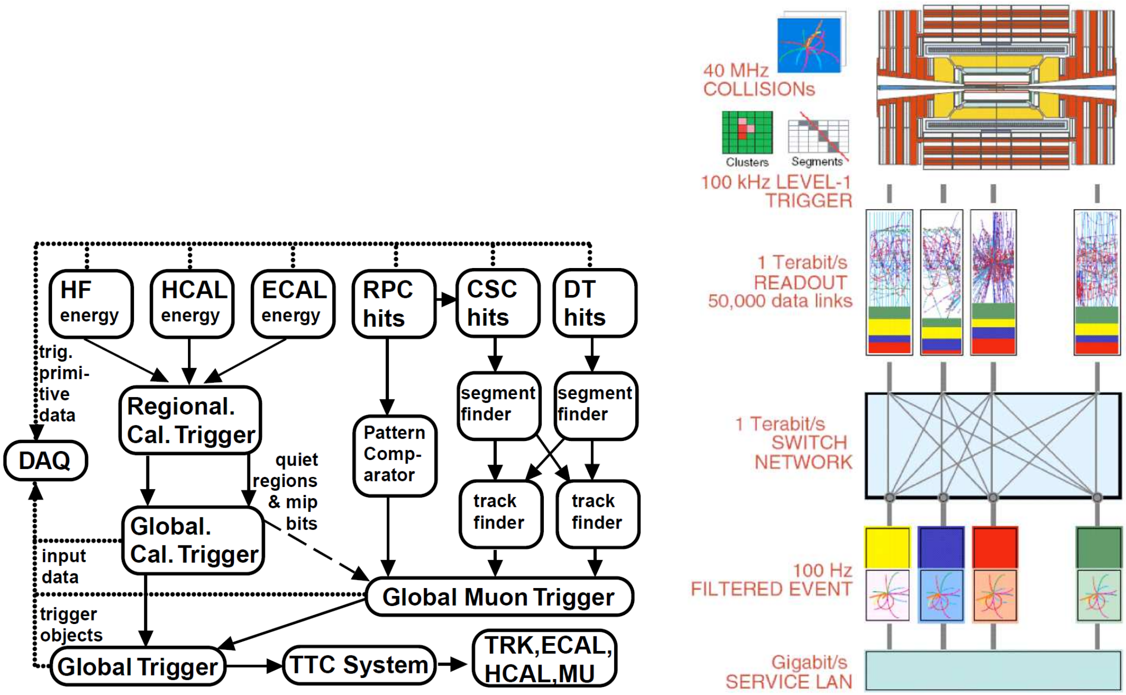
The resistive plate chambers are gaseous parallel-plate detectors, which provide a timing resolution of an ionising event, that is much shorter than the 25 ns of the LHC bunch crossing. Therefore, RPCs are especially dedicated to the triggering of muons and complement the muon system in the barrel and the endcap regions. RPCs cover a range in pseudorapidity up to $|\eta| < 1.6$. RPCs also help to resolve ambiguities in the CSCs [24].

4.2.5 Trigger and Data Acquisition

At the nominal LHC bunch crossing rate, proton-proton collisions are provided every 25 ns, which results in a collision rate of 40 MHz. For each bunch crossing, on average 20 events are expected to occur, which results into a total input rate of the order of 10^9 *interactions per second*. The task of the **online trigger and data acquisition system** is to select the most interesting events at a rate of the order of 10^2 events per second, which are written to archival media (mass storage). Hence, a reduction factor of the order 10^7 is needed. The trigger and data acquisition system comprises custom detector electronics, a read-out network and an online event filter system, which is illustrated in figure 4.8 (b). The reduction of the input rate is implemented in two steps: a **level-1** (L1) trigger system, which is based on custom detector electronics and a **high-level trigger** (HLT) system, which is based on a farm of commercial processors in the online event filter system. A comprehensive description of the CMS trigger system is given in [25].

Level-1 Trigger

The L1 trigger system is hosted in a service cavern underground next to the detector cavern. The minimal processing time, including transit time from the front-end detector electronics to the service cavern, L1 trigger calculation time and return back to the front-end electronics, is 3.2 μs. This time interval includes a total latency of the L1 trigger calculation of less than 1 μs. The L1 trigger system uses a coarser subset of the full event information from the calorimetry, the muon system and some correlated information between both subsystems (*global trigger*). The high-resolution data is stored intermediately in pipelined memory at the front-end electronics, while the L1 trigger decisions are made. A L1 trigger decision interval corresponds to 128 beam crossings at a bunch crossing time of 25 ns. The trigger decision is based on reduced information from the detector output, which is used to build trigger objects, such as muons, electrons and photons, as well as jets and missing energy, called the "*trigger primitives*". Figure 4.8 (a) shows a sketch of



(a) (b)

Figure 4.8: The online trigger and data acquisition system comprises custom detector electronics, a read-out network and an online event filter system. (a) Schematic view of the level-1 trigger system, which is based on custom detector electronics. (b) The read-out network and online event filter system [25].

the L1 trigger decision algorithm, which is described in [25]. The event output rate of the L1 trigger system, which is forwarded to the HLT event filter farm, is about 100 kHz, which corresponds to a reduction factor of the order of 10^3.

High-Level Trigger

Each event in proton-proton collisions, which is kept intermediately in pipelined memory in the detector electronics, has a size of 1.5 MB. Based on the L1 trigger decision, the stored events are transferred from the detector front-end buffers to the event filter units in the online processor farm by a **read-out network**, called *switch*. The use of a software-based HLT, which is implemented on a processor farm, allows for maximal flexibility and benefits from the evolution of the computer technology. The idea of the HLT *reconstruction algorithm* is to use only those objects and regions of the detector, that are actually needed for reconstruction [22]. Finally, after the reconstruction of an event, the full information including also tracking information is forwarded to the storage system and further offline processing. The HLT system provides an additional reduction factor of the order of 10^3 down to 100 - 200 events per second, which is manageable by the mass storage and offline computing systems, which are described in the following.

CMS Computing System

The CMS experiment uses a world-wide, tiered approach for the processing of the data, recorded by the trigger and data acquisition system. The CMS computing system is composed of three layers (*tiers*), which are referred to as Tier-0 (T0), Tier-1 (T1) and Tier-2 (T2). Figure 4.9 shows an overview of the workflow between the CMS computing centres. **Tier-0** is a high-performance computing centre located at the CERN cite (Geneva). Its main purpose is the storage and reprocessing of the '*raw*' data output from the CMS online trigger and data acquisition system, as described above. The **CERN Analysis Facility** (CAF) is also hosted at CERN and provides large computing resources for the purpose of prompt reconstruction and prompt calibration of the recorded data. In the following, the raw data samples are distributed throughout the next layer of computing centres, the **Tier-1** centres. The T1 centres are seven large computing centres, hosted at CMS collaborating countries world-wide. Each of them receives a subset of the data samples, produced at T0. They also provide a tape archive for the *full event data format* (raw and reconstructed data). They also provide substantial CPU power for further reprocessing of the data sets and data-intensive analysis tasks, as well as a second secure copy of the raw data. Finally, the **Tier-2** centres provide computing resources for user analysis and the production of simulated data samples (*Monte Carlo* production).

4.2.6 Data Quality Monitoring

The data quality monitoring (DQM) system provides an unique tools for the assessment of the status of the detector subsystems in real-time (**online DQM**), as well as the monitoring and certification of the reconstruction and calibration of the recorded data (**offline DQM**). An overview of the DQM workflow in shown in figure 4.10. The tools for the creation of control histograms are called *DQM applications*. The histograms of

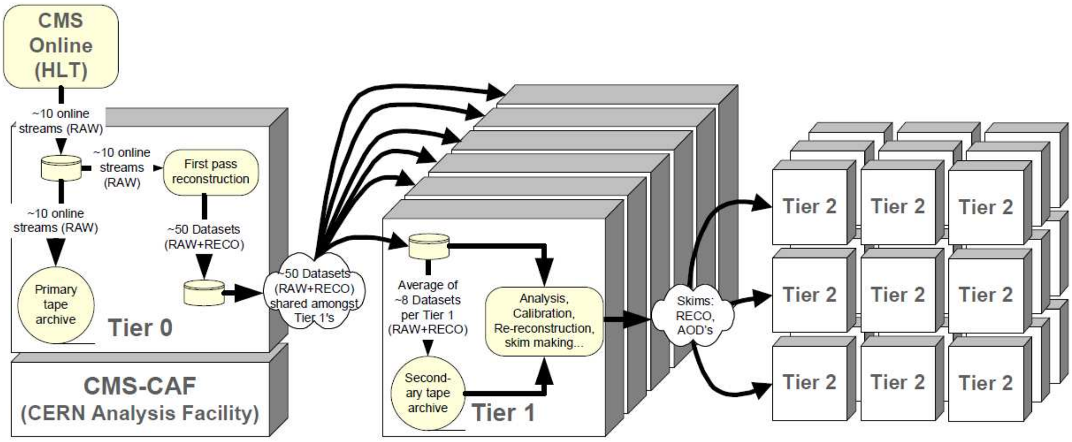


Figure 4.9: Overview of the workflow between the CMS computing centres: the CMS collaboration uses a world-wide, tiered approach for the processing of the data, recorded by the trigger and data acquisition system. The CMS computing system is composed of three layers, which are referred to as Tier-0, Tier-1 and Tier-2 centres [24].

the online and offline DQM are visualised by one centralised, web-based *graphical user interface* (GUI). The results of the certification procedure, which is carried out in daily shift operations, is stored in a central certification data base, which is referred to as the *run registry*. Based on this certification procedure, a list of certified, good runs is created and defines the input for sophisticated, offline physics analysis, as presented in this thesis. A report of the operations of the DQM system is given in [26].

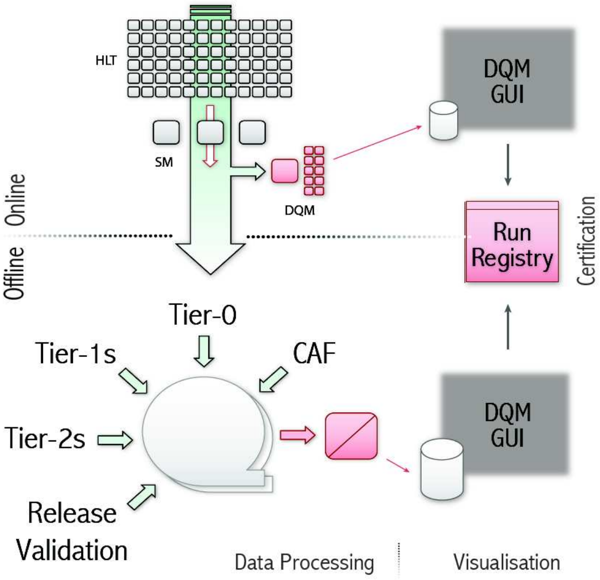


Figure 4.10: Overview of the workflow of the data quality monitoring system [26].

The purpose of the **online DQM** is the *real-time control* and monitoring of the CMS detector. The monitoring is performed in shift operation, for 24-hours coverage on each day. For the online DQM, a subset of the event data from the HLT output stream is processed in parallel by several independent DQM applications, each corresponding to one singular or more subsystems, such as the CSCs or silicon strip tracker (SiStrip). The DQM applications create control histograms in real-time during the data taking, which are called *monitor elements*. These monitor elements are assessed via the DQM GUI. The GUI plays a central role in the DQM procedure. It allows for the monitoring of the detector status in real-time in the online DQM, as well as for the validation of reconstructed objects in the offline DQM. It provides a compact overview of the status of all subsystems and provides histograms from actual data taken, as well as from archived data taking periods. The **offline DQM** enables the possibility to monitor *reconstructed physics objects*, such as muons, photons or jets, with a relatively short latency close to the data taking. The aim is to provide fast feedback of possible problems, due to the reconstruction procedure. An anew evaluation of the detector components with full event statistics of a given data taking period is also performed.

An overview of the **prompt data monitoring** and validation for a top quark-specific event topology with the early data, recorded by the CMS detector, is given in [27]. The aim is to monitor rates, topology and trigger efficiencies for top-like dilepton events in the online and offline DQM, as well as the prompt validation of dilepton object reconstruction in the offline DQM sequence.

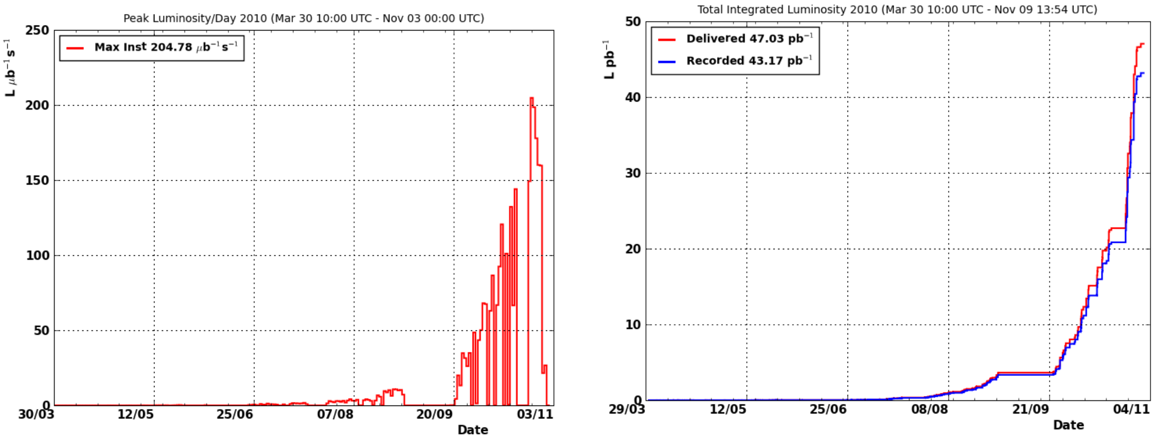
4.2.7 Luminosity Measurement

The measurement of the luminosity at the CMS experiment is an important component in order to monitor the performance of the LHC in real-time. It also provides the *overall normalisation* for most physics analysis, such as the measurement of cross sections. The instantaneous luminosity of the LHC is determined from the **beam size**, measured in horizontal and vertical separation scans (*Van Der Meer scans*), and the **beam current**. The total systematic error of the luminosity measurement is 3.6%, which is described in [28] in further detail. An overview of the measurement of the luminosity is given in [29]. Figure 4.11 (a) shows the instantaneous luminosity $\mathcal{L}$ as a function of time for the complete data taking period in 2010, whereas the corresponding *integrated luminosity* $L = \int dt\, \mathcal{L}$ is shown in 4.11 (b).

Online and Offline Methods

The CMS **online luminosity** measurement uses the hadron forward (HF) calorimeter in order to measure the instantaneous luminosity in real-time, which is referred to as the **online HF** method. Therefore, two methods are implemented into the HF firmware. The first method uses the average fraction of empty calorimeter towers in order to derive the mean number of interactions per bunch crossing. The second method assumes a linear relationship between the average transverse energy $\mathrm{E_T}$ per tower and the luminosity. Both methods are also used to estimate residual background effects due to beam-gas and beam scraping interactions.

In addition to the online HF measurement of the luminosity, two **offline methods** are applied as a cross-check to the online method. In general, the offline methods have a significantly longer latency, due to the prompt reconstruction. On the other hand, they provide a better background rejection of beam-gas and non-collision events. The **offline**



(a) (b)

Figure 4.11: (a) Instantaneous luminosity $\mathcal{L}$ as a function of time and (b) the integrated luminosity $L = \int dt\, \mathcal{L}$ for the complete data taking period in 2010.

HF method exploits the energy depositions in the HF and is based on the coincidence of the total sum of transverse energy ($\Sigma\,\mathrm{E_T}$) of at least 1 GeV in the HF *on both sides* of the CMS detector. An additional timing requirement is applied, in order to reduce background from non-collision events. The second method, called the **vertex method**, is based on information from tracking and vertex finding and provides an orthogonal set of systematic uncertainties, which complements the online and offline HF methods. At least one reconstructed vertex with at least two associated tracks is required. The vertex method provides a good efficiency for minimum bias events and is highly suppressive against non-collision background events [29].

Normalisation and Van Der Meer Scans

In order to obtain the total luminosity, which is needed for physics analysis, the luminosity measured by the online and offline methods is normalised to an absolute scale. The scale is provided by a separation scan method, referred to as *Van Der Meer* (VdM) scan, which is named after Simon van der Meer. This scale is used for the determination of the **absolute luminosity**. The aim of this method is to measure the size and the shape of the interaction region as a function of the transverse beam separation. This information is used to determine the **beam profile** $F(x, y) = f_x(x)f_y(y)$ transverse to the beam direction, which is described in [29]. Finally, the absolute luminosity is calculated from the beam size, derived from the VdM scan, and the beam current. The total systematic uncertainty of the luminosity measurement yields a relative error of 3.6%, as described in [28].

Chapter 5

Simulation and Event Reconstruction

When you're a Jet, you're a Jet all the way.
West Side Story

An event is the result of a single proton-proton collision inside the CMS detector or a simulated collision event. **Monte Carlo generators** are used to simulate events, such as top-antitop quark pair production. The calculation of the production process originating from an inelastic proton-proton scattering and the subsequent decay into final states of *stable* quarks and leptons, is called the *generation* step. The response of the CMS detector is simulated in the *simulation* step. This information is converted into a data format, which is similar to the real detector output (RAW data), referred to as the *digitisation* step. Figure 5.1 shows an illustration of the *full detector simulation* before the reconstruction. The raw data format, which originates from the full detector simulation (containing generation, detector simulation and digitisation) or from the real detector output, is committed to the *reconstruction* step. The aim of the **event reconstruction** is to determine the detector response, reconstruction efficiencies, as well as the resolution of reconstructed objects.

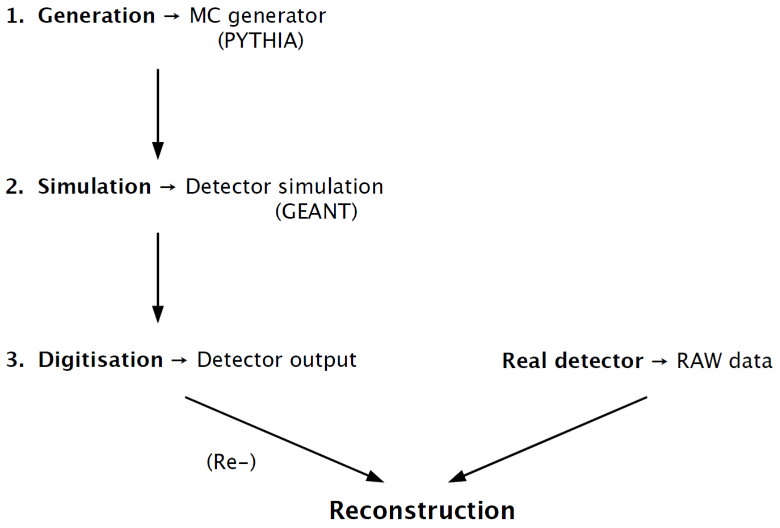


Figure 5.1: The generation and reconstruction of a Monte Carlo simulated event.

5.1 Monte Carlo Simulation

Different Monte Carlo (MC) generators are used in order to simulate signal and background processes for physics analysis. The correct theoretical understanding of the signal production process and all relevant background processes is a key ingredient to the interpretation of the result of the measurement. For the **generation** of simulated data samples, which are used for the analysis presented in this thesis, the event generator PYTHIA6 is used [6]. Additional *higher-order radiation effects* are calculated by the generator MADGRAPH [30] on matrix element level and are further processed by PYTHIA6 for the simulation of the full parton showering process. Finally, for the correct treatment of the decay of τ leptons in several final states, the TAUOLA decay library is used [31].

PYTHIA6

The main task of MC generators is to provide a complete description of the event structure at collider experiments. Therefore, PYTHIA starts with the calculation of the matrix element of an inelastic proton-proton scattering process, called the *hard process*, at the *leading-order* (LO) of the strong coupling constant α_{s}. In terms of initial state and final state objects ($A \rightarrow B$), PYTHIA is optimised for the calculation of processes with two initial partons resulting into one or two partons in the final state, referred to as ($2 \rightarrow 1$) or ($2 \rightarrow 2$) processes. In general, the cross section at LO of any process with two initial partons i, j and one final parton k is given by equation 2.23, where f_i^1 and f_j^2 are the parton density functions of the colliding particles 1 and 2, introduced in section 2.4, and $\widehat{\sigma}_{ij \rightarrow k}$ is the cross section of the hard process ($ij \rightarrow k$) calculated in LO of perturbative QCD [6]. In the following, several **higher-order corrections** are applied to the LO approximation, in order to obtain a complete description of the process.

The initial and final state partons give rise to the emission of bremsstrahlung, which is referred to as *initial state* and *final state radiation* (ISR/FSR). The higher-order corrections are obtained either from the **matrix element method**, where additional Feynman diagrams are calculated order by order, or from perturbative calculations, such as **parton showering**, which is a key feature of PYTHIA. Using the parton showering method, multijet events are generated from branchings of one parton into two ($a \rightarrow bc$). The branching process is described by the *splitting kernels* $P_{a \rightarrow bc}(z)$, where the daughter parton b has the energy fraction z of its mother parton a. In general, the separation of radiation into ISR and FSR showers is arbitrary. Using parton showers, in order to model the perturbative corrections to the LO approximation, is preferred, because corrections are relatively simple to calculate, compared to the matrix element method. The LO cross section is referred to as the *Born cross section* σ_{Born}. By considering one additional virtual or real emission, the **next-to-leading order** (NLO) cross section σ_{NLO} is obtained by the following equation:

$$\sigma_{\text{NLO}} \;=\; \sigma_{\text{Born}} \;+\; \sigma_{\text{virtual}} \;+\; \sigma_{\text{real}} \;. \tag{5.1}$$

Further corrections have to be applied due to the **hadronisation process**. As already introduced in section 2.3.3, the string fragmentation model is used, in order to approximate the long-range confinement of the strong force (LUND **string model**). Combining the string fragmentation model and the perturbative corrections obtained from parton shower models, an almost complete description of hard interactions at hadron colliders is obtained.

MadGraph

The MadGraph/MadEvent MC event generator (MadGraph) is used for the calculation of higher-order corrections to the hard process on matrix element level, as described above [30]. Therefore, additional radiative processes (bremsstrahlung) are calculated at tree-level and further processed with the full parton showering and hadronisation infrastructure of a general-purpose event generator, such as Pythia6. It should be noted, that this procedure does *not* provide a complete NLO description of the hard process, which also requires the calculation of loop diagrams. An important aspect of the method used by MadGraph is the *correct matching* of a calculated higher-order matrix element to a parton shower, in order to avoid double-counting.

Tauola

The Tauola MC package is an universal interface for the processing of the decay of τ leptons, as described in [31]. This package incorporates a substantial amount of results from high-precision τ lepton measurements. The universal interface of Tauola requires the general-purpose MC generator, such as Pythia6, to produce *stable* τ leptons. The event content, which is provided by the MC generator, is searched through for all stable τ leptons and neutrinos. It is required, that a τ-flavoured pair of particles ($\tau^+\tau^-$ or $\tau\nu_\tau$) originates from the same *mother* particle. Finally, the decay of the pair is performed by Tauola.

5.2 Object Reconstruction

For each event, either from proton-proton collisions or cosmic muon data taking, the CMS detector provides a set of detector signals. In order to perform an analysis, the raw data have to be converted into classes, which are referred to as physical objects, e.g. electrons, photons or jets. This step, the transformation of raw-level data into reconstructed physics objects, is called *reconstruction*. The process of reconstruction uses the information of one or more subdetectors of the CMS experiment to build different classes of *candidates*. Figure 5.2 shows a slice through the CMS detector, which illustrates the use of the different subdetectors for the reconstruction of the different particle candidates.

In the first step of the reconstruction, hits in the inner (tracker) or outer tracking (muon) systems are combined, in order to form **charged-particle tracks**, as well as the clustering of single-cell energy deposits into **calorimeter energy clusters**. In the following, this information is transformed into physics object candidates, like electrons or muons. **Neutral particles**, like photons or neutral hadrons, are reconstructed from an ECAL or HCAL clusters. **Charged particles** consist of the combination of a reconstructed track and a calorimeter cluster. For example, electrons are expected to have a track pointing towards the direction of a cluster in the ECAL, without any matching energy deposit in the HCAL or any reconstructed track in the muon system. Muons are expected to traverse the calorimetry without significant loss of energy. Therefore, only a minimal energy deposit in the HCAL or ECAL is required for the reconstruction of a muon candidate. The reconstruction of charged hadrons uses information from all subdetectors, except for the muon system.

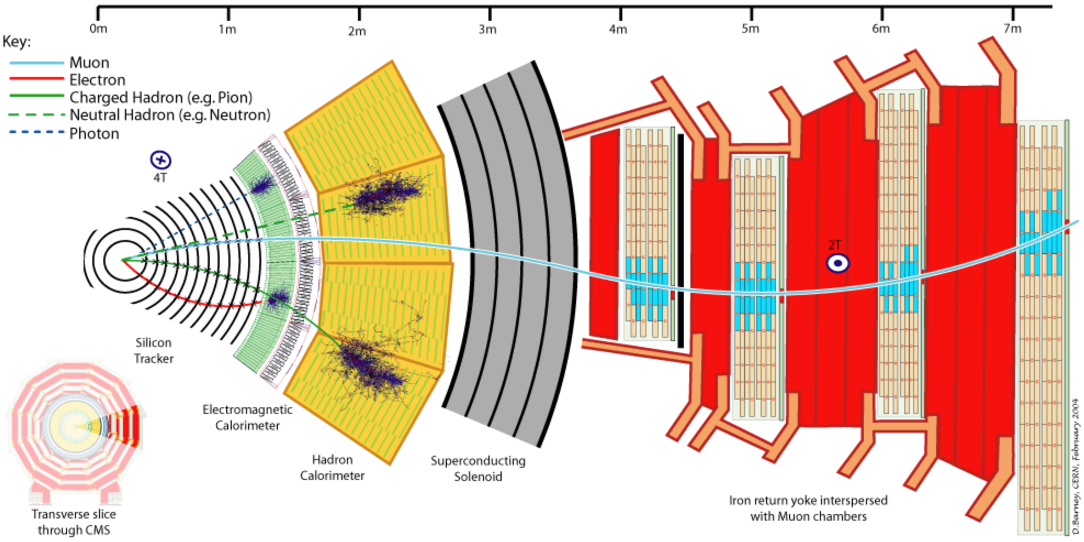


Figure 5.2: Slice through the CMS detector. It is illustrated, how the different subdetectors are used for the reconstruction of the different particle candidates [32].

5.2.1 Muon Reconstruction

Muon Resolution

The quality of the reconstruction of muon candidates depends on the resolution and efficiency of the used subdetectors. The muon momentum is calculated from its measured bending angle, either individually from the inner tracking system or the muon system alone, or a combination of both tracking systems. Figure 5.3 shows the momentum resolution for muons in the barrel region $|\eta| < 0.8$ (left) and the endcap regions $1.2 < |\eta| < 2.4$ (right).

The inner tracking system provides the best momentum resolution for low p_{T} up to several hundred GeV in both regions of the muon system, which is shown in figure 5.3. In the central part of the detector, the muon momentum resolution is improved by combining information from the tracker and the muon detectors for muon momenta of $p_{\mathrm{T}} \gtrsim 200$ GeV/c. In the endcap region, the inner tracking system is dominating the momentum resolution up to momenta in the TeV range. Reconstructed tracks from the muon system provide an unambiguous identification of muons, whereas the inner tracking system improves the measurement of the muon momentum significantly. From that fact, three different **muon reconstruction algorithms** arise, which are commented in the following subsection. It should be noted, that the overall muon momentum resolution $\frac{\Delta p}{p}$ is below 5% for muon momenta up to $p_{\mathrm{T}} \lesssim 500$ GeV/c and below 2% for $p_{\mathrm{T}} \lesssim 200$ GeV/c in the central detector.

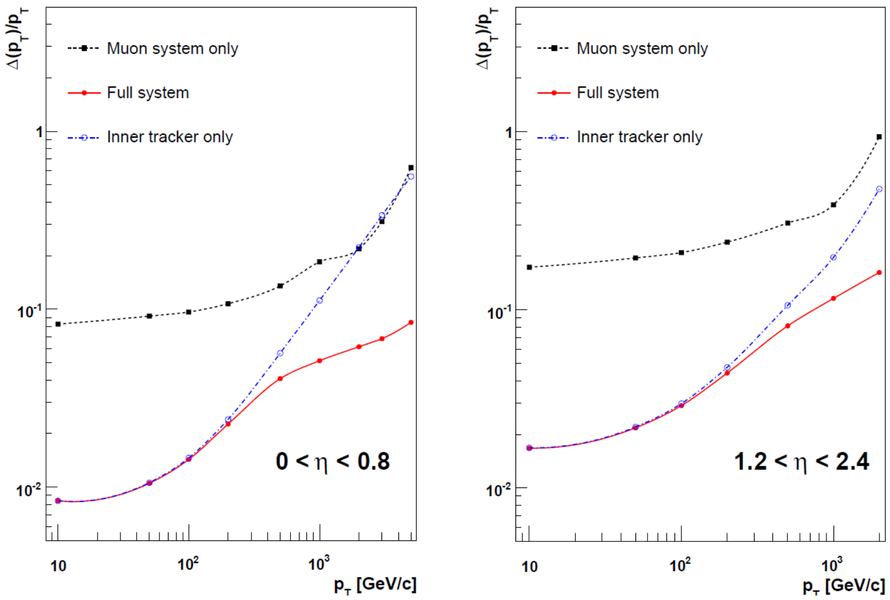


Figure 5.3: Muon transverse momentum resolution for the barrel region $|\eta| < 0.8$ (*left*) and the endcap regions $1.2 < |\eta| < 2.4$ (*right*). In both regions of the muon system, the inner tracking system is providing the best momentum resolution for low p_T up to several hundred GeV [24].

Reconstruction Algorithms

Muons are reconstructed using reconstructed tracks from the muon detectors alone or combining them with reconstructed tracks from the silicon tracker. Tracks are reconstructed independently in the inner tracking system (tracker tracks) and in the muon system (stand-alone muon tracks) [33]. The concept of regional or local reconstruction is used for muon reconstruction. Seeds in the muon system define regions of interest, where a local reconstruction is performed, such as the software algorithm only reconstructs charged-particle tracks compatible with hits in the muon chambers. **Stand-alone muon** reconstruction only uses information from the muon system, whereas **global muon** reconstruction also uses information from the inner tracking system. The latter method, which includes tracker tracks, is subdivided into an outside-in (*global muon*) and an inside-out (*tracker muon*) approach.

The global muon reconstruction starts with a locally reconstructed stand-alone muon in the muon system. The muon trajectory is extrapolated back onto the surface of the outer silicon tracker volume, which is therefore referred to as an *outside-in* approach. If a matching tracker track is found, a *global muon track* is fitted, which combines tracker and stand-alone muon tracks. As mentioned above, for muon momenta up to $p_\mathrm{T} \lesssim 200$ GeV/c, the momentum resolution is determined by the tracker fit. The *inside-out* approach reconstructs tracker muons, which use tracker tracks as a seed and are extrapolated to the muon system, taking into account the expected energy loss traversing the calorimetry and solenoid coil. If at least one matching muon segment is found, the tracker track is considered as *tracker muon track*.

Muon Identification

The standard muon reconstruction algorithms contain *by design* a basic muon identification, given that any muon candidate track found in the inner or outer tracking detectors has to be compatible with a muon hypothesis. In addition to that, a further **muon quality and identification** (muon ID) selection is applied to reconstructed muon candidates in order to provide additional information for each muon. The **muon track quality** comprises the *number of hits* associated to the candidate track (stand-alone, tracker or global track), the χ^2 of the *track fit* and the *transverse impact parameter* with respect to the primary vertex [33].

Muon Isolation

Muon candidates, which have been successfully identified and reconstructed, are expected to be real muons. At this selection step, most of the muons originate from the leptonic decays of heavy quarks, such as b and c quarks. There are also contributions from muonic kaon and pion decays [22]. Such muons are usually accompanied by other particles from the fragmentation process and the subsequent decays, which form jets. High-p_T muons, which originate from heavier particles, like W and Z bosons, are expected to have no such activity in their environment. The absence of particles in the vicinity of the muon is defined as **isolation**. The definition of isolation allows for the separation of muons originating from the decay of light hadrons and those from the leptonic decay of vector bosons or heavier objects. Different muon isolation algorithms compare the amount of calorimeter energy deposits and charged-particle track momenta in a *defined cone* around the muon direction to the momentum of the muon itself. The energy deposited by the muon and its transverse momentum is not considered for the determination of the isolation variables, referred to as *veto cone*, which is illustrated in figure 5.4.

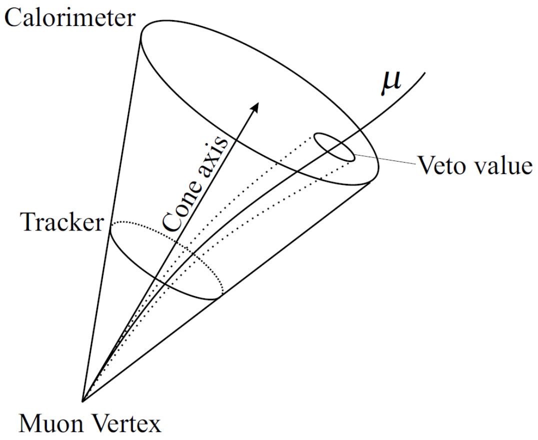


Figure 5.4: Muon isolation algorithms compare the amount of calorimeter energy deposits and charged-particle track momenta in a defined cone around the muon direction to the momentum of the muon itself. The energy deposited by the muon and its transverse momentum is not considered for the determination of the isolation variables, called veto cone [24].

The **combined isolation** is defined as:

$$I_{\mathrm{comb}} \;=\; I_{\mathrm{ECAL}} + I_{\mathrm{HCAL}} + I_{\mathrm{Trk}} \;,$$

where the *calorimeter isolations* I_{ECAL} and I_{HCAL} are the sum of transverse energy $\mathrm{E_T}$ in the ECAL and HCAL in a cone of size $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.3$ in $(\eta - \phi)$ space around the muon direction. Respectively, the *tracker isolation* I_{Trk} is defined as the sum of transverse momentum p_{T} within a cone $\Delta R < 0.3$. The track of the muon, as well as the energy in the ECAL and HCAL associated to the muon are excluded from the computation of the isolation [33]. In order to enhance the separation power of the isolation requirement between signal and background, the isolation is normalised to the p_{T} of the muon. This quantity is referred to as the **combined relative isolation**, which is defined as:

$$I^{\mathrm{rel}}_{\mathrm{comb}} \;=\; \frac{I_{\mathrm{comb}}}{p_{\mathrm{T}}(\mu)} \;=\; \frac{I_{\mathrm{ECAL}} + I_{\mathrm{HCAL}} + I_{\mathrm{Trk}}}{p_{\mathrm{T}}(\mu)} \;.$$

5.2.2 Electron Reconstruction

In contrast to muons, electrons do not have a unique signature in an exclusive subsystem, like the muon system. Hence, there is no inherent electron identification embedded in their reconstruction algorithm. Therefore, the rate of wrongly reconstructed electrons is relatively higher compared to muons. A reconstructed electron consists of the combination of a charged-particle track in the inner tracking system originating from the IP and a corresponding energy deposition in the ECAL. The reconstruction of electrons is compromised by the fact, that there is a large amount of tracker material between the IP and the ECAL. By traversing the tracker material, electrons radiate a considerable amount of bremsstrahlung, which causes a spread of the energy localisation in the ECAL crystals in the ϕ-direction. The average amount of tracker material traversed by the electrons varies between $0.3\,X_0$ in the barrel region and up to $1.5\,X_0$ in the overlap region around $|\eta| = 1.5$.

Electron Energy Resolution

The quality of the electron reconstruction and the correct determination of its energy mainly depend on the energy resolution of the ECAL. The **energy resolution** of an ECAL supermodule for incident electrons $\frac{\sigma_{\mathrm{E}}}{E}$ has been measured with a test beam. The result is shown in figure 5.5. The energy resolution is below 1% for measured energies of greater than 20 GeV. It is parametrised as a function of energy, considering a stochastic term S, a noise term N and a constant term C, as described in [22]:

$$\left(\frac{\sigma_{\mathrm{E}}}{E}\right)^2 \;=\; \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \;.$$

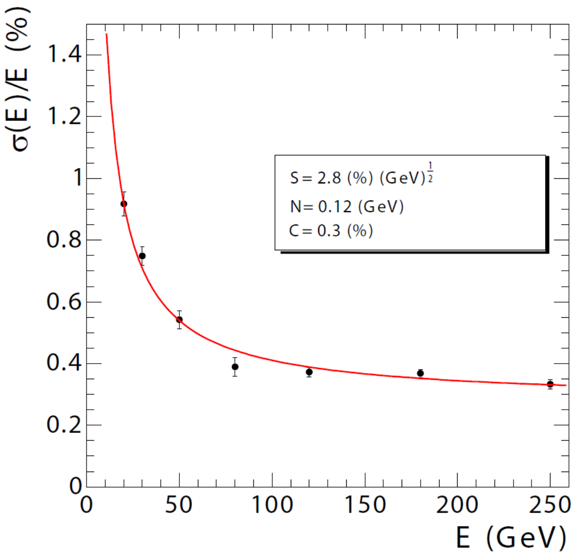


Figure 5.5: Energy resolution of an ECAL supermodule for incident electrons $\frac{\sigma_E}{E}$, which has been measured with a test beam [24]. The energy resolution is parametrised as a function of energy, considering a stochastic term S, a noise term N and a constant term C, as described in [22].

Electron Clustering

An electromagnetic shower caused by a single electron or photon deposits its energy in several crystals in the ECAL [22]. In an ideal environment, about 94% of the electron energy is contained in an array of 3 × 3 crystals, while 97% is clustered in an 5 × 5 array. A single elementary cell of ECAL crystals is referred to as a *basic cluster*. In the real experiment, electrons traverse the silicon tracker first and radiate bremsstrahlung. The emission of **bremsstrahlung** leads to a widening of the electron energy distribution in the ECAL along the track curvature, perpendicular to the magnetic field. Hence, the energy deposit is spread in ϕ-direction. The distribution of the electron energy among a *cluster of clusters* due to bremsstrahlung photon emission is called "*supercluster*", illustrated in figure 5.6.

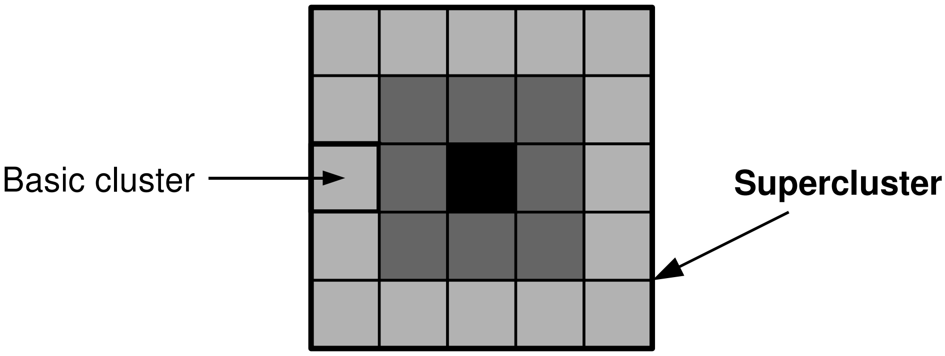


Figure 5.6: An electromagnetic supercluster is composed of an 5 × 5 array of ECAL basic clusters.

An **electromagnetic supercluster** consists of an 5×5 array of basis clusters and ideally contains the complete energy from the emitted photons, which is expressed as $\sum E^{\gamma}_{\text{brem}}/E^{\text{e}}$. Figure 5.7 shows the distribution of bremsstrahlung emission as a fraction of the initial electron energy. In order to collect this energy spread in ϕ-direction, the *hybrid* and *island superclustering algorithms* are used. The hybrid algorithm is designed to reconstruct relatively high-E_T electrons in the barrel, while the island algorithm is better suited for low-energetic electron reconstruction in the endcap region [22]. The superclustering algorithms require a basic cluster as *seed cluster* above an energy threshold of $E_{\text{T}}^{\text{seed}} > 1$ GeV. Those seed clusters initiate a dynamical recovery procedure of the lost photon energy.

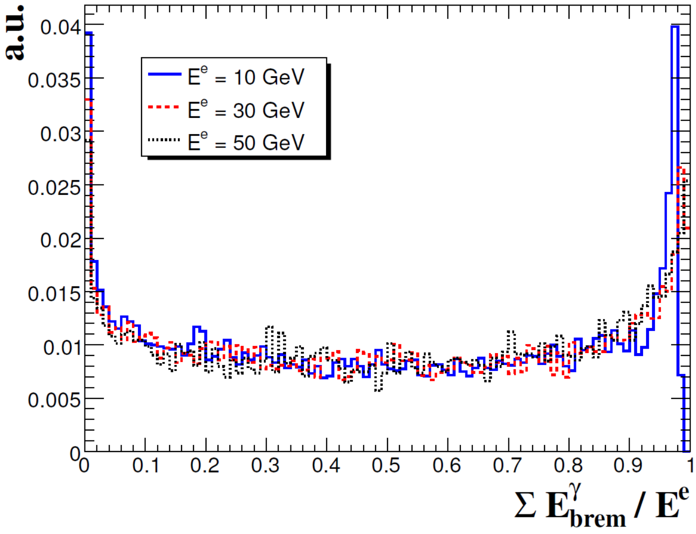


Figure 5.7: Distribution of bremsstrahlung emission as a fraction of the initial electron energy. In order to collect this energy spread in ϕ-direction, two superclustering algorithms (*hybrid* and *island*) are available. [34].

Electron track reconstruction

The reconstruction of an ECAL supercluster initiates a supercluster-driven track reconstruction. The supercluster-driven seed finding method first searches for two hits in the pixel detector. It is assumed, that the energy weighted average impact point of the electron and its associated bremsstrahlung photons coincides with the impact point of an ideal, non-radiating electron of the same momentum [22]. The energy-weighted mean position of the supercluster is extrapolated backward towards the innermost layer of the pixel detector. If a compatible hit is found within a loose $\Delta\phi - \Delta z$ window, the predicted trajectory of the found hit and the supercluster position is propagated towards the next pixel layer, in order to search for a second pixel hit within a narrower $\Delta\phi - \Delta z$ window. The two hits found in the pixel detector are used as input for a track finding algorithm in the silicon strip tracker. Tracks are reconstructed using a Gaussian Sum Filter, referred to as *GSF tracks*. In contrast to a simple Kalman Filter (KF), which uses a global least-square minimisation method, based on a linear model for the track evolution, the GSF uses a non-linear filter approach, which also accounts for non-Gaussian fluctuations.

In this way, a better description of the propagation of the electron is achieved. A more detailed description of the electron track reconstruction is given in [34]. In summary, the track reconstruction procedure comprises four modular components:

1. **Seed generation**: initial track (*seed*) finding
2. **Trajectory builder**: building all possible trajectories for a given seed
3. **Trajectory cleaner**: solving ambiguities among possible trajectories
4. **Trajectory smoother**: final track fit (backward fit)

Electron Classification

As an example, approximately 35% of the electrons radiate about 70% of their initial energy before reaching the ECAL. In nearly 10% of all cases, this fraction is above 95%. According to the amount of radiation losses in the inner tracker, electrons are grouped into disjoint **electron classes**, introduced in [34]. Therefore, several tracking and calorimetry variables are introduced, which measure the amount of radiated energy. The first variable, which measures the amount of bremsstrahlung f_{brem}, is defined as the relative difference of the electron momentum at the origin p_{in} and the momentum at the tracker surface p_{out}:

$$f_{\mathrm{brem}} = \frac{p_{\mathrm{in}} - p_{\mathrm{out}}}{p_{\mathrm{in}}} .$$

Furthermore, the total energy of the supercluster E_{SC} should not differ significantly from the inner track momentum, measured as $\frac{E_{\mathrm{SC}}}{p_{\mathrm{in}}}$. Also a matching between the supercluster position in ϕ and the reconstructed track is used. The first category represents *low-radiative* electrons, also referred to as "*golden electrons*". Electrons in this class have a bremsstrahlung fraction below 20% ($f_{\mathrm{brem}} < 0.2$) and their corresponding supercluster is constituted by a single seed cluster. Also the $\frac{E_{\mathrm{SC}}}{p_{\mathrm{in}}}$ value is greater than 0.9. The next two categories are referred to as *big brem* and *narrow* electrons. They contain a good energy-momentum matching $0.9 < \frac{E_{\mathrm{SC}}}{p_{\mathrm{in}}} < 1.1$, despite a large bremsstrahlung fraction of $f_{\mathrm{brem}} > 0.5$. The remaining electrons are grouped into the final category of *showering electrons*. Electrons from this class have the largest fraction of bremsstrahlung loss and a bad matching between supercluster energy and track momentum.

Electron Isolation

As already introduced in the context of muon reconstruction, lepton isolation is a simple and powerful tool in order to suppress background from QCD processes, like the production of *prompt* electrons from leptonic decays of b and c quarks, as well as electrons misreconstructed from hadronic jets. As a first preselection step, in order to separate *real* electrons from non-genuine or "*fake*" electrons, also emerging from early photon conversions in the tracker material, a simple **track-based electron isolation** is applied. Requiring only track isolation also avoids complications due to severe bremsstrahlung or photon conversions [22]. The electron track isolation is defined as the sum of transverse track momenta $p_{\mathrm{T}}^{\mathrm{track}}$ in a fixed cone of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.3$ in $(\eta - \phi)$ space

around the electron candidate, excluding the electron track itself. The sum is normalised to the electron momentum: $\sum \left(p_{\mathrm{T}}^{\mathrm{track}}\right) / p_{\mathrm{T}}^{\mathrm{e}}$.
In the event selection, presented in section 6, the electron isolation is defined analogously to the muon isolation defined above as a **relative combined isolation** separately for the barrel and endcap regions, taking also into account the relative calorimeter isolation. For the reconstruction of electrons, also a requirement on the *impact parameter* with respect to the beam spot is applied. In the following, track-based electron isolation is complemented by electron identification requirements.

Electron Identification

The electron identification makes use of the full set of separating variables. It is expected, that the performance of electron identification, in terms of efficiency and purity, depends on the preselection, such as classification and isolation. As a first selection step, electrons are separated into different classes, based on an energy-momentum matching between the reconstructed electron track and the corresponding supercluster. Also a loose track isolation of $\sum p_{\mathrm{T}}^{\mathrm{track}} / p_{\mathrm{T}}^{\mathrm{e}} < 0.5$ with cone size $R = 0.35$ is applied, which significantly reduces the background from QCD *fake* electrons, as described in [34]. In the following, further **electron quality variables** are introduced, in order to reconstruct electron candidates with optimal purity:
A geometrical matching between the *extrapolation* of the track position at the primary vertex to the ECAL $\eta_{\mathrm{in}}^{\mathrm{extrap.}}$ and $\phi_{\mathrm{in}}^{\mathrm{extrap.}}$ to the energy-weighted position of the supercluster η_{SC} and ϕ_{SC} is performed, as described in [34]. The resulting identification variables, labelled as $\Delta\eta_{\mathrm{in}}$ and $\Delta\phi_{\mathrm{in}}$, are defined as:

$$|\Delta\eta_{\mathrm{in}}| = |\eta_{\mathrm{SC}} - \eta_{\mathrm{in}}^{\mathrm{extrap.}}| \ ,$$

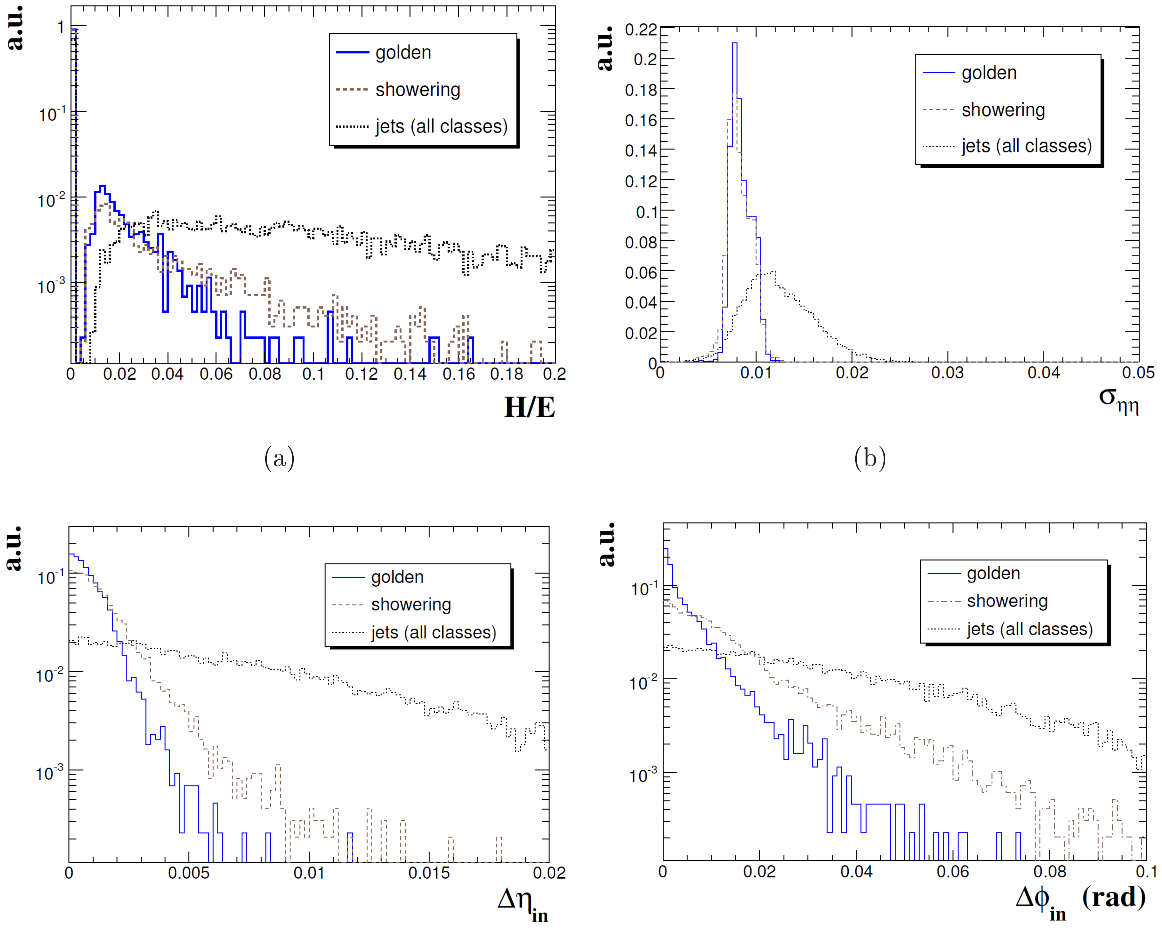
$$|\Delta\phi_{\mathrm{in}}| = |\phi_{\mathrm{SC}} - \phi_{\mathrm{in}}^{\mathrm{extrap.}}| \ .$$

Electrons are expected to have no significant energy deposited in the HCAL, therefore an upper limit on the fraction of **energy in the HCAL tower** just behind the seed cluster is required. The corresponding variable is the ratio H/E, which is expected to be $H/E < 0.2$ for well identified electron candidates. In order to further discriminate the jet background, electron **shower shape** variables are used. For each supercluster with a given seed cluster s and crystals i, the shape variables $\sigma_{\eta\eta}$ and $\sigma_{\phi\phi}$ are defined as [34]:

$$\sigma_{\eta\eta} = \sum_{\text{crystals } i} (\eta_i - \eta_s)^2 \, E_i \cdot \frac{1}{E_s} \ ,$$

$$\sigma_{\phi\phi} = \sum_{\text{crystals } i} (\phi_i - \phi_s)^2 \, E_i \cdot \frac{1}{E_s} \ .$$

Due to effects from bremsstrahlung, particularly the ϕ projection appears distorted. Therefore, only the shape variable $\sigma_{\eta\eta}$ is used for the electron selection in the analysis, described in chapter 6. All relevant electron identification variables are shown for different electron classes in figures 5.8 (a) - (d).



(a) (b)

(c) (d)

Figure 5.8: **Electron identification variables**: As a first selection step, electrons are separated into different classes, labelled as *golden* and *showering*, based on an energy-momentum matching between the reconstructed electron track to the corresponding supercluster. Electrons are expected to have no significant energy deposited in the HCAL, therefore an upper limit on the fraction of energy in the HCAL tower just behind the seed cluster is required. The corresponding variable is the ratio H/E, shown in (a). In order to further discriminate the jet background, electron shower shape variables are used. As an example $\sigma_{\eta\eta}$, is shown in (b). A geometrical matching between the extrapolation of the track position at the primary vertex to the ECAL to the energy-weighted position of the supercluster is performed. The resulting identification variables, labelled as $\Delta\eta_{\rm in}$ and $\Delta\phi_{\rm in}$, are shown in (c) and (d) [34].

Cut-Based Electron Selection

Several electron identification (ID) and reconstruction requirements have been discussed above. In order to select well identified electrons, a robust, **cut-based electron selection** is used. The selection criteria are optimised with respect to the selection efficiency of $\mathrm{W}^{\pm} \rightarrow \mathrm{e}^{\pm}\nu_{\mathrm{e}}$ against QCD events as described in an analysis from the electroweak working group [35]. A detailed description of the electron ID requirements is given in section 6.4. The application of a simple cut-based electron ID results in an overall efficiency for *fake* electrons from QCD background at the permill level, while obtaining an identification efficiency for electrons with momentum $5 < p_{\mathrm{T}} < 50$ GeV/c of about 90%, as described in [34].

5.2.3 Jet Reconstruction

At a hadron collider experiment such as CMS, every recorded event contains a huge number of reconstructed charged-particle tracks and energy deposits in the calorimetry. Any coloured quark or gluon produced in a proton-proton collision hadronises into colourless mesons or baryons, detected as a jet in the CMS experiment. In order to reconstruct the kinematics of a $\mathrm{t\bar{t}}$ quark pair event, it is important to understand the jet composition in such events. A key problem is the correct association of a quark or gluon produced in a hard scattering process with a jet, measured in the detector. Therefore, in CMS, three different approaches for the reconstruction of jets are used: *calorimeter (calo) jet* reconstruction using only information from the calorimeters, *Jet-Plus-Track* (JPT) reconstruction, which uses calorimeter jets and combines them with the information of associated tracks from the inner tracking system, and the *Particle-Flow* (PF) reconstruction, which includes information from all subdetectors to subsequently reconstruct each individual particle in the event. For the analysis presented in this thesis, Particle-Flow reconstructed jets are used. The PF algorithm is explained in further detail in section 5.2.6.

Jet Energy Resolution

The jet energy resolution of the HCAL has been studied using simulated QCD dijet events by comparing the reconstructed jet transverse energy $E_{\mathrm{T}}^{\mathrm{reco}}$ to the generated transverse energy $E_{\mathrm{T}}^{\mathrm{MC}}$. The result is shown in figure 5.9. The jet transverse energy resolution shows similar performance in all different detector components, the barrel, the endcaps and in the forward direction. The distribution of the resolution $\sigma\left(\frac{E_{\mathrm{T}}^{\mathrm{reco}}}{E_{\mathrm{T}}^{\mathrm{MC}}}\right)/\langle\frac{E_{\mathrm{T}}^{\mathrm{reco}}}{E_{\mathrm{T}}^{\mathrm{MC}}}\rangle$ is parametrised with the following functional form:

$$\frac{\sigma\left(\frac{E_{\mathrm{T}}^{\mathrm{reco}}}{E_{\mathrm{T}}^{\mathrm{MC}}}\right)}{\langle\frac{E_{\mathrm{T}}^{\mathrm{reco}}}{E_{\mathrm{T}}^{\mathrm{MC}}}\rangle} = \frac{A}{E_{\mathrm{T}}^{\mathrm{MC}}} + \frac{B}{\sqrt{E_{\mathrm{T}}^{\mathrm{MC}}}} + C \ ,$$

where the term A accounts for energy fluctuation, such as electronic noise or underlying event energy, B is the stochastic response of the measurement and C represents non-linearities in the detector response.

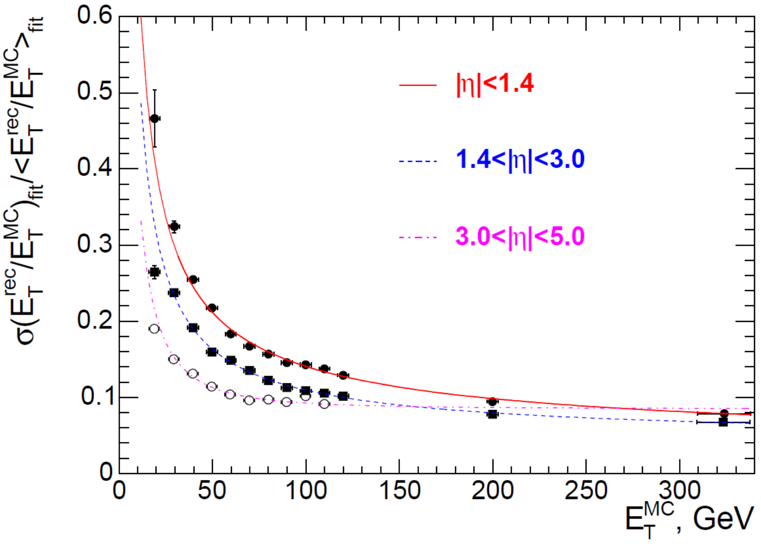


Figure 5.9: The jet energy resolution of the HCAL has been studied using simulated QCD dijet events by comparing the reconstructed jet transverse energy $E_T^{\rm reco}$ to the generated transverse energy $E_T^{\rm MC}$ [22].

Tower definition

A calorimeter jet is defined by an accumulation of energy deposited by hadrons in the calorimeters (HCAL and ECAL). The readout cells in the HCAL are arranged in towers in $(\eta - \phi)$ space in the direction towards the interaction point, as described in [22]. A *calorimeter tower* combines the information from the ECAL and HCAL by adding the signals from the much finer binned ECAL into one corresponding bin in the HCAL. Therefore, the size of a HCAL cell determines the size of a calorimeter tower, where one HCAL cell corresponds to an 5×5 array of ECAL cells. *Jet clustering algorithms*, which are described in the following subsection, use calorimeter towers as input. The total energy of a tower is given by the sum of all corresponding readout cells above a certain threshold, in order to suppress detector noise. A tower is further treated as a massless particle, with the energy given by the tower energy and the direction defined by the centre of the tower and the nominal interaction point.

Jet Algorithms

In CMS, three different jet reconstruction methods are used, which are distinguished by the set of subdetectors, which are used for jet finding. *Jet clustering* or *cone algorithms* are used as input to all jet reconstruction methods. Different jet algorithms usually result in a different number of jets in the event after full reconstruction. Cone algorithms are usually simpler and faster and therefore used in software-based triggering systems for jet reconstruction, while clustering algorithms are mostly used for offline analysis. Cone jet algorithms group input objects together intermediately and the final jet quantities are calculated once at the end of the jet finding procedure. On the other hand, clustering algorithms merge input objects into possible final jets in each iteration step and the quantities of a potential jet have to be calculated during the clustering.

Clustering Algorithms There are three clustering algorithms available at CMS for jet reconstruction, as described in [36]: the k_{T} -algorithm, the Cambridge/Aachen algorithm and the Anti-k_{T} -algorithm. Each of the clustering algorithms combines the input four vectors of two particles, i and j, pair-wise according to the distance of the two particles d_{ij} and their individual distance to the beam $d_{i\mathrm{B}}$:

$$d_{ij} = \min\left((k_{\mathrm{T}})_i^{2p}, (k_{\mathrm{T}})_j^{2p}\right) \frac{\Delta_{ij}^2}{D^2} \ ,$$

$$d_{i\mathrm{B}} = (k_{\mathrm{T}})_i^{2p} \ ,$$

where Δ is defined as $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i + \phi_j)^2$ and $(k_{\mathrm{T}})_i$, y_i and ϕ_i are the particles transverse momentum, rapidity and azimuthal angle. For all input particles the **smallest distance** among d_{ij} and $d_{i\mathrm{B}}$ is determined. If d_{ij} is minimal, the two particles i and j are recombined, adding their four momenta. If $d_{i\mathrm{B}}$ is minimal, particle i is removed from the list of particles and called a *jet*. After each step, all distances are recalculated and the procedure is repeated until no particles are left [36].
All three algorithms are **infrared- and collinear-safe**, such that no soft emission or collinear splitting changes the result of the clustering. This is an important feature of jet algorithms for the interpretation of the relation between partons and jets in terms of theoretical modelling of experimental data. The *parameter p* distinguishes the three algorithms: (p = 1) corresponds to the k_{T} -algorithm, while (p = 0) corresponds to Cambridge/Aachen and (p = -1) to the Anti-k_{T} -algorithm.

Cone Algorithms Two cone algorithms are provided by the CMS group [36]: the *Iterative Cone* and the *Seedless Infrared-Safe* (SIS) Cone algorithm. In the Iterative Cone algorithm, the most energetic particle is used as a seed and all particles within a fixed cone of radius R are clustered into a jet candidate. The sum of all clustered particles gives rise to a new jet axis, which has to be compatible within precision with the seed axis. If the two axes agree, the jet candidate is called a *jet* and the jet constituents are removed from the list of input object. The procedure is repeated with the remaining, most energetic particle as the next seed. If the two axes do not agree, the procedure is repeated with the clustered jet candidate as a new seed, until a stable cone is found. Like most of the commonly used cone-type algorithms, the Iterative Cone algorithm is simple and fast to process, but it is *not found* to be infrared- or collinear-safe. Therefore, a *seedless infrared-safe* (SIS) cone algorithm is proposed, which introduces a non-iterative approach.

Two characteristic **distance parameters** were introduced for cone-type and clustering algorithms: the *jet size parameter* R and the *rescaling parameter* D for sequential clustering. Within the CMS framework, several values for these distance parameters are available: R = 0.5 for Iterative Cone and R = 0.5, 0.7 for SIS Cone and D = 0.4, 0.6 for the k_{T} -algorithms. The **Anti-k_{T} algorithm**, which uses the **inverse square** of the particles four momenta as a distance scale, is expected to have similar performance as the Iterative Cone algorithm, while also being infrared- and collinear-safe. Therefore, this algorithm is preferred and is used together with the Particle-Flow jet reconstruction algorithm for the analysis.

Calorimeter Jets

The calorimeter jet reconstruction exclusively uses *calorimeter towers* as input. As described above, a single calorimeter tower consists of one HCAL cell and an 5×5 array of ECAL cells. Beyond the coverage of the ECAL, a calorimeter tower consists of one HCAL cell [36]. Several subdetector specific energy thresholds are defined in order to reject electronic noise. Any tower, which passes the minimal $\mathrm{E_T}$ -requirements is considered as input to one of the jet finding algorithms (clustering or cone-type), described above.

Jet-Plus-Track Jets

In the Jet-Plus-Track (JPT) algorithm, jets are first reconstructed as *calorimeter jets* with a fixed-cone (*Iterative Cone*) jet finding algorithm. In order to correct for the resolution and the p_T -response of calorimeter jets, information from the inner tracking system is used. First, charged-particle tracks are associated to the reconstructed calorimeter jets, based on the $(\eta - \phi)$ distance between the jet axis and the track direction at the primary vertex. If the projection of the track onto the surface of the calorimeter points into the cone of the jet, the track is classified as *in-cone* track. If the the track is bent out of the jet cone by the magnetic field, the track is classified as *out-cone* track [36]. The track momenta of both, in-cone and out-cone tracks, are added to the energy of the calorimeter jet, while the expected energy deposit of the in-cone tracks is furthermore subtracked. The resulting energy of the JPT jet is greater than or equals the initial calorimeter jet energy, corrected by the charged-particle track momenta, which presumably are bend out of the fixed jet cone and therefore can only be used with the Iterative Cone algorithm.

Particle-Flow Jets

The Particle-Flow (PF) jet reconstruction algorithm is currently the only method, which makes use of all subdetector systems of the CMS experiment. The aim of the algorithm is to correctly identify each object in the event on the *particle level*. As already introduced (see figure 5.2), by combining subsequently all subdetector information, five categories of particles are distinguished, based on their different detector response:
Firstly, muons are identified, due to their unique signature in the muon system, despite *punch-through* effects of hadrons, which give rise to misidentified muon candidates. The reconstruction of muons combines information from all CMS subdetectors (inner tracking system, calorimetry and muon system). Secondly, electrons and charged hadrons are reconstructed, using information from the inner tracking system and the full calorimetry (ECAL and HCAL). Finally, photons and neutral hadrons are reconstructed, using either the ECAL or (in case of neutral hadrons) the HCAL alone. The reconstruction of individual particles by the PF algorithm is described in further detail in section 5.2.6.

Jet Energy Corrections

In order to correct reconstructed jets for their mismeasured energies, three different corrections are applied successively to each jet. In CMS, a factorised approach is used, which corrects jets to the parton-level (MC) information: first, energy contributions due to electronic noise and pileup are subtracted (*offset correction*). Secondly, a *relative correction* is applied, to correct for non-uniformities in the jet response in η, compared to the average response in the barrel calorimeters. Finally, the energy of reconstructed jets is corrected to the simulated response of associated generated jets, referred to as *absolute correction.*

5.2.4 B Jet Identification

The identification of jets containing decay products of b hadrons, referred to as *b tagging*, is an important aspect of the full event reconstruction procedure, in order to separate these jets from jets associated with gluons or light quarks [37]. Therefore, different algorithms have been developed, using information from the inner tracking system due to the long lifetime of the b hadron, as well as the presence of non-isolated leptons, stemming from the semileptonic decay of b hadrons. In the following, two algorithms are considered, both exploiting the relatively long lifetime of the b hadron, due to its weak decay, which are using the **track impact parameter** and **reconstructed secondary vertices** as input for the construction of a discriminating variable, referred to as b tagging "*discriminator*". This numerical output is required to be above a certain threshold, called *operating point*, in order that a jet is considered as b jet.

Track Counting Algorithm

The track counting (TC) algorithm is an approach to b jet identification, which uses an effective single-track observable: the distance between the track and the associated vertex at the point of closest approach, called the **impact parameter** (IP), as described in [37]. The variable, which is used as the discriminator for the TC algorithm, is the **significance** of the impact parameter $\frac{\mathrm{IP}}{\sigma_{\mathrm{IP}}}$, where σ_{IP} is the uncertainty of the IP. The tracks contained in a jet are ordered in decreasing significance. If a jet contains at least N tracks with a significance of the impact parameter above the discriminator value S, the jet is identified as a b jet by the TC algorithm. Hence, the discriminator variable corresponds to the significance of the IP of the N^{th} track associated to the jet. According to $N = 2$ and $N = 3$, two variants of the TC algorithm are distinguished, referred to as *track counting high efficiency* (TCHE) and *track counting high purity* (TCHP).

Simple Secondary Vertex Algorithm

In the case of the simple secondary vertex (SSV) algorithm a direct search for a secondary vertex from the decay of the b hadron is performed. For a given set of tracks, the same reconstruction procedure as for primary vertex reconstruction is used. If at least 65% of all tracks are shared with the primary vertex, the vertex is removed from the list of candidates [37]. The most important variables used for the vertex reconstruction are the track multiplicity and the **three-dimensional flight distance**. The flight distance is used as the discriminator for the SSV b tagging algorithm, whereas a minimal number of at least two tracks associated to the reconstructed secondary vertex is required. Similar to the TC algorithm, a number of tracks $N \geq 2$ corresponds to the SSV *high efficiency* (SSVHE) version of the tagging algorithm. Also a version with $N \geq 3$ is available, which is referred to as the SSV *high purity* (SSVHP) algorithm.

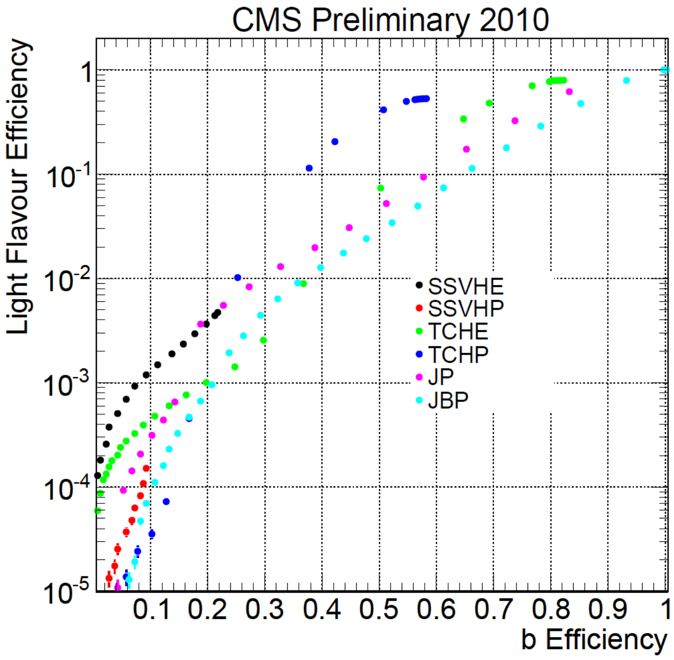


Figure 5.10: Acceptance of light jets (*mistag rate*) versus b-tagging efficiency [37].

The performance of the algorithms described above is characterised by the selection efficiency of real b jets (*b tagging efficiency*) and the acceptance of light-flavour jets, called the *mistag* rate, which measures the purity of the selected b jet sample. In figure 5.10, the mistag rate is shown versus the b tagging efficiency. For each of the algorithms, distinct operating point are defined, which correspond to a specific value of the discriminating variable. These operating points are labelled "*loose*", "*medium*" and "*tight*". Based on an estimation from simulated data samples, the operating points correspond to an acceptance of light-flavour partons of 10%, 1% and 0.1% respectively, given in [37].

5.2.5 Missing Transverse Energy

The primary goal of the event reconstruction is to obtain a complete picture of the event and the kinematics, including neutrinos or other weakly-interacting stable particles, which give rise to apparent missing energy in the event [22]. It is impossible to detect particles produced in the longitudinal direction of the beam pipe, hence the balance of energy deposits projected into the transverse plane to the beam direction (spanned by $\vec{i}$ and $\vec{j}$) is measured with good accuracy. The **missing transverse energy vector** $\vec{\mathrm{E}}_{\mathrm{T}}^{\mathrm{miss}}$ is measured as the sum of the projection of all individual calorimeter towers i, as described in [22]:

$$\text{with} \quad E_{\mathrm{x}}^{\mathrm{miss}} = \sum_{\text{towers } i} \left(E_i \sin(\theta_i) \cos(\phi_i)\right)$$

$$\text{and} \quad E_{\mathrm{y}}^{\mathrm{miss}} = \sum_{\text{towers } i} \left(E_i \sin(\theta_i) \sin(\phi_i)\right)$$

$$\Longrightarrow \quad \vec{\mathrm{E}}_{\mathrm{T}}^{\mathrm{miss}} = E_{\mathrm{x}}^{\mathrm{miss}} \cdot \vec{i} + E_{\mathrm{y}}^{\mathrm{miss}} \cdot \vec{j} \tag{5.2}$$

Also reconstructed muons contribute to the calculation of $\mathrm{E}_\mathrm{T}^\mathrm{miss}$ by taking into account their reconstructed track p_T instead of the expected calorimeter deposit. It is expected, that the resolution of the missing energy depends on the overall activity in the calorimetry, measured by the scalar sum of transverse energy in all calorimeter cells $(\sum E_\mathrm{T})$ [22]. The resolution is found to follow the form $\sigma(\Sigma E_\mathrm{T}) = C \cdot \sqrt{\Sigma E_\mathrm{T}}$, where C is a constant, which depends on the jet resolution, shown above. Figure 5.11 shows the resolution $\sigma(\Sigma E_\mathrm{T})$ as a function of the total sum of all calibrated energy deposits, found by the Particle-Flow algorithm, which is described in the next section.

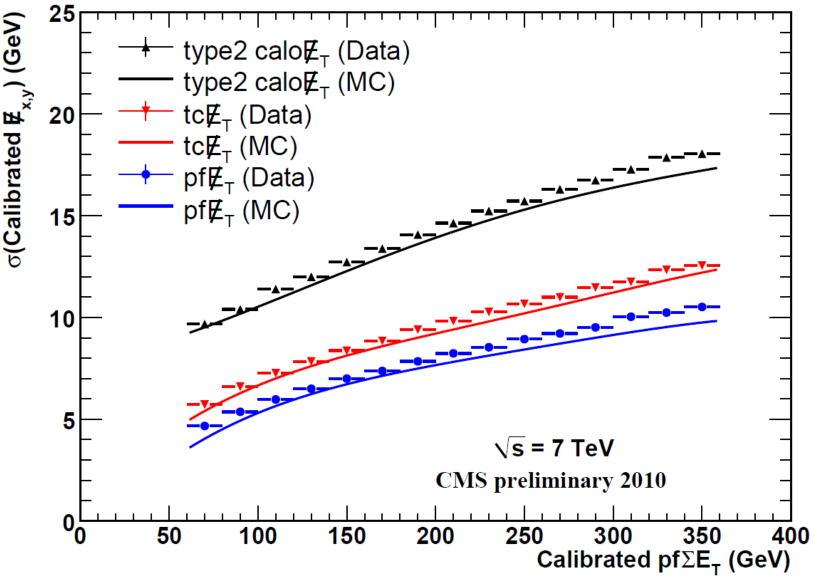


Figure 5.11: Missing transverse energy resolution $\sigma(\Sigma E_\mathrm{T})$ as a function of the total sum of all calibrated energy deposits, found by the Particle-Flow algorithm [38].

5.2.6 Particle-Flow Reconstruction

The Particle-Flow (PF) algorithm attempts to identify and reconstruct each individual particle in the event, utilising information from *all CMS subdetectors*. Figure 5.2 gives an overview of the different contributions of each subsystem to the event reconstruction procedure, which has been described above. One important aspect of the PF reconstruction algorithm is to provide a better performance of reconstructed jets and missing transverse energy, containing an improved resolution and an intrinsic calibration. A complete description of the PF algorithm is given in [39].

The PF algorithm starts with an improved reconstruction of charged-particle tracks (*iterative tracking*) and calorimeter clusters (*topological clustering*), as well as the reconstruction of muon tracks. These objects are called the "*fundamental elements*", which serve as input to the PF reconstruction algorithm. Before the PF algorithm is applied, the fundamental elements (tracks & clusters) are linked to each other, in order to form "*blocks*", which constitute an early stage of reconstructed particles. Special emphasis is put on the identification of *charged and neutral hadrons*, as well as *photons*. The standard object reconstruction algorithms only utilise information from the calorimetry, whereas the PF reconstruction algorithm incorporates the superior momentum and direction measurement of the inner tracking system. Therefore, an improved performance of the separation and identification of reconstructed jets and photons is expected. As an example, a typical jet energy fraction of 65% is carried by charged particles, 25% by photons and 10% by neutral hadrons, as described in [39]. Finally, the amount of **missing transverse energy** in the event is calculated from the modulus of the vectorial sum of the transverse momenta of all reconstructed particles, which is also expected to have an improved performance compared to MET solely reconstructed from the calorimetry, as described above.

Fundamental Elements

As a first approach to the complete reconstruction of all individual particles in the event, fundamental elements are build, which comprise charged-particle tracks, calorimeter clusters and tracks in the muon system. These objects are used as input to the *link algorithm*, which is described later. Two improved methods for the reconstruction of tracks and clusters are described in the following.

Iterative Tracking The momentum of charged hadrons and the direction of charged particles at the vertex is measured to much higher accuracy in the silicon tracker, compared to the calorimeters over a large range of energies. In order to optimise the reconstruction efficiency of **charged-particle tracks**, while simultaneously minimising the rate of misidentified charged tracks (fake rate), an iterative track reconstruction procedure is applied, in order to build the track input collection to the PF reconstruction algorithm. First, the track finding procedure is seeded with relatively tight reconstruction criteria, which minimises the fake rate, while obtaining a moderate efficiency for real tracks. As a next step, all unambiguous hits found in the previous iteration are removed from the list of hits and the track finding procedure is continued with looser track seeding criteria. This leads to an increase of the tracking efficiency, while the previous removal of hits keeps the fake rate small, due to the reduced combinatorics.

Topological Clustering The reconstruction of calorimeter clusters is of special interest, in order to measure the energy and direction of all **neutral particles**, such as photons and neutral hadrons. The main task of the improved clustering algorithm is to separate neutral particles from charged hadrons, as well as the reconstruction of electrons. As a first step, the clustering procedure is seeded by finding a local maximum in the calorimeter cells. In the following, the *topological clustering* algorithm adds cells with at least one side in common with a cell in the cluster and with an energy above a given threshold to the **PF cluster**. The advantage of the topological clustering algorithm is, that the energy of each calorimeter cell is shared among all PF clusters according to the cell-cluster distance, which provides an iterative determination of cluster energies and positions.

Link Algorithm

One given particle gives rise to the construction of several PF elements, as described above. Before the fundamental elements are handed over to the core of the PF algorithm, the PF reconstruction algorithm, they are linked to each other, in order to form *blocks*. In the following, these blocks are the input to the reconstruction and identification algorithm. First, charged-particle tracks are linked to calorimeter clusters, by extrapolating tracks from the inner tracking system into the calorimeter, similar to the electron reconstruction algorithm, described above. As a next step, in order to collect calorimeter clusters caused by the radiation of **bremsstrahlung photons**, tangents to the track at the intersection point of each layer of the silicon tracker are extrapolated into the ECAL. If the extrapolated tangent track position is within the boundaries of a cluster, the cluster is linked to the track. This is also part of the track-to-cluster linking procedure. Similarly, also links are established between two calorimeter clusters from the ECAL and HCAL or preshower, if the position of a cluster lies within the envelope of a less granular calorimeter cell. Finally, charged-particle tracks are linked to muon tracks, in order to create **global muon** candidates.

Particle-Flow Reconstruction and Identification

The quintessential part of the Particle-Flow algorithm is the reconstruction and identification algorithm, which uses the fundamental particles as input, as described above. The reconstruction sequence starts with the construction of **PF muons** from each global muon candidate and the muon track is removed from the block. As a next step, each successfully identified electron creates a **PF electron** and the corresponding track and ECAL cluster are removed from the block. In the following, a tightening of the track quality criteria is applied to the list of input tracks, as well as a calibration of the calorimeter cluster energies [39]. The remaining elements give rise to charged and neutral hadrons, as well as photons. Each remaining charged track in the block gives rise to a **PF charged hadron**. Finally, if the calibrated energy of a remaining PF cluster is significantly larger than any charged-particle momentum, a **PF photon** or a **PF neutral hadron** is created.

Chapter 6

Event Selection

The event signature of a dileptonically decaying $t\bar{t}$ quark pair consists of two isolated, oppositely-charged leptons with high transverse momentum, two high-energetic b jets and large missing transverse energy $\not\!\!E_{\mathrm{T}}$, due to the neutrinos from the W boson decays. In this analysis, only dilepton events with at least one muon and one electron are selected. The event selection is applied in several steps, starting with the selection of events with at least one muon with $p_{\mathrm{T}} > 20$ GeV/c, which has been reconstructed by *both* the global and tracker muon algorithms (*step0*). The next selection step requires one well identified and isolated muon, which mainly suppresses background from QCD multijet events (*step1*). As a next selection step, one identified and isolated electron is required, which has to give rise to one muon-electron pair with unlike charge (*step2*). Finally, at least two jets with high transverse momentum are required (*step4*). For the muon-electron channel, no missing transverse energy is required on top of the dilepton and jet selection. In addition to the dilepton-plus-jets selection, at least one jet is required, which is identified as b jet by a b-tagging algorithm, in order to study the efficiency of the b jet identification algorithm. The same selection is applied to data and simulated samples and corresponds to the reference selection described in the "*Pass6*" top-dilepton reference analysis, as described in [40]. The number of observed events in the data is compared to the expected number of events from simulation after the following, sequential selection steps:

1. **One high-p_{T} muon (step0)**: at least one muon is reconstructed as global and tracker muon with $p_{\mathrm{T}} > 20$ GeV/c
2. **One isolated muon (step1)**: at least one muon is passing the full muon selection requirements (trigger, kinematic, identification and isolation requirements)
3. **Two isolated leptons (step2)**: one pair of an identified and isolated, oppositely-charged muon and electron with an invariant dilepton mass greater than 12 GeV/c^2, referred to as *full $\mu^{\pm}e^{\mp}$ selection*
4. **Two jets (step4)**: full $\mu^{\pm}\mathrm{e}^{\mp}$ selection with at least two additional AK5 PF jets fulfilling all jet requirements, referred to as *dilepton-plus-jets selection*
5. **One b-tag**: dilepton-plus-jets selection with at least one additional b-tagged jet (TCHE *loose* working point)

6.1 Data Samples

Collision data

The analysis is performed on the full data set of proton-proton collisions, recorded in 2010. In order to select dileptonically decaying signal events, data samples recorded with a single-muon or single-electron trigger are used. The collision data sets used in this analysis are listed in table 6.1. These samples are reprocessed with CMS software version `CMSSW_3_8_6`. The list of *certified, good runs* given in [41] is used. The selected runs, which are used as input to the event selection, correspond to an integrated luminosity of 35.9 pb^{-1} with a relative uncertainty of 3.6%.

Sample	Run range
`/Mu/Run2010A-Nov4ReReco_v1/AOD`	135 821 - 144 114
`/EG/Run2010A-Nov4ReReco_v1/AOD`	135 821 - 144 114
`/Mu/Run2010B-Nov4ReReco_v1/AOD`	146 240 - 149 711
`/Electron/Run2010B-Nov4ReReco_v1/AOD`	146 240 - 149 711

Table 6.1: Collision data samples selected by single-muon or single-electron triggers of the 2010 data taking period, which are used in the analysis.

Simulated data

The simulated data samples used in this analysis correspond to the *Fall10* production series of Monte Carlo samples [42], as described in the *Pass6* reference analysis of the top-dilepton working group [40]. The selected samples comprise the inclusive $t\bar{t}$ quark pair production, as well as all relevant background processes. The background processes used in this analysis are:

• Drell-Yan $Z^0/\gamma^\star$ processes	(DY)
• Associated W boson and single-top quark production	(tW)
• W boson production plus additional jets	(W+Jets)
• Diboson production WW, WZ and ZZ	(VV)
• QCD multijet events	(QCD)

The inclusive $t\bar{t}$ quark pair sample is split on generator level into the muon-electron signal channel, including leptonic tau decays ($\tau\tau$, τe & $\tau\mu$), and all other $t\bar{t}$ quark pair decays, which are considered to be background processes to the signal selection. For the simulation of the QCD background, a muon-enriched data sample with a muon transverse momentum of greater than 15 GeV/c is used. The inclusive $t\bar{t}$ quark pair, the single-top quark tW and the DY background samples are generated with MadGraph [30] and interfaced with Pythia [6] for hadronisation and parton fragmentation. For low dilepton invariant masses (generated $Z^0/\gamma^\star$ mass smaller than 50 GeV/c^2), additional DY samples generated with Pythia are used. The diboson and QCD samples are also produced with Pythia. The production cross sections used for the generated samples are given at [40], based on [43] and [18]. A summary of all simulated data samples is given in table 6.2:

Process	Sample	σ [pb]
Inclusive $\mathrm{t\bar{t}}$	`TTJets_TuneD6T_7TeV-madgraph-tauola`	157.5
single-top tW	`TToBLNu_TuneZ2_tW-channel_7TeV-madgraph`	10.6
Diboson WW	`WWTo2L2Nu_TuneZ2_7TeV-pythia6`	4.5
Diboson WZ	`WZTo3LNu_TuneZ2_7TeV-pythia6`	0.6
Diboson ZZ	`ZZtoAnything_TuneZ2_7TeV-pythia6-tauola`	7.4
W+Jets (W→ eν)	`WToENu_TuneZ2_7TeV-pythia6`	10 438
W+Jets (W→ $\mu\nu$)	`WToMuNu_TuneZ2_7TeV-pythia6`	10 438
W+Jets (W→ $\tau\nu$)	`WToTauNu_TuneZ2_7TeV-pythia6-tauola`	10 438
QCD (μ enriched)	`QCD_Pt-20_MuEnrichedPt-15_TuneZ2_7TeV-pythia6`	84 679
	10 GeV/c^2 < $M_{Z^0/\gamma^\star}$ < 20 GeV/c^2	
$Z^0/\gamma^\star \rightarrow e^+e^-$	`DYToEE_M-10To20_TuneZ2_7TeV-pythia6`	3 457
$Z^0/\gamma^\star \rightarrow \mu^+\mu^-$	`DYToMuMu_M-10To20_TuneZ2_7TeV-pythia6`	3 457
$Z^0/\gamma^\star \rightarrow \tau^+\tau^-$	`DYToTauTau_M-10To20_TuneZ2_7TeV-pythia6-tauola`	3 457
	20 GeV/c^2 < $M_{Z^0/\gamma^\star}$ < 50 GeV/c^2	
$Z^0/\gamma^\star \rightarrow e^+e^-$	`DYToEE_M-20_TuneZ2_7TeV-pythia6`	1 666
$Z^0/\gamma^\star \rightarrow \mu^+\mu^-$	`DYToMuMu_M-20_TuneZ2_7TeV-pythia6`	1 666
$Z^0/\gamma^\star \rightarrow \tau^+\tau^-$	`DYToTauTau_M-20_TuneZ2_7TeV-pythia6-tauola`	1 666
	50 GeV/c^2 < $M_{Z^0/\gamma^\star}$	
$Z^0/\gamma^\star \rightarrow e^+e^-$, $\mu^+\mu^-$	`DYJetsToLL_TuneZ2_M-50_7TeV-madgraph-tauola`	3 048
$Z^0/\gamma^\star \rightarrow \tau^+\tau^-$	`DYJetsToLL_TuneD6T_M-50_7TeV-madgraph-tauola`	3 048

Table 6.2: Summary of simulated data samples and corresponding cross sections used in this analysis.

The MC samples are simulated with different "*tunes*", which have different amounts of initial and final state radiation, in order to describe correctly the particle multiplicities in proton-proton interactions. Some samples are simulated using the *D6T* tune, because the corresponding *Z2* tune samples are affected by a problem: a large fraction of taus were decayed by GEANT, which makes it impossible to find their decay products on generator level. To compare simulated events to events from collision data, the simulated data samples are normalised to the integrated luminosity of 35.9 pb^{-1}. The conditions for calibration and alignment of the detector defined for early running are taken into account. This analysis is performed with `CMSSW_3_8_6` and the list of tags is given in [40], using the *Physics Analysis Toolkit* (*PAT*) layers 0 and 1 as the initial step [44].

6.2 Event Cleaning and Trigger Selection

Data events are required to fulfil a *good run* selection, mentioned in the previous section 6.1. Events originating from proton-proton collisions are selected. As a first step, the events are required to have a significant fraction of high-purity tracks (> 25%) with respect to the total number of tracks in events with at least 10 tracks. Also events with significant noise in the hadronic calorimeters are removed. In addition, the presence of a well identified primary vertex with ndof > 4, $|\rho| < 2$ cm and $|z| < 24$ cm is required. Due to the increase of the instantaneous luminosity over several orders of magnitude in the 2010 running period, the trigger configuration changed during the data taking period. As a result, a *run-dependent trigger selection* with increasing p_T -thresholds is applied to the collision data samples. In this analysis, events selected by a **single-muon** or **single-electron trigger** are considered. The list of used triggers, together with the run ranges and corresponding integrated luminosities is shown in tables 6.3 and 6.4. Dilepton events in the muon-electron final state are required to be selected *either* by a muon *or* an electron trigger. In order to avoid multiple event selection, events from the electron data sample are not selected, if the event has also been triggered by the muon trigger of the corresponding data period. Simulated events are selected using the `HLT_Mu9` or `HLT_Ele10_SW_L1R` trigger. A correction is applied to the trigger efficiencies, in order to account for the difference between the triggers used in the data samples and in the simulated samples. Further details are given in section 7.1.

Run range	Trigger path	L_{int} [pb^{-1}]
136 035 - 144 999	`HLT_Mu9`	3.2
145 000 - 147 119	`HLT_Mu11`	5.0
147 120 - 149 294	`HLT_Mu15_v1`	27.7

Table 6.3: Single-muon triggers used in the analysis.

Run range	Trigger path
< 138 000	`HLT_Ele10_LW_L1R`
138 000 - 141 899	`HLT_Ele15_LW_L1R`
141 900 - 143 999	`HLT_Ele15_SW_L1R`
144 000 - 144 114	`HLT_Ele15_SW_CaloEleId_L1R` OR `HLT_Ele20_SW_L1R` OR `HLT_DoubleEle10_SW_L1R`
145 000 - 147 119	`HLT_DoubleEle10_SW_L1R` OR `HLT_Ele17_SW_CaloEleId_L1R`
147 120 - 148 100	`HLT_DoubleEle15_SW_L1R_v1` OR `HLT_Ele17_SW_TightCaloId_SC8HE_L1R_v1` OR `HLT_Ele17_SW_TightEleId_L1R`
> 148 100	`HLT_DoubleEle17_SW_L1R_v1` OR `HLT_Ele17_SW_TightCaloEleId_Ele8HE_L1R_v1` OR `HLT_Ele17_SW_TightCaloEleId_Ele8HE_L1R_v2` OR `HLT_Ele22_SW_TighterEleId_L1R_v2` OR `HLT_Ele22_SW_TighterEleId_L1R_v3` OR `HLT_Ele17_SW_TighterEleIdIsol_L1R_v2` OR `HLT_Ele17_SW_TighterEleIdIsol_L1R_v3`

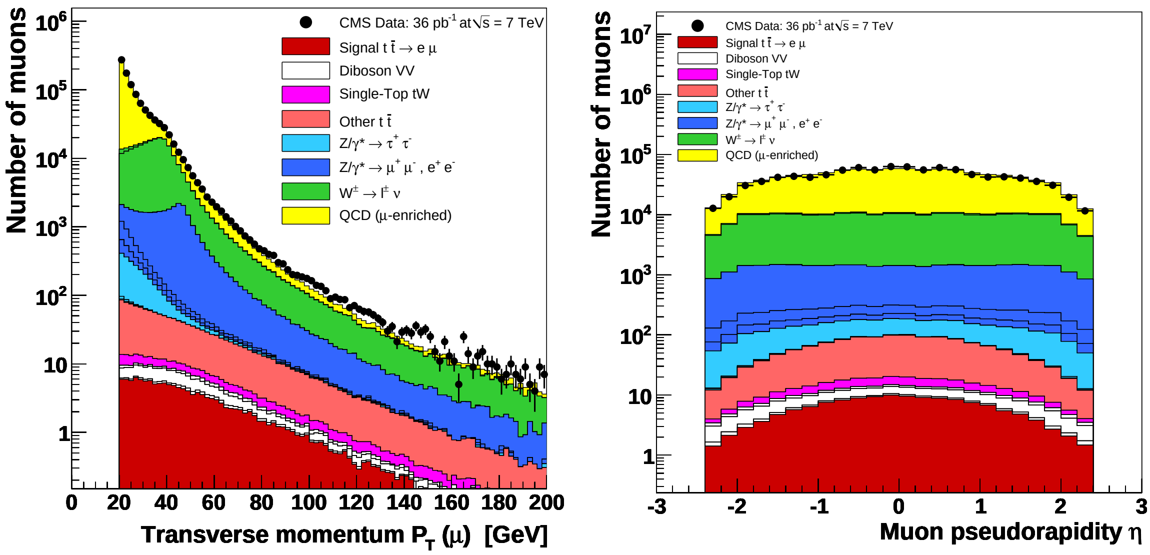
Table 6.4: Single-electron triggers used in the analysis.

6.3 Muon Selection

Muon candidates selected for this analysis are required to have $p_{\mathrm{T}} > 20$ GeV/c and to be within the pseudorapidity region $|\eta| < 2.4$. Several muon identification requirements are applied to the candidates. The muons have to be reconstructed by both **muon reconstruction** algorithms, the *tracker muon* and *global muon* algorithm. In order to remove misidentified muons or muons from decay-in-flight processes, a minimal number of hits in the silicon tracker associated to the muon candidate of at least 11 hits in the silicon tracker ($N_{\mathrm{hits}}^{\mathrm{trk}} > 10$) is required, as well as a transverse impact parameter of the muon track relative to the beam spot of less than 200 μm ($|d_0^{\mathrm{BS}}| < 0.02$ cm). The combination of silicon track hits and hits in the muon chambers (global track fit) is required to have a good quality of $\chi^2/\mathrm{ndof} < 10$ and at least one hit in the muon detectors. The **muon isolation** is defined as the combined relative isolation, given in equation 6.1:

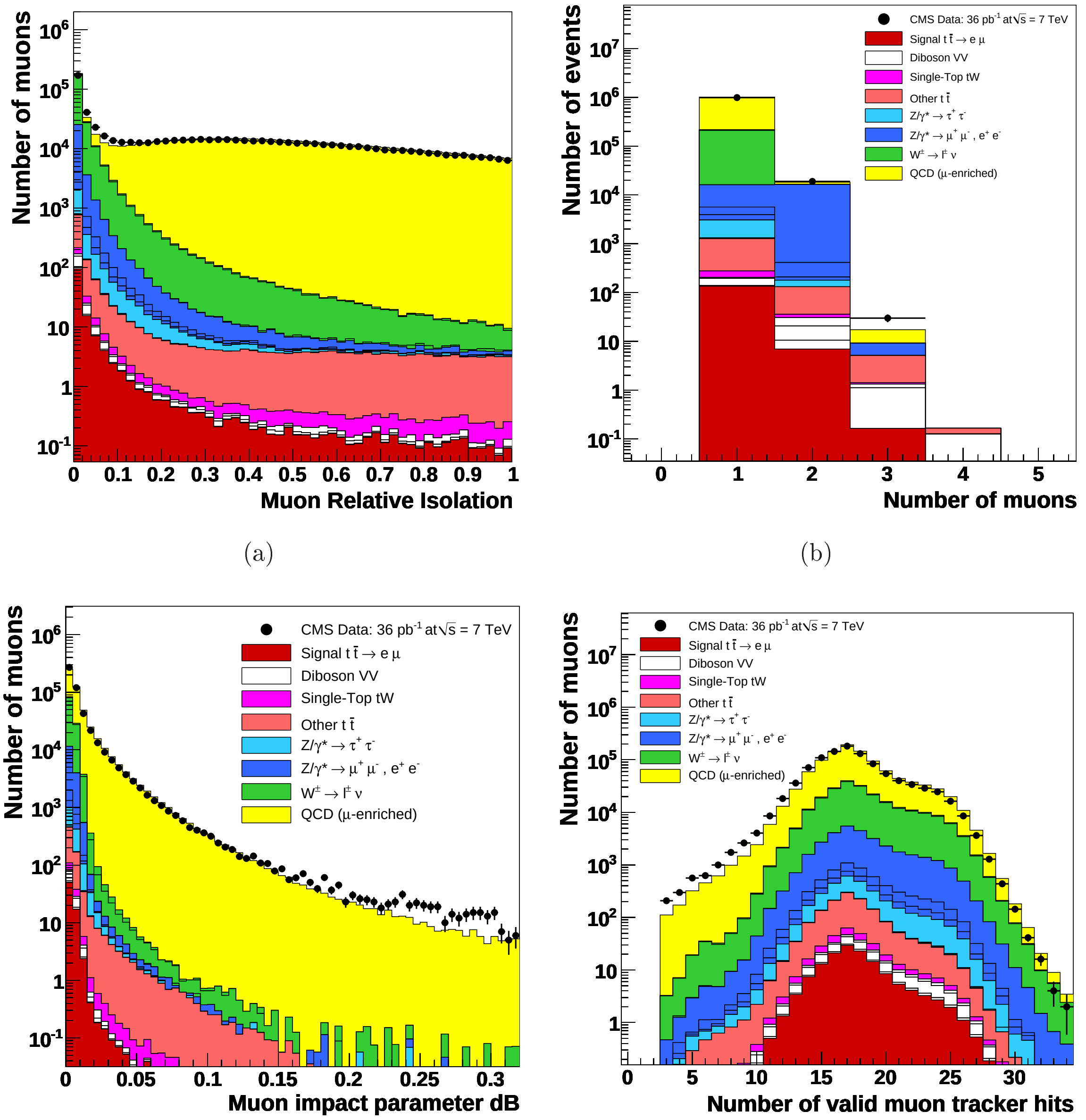
$$I_{\mathrm{comb}}^{\mathrm{rel}} = \frac{I_{\mathrm{ECAL}} + I_{\mathrm{HCAL}} + I_{\mathrm{Trk}}}{p_{\mathrm{T}}(\mu)} \tag{6.1}$$

The values for calorimeter isolation I_{ECAL} and I_{HCAL} are defined as the amount of transverse energy deposits inside a cone in $(\eta - \phi)$ space of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.3$ around the muon candidate. Similarly, the value for the tracker isolation I_{Trk} is defined as the sum of transverse momentum of the tracks in a cone of $\Delta R < 0.3$ around the muon. The track and energy deposits associated with the muon candidate are excluded. Isolated muons are required to have $I_{\mathrm{comb}} < 0.15$.



(a) (b)

Figure 6.1: (a) Muon transverse momentum p_{T} and (b) pseudorapidity η after applying the kinematic requirements $p_{\mathrm{T}} > 20$ GeV/c and $|\eta| < 2.4$ for reconstructed tracker and global muons (**step0**). Simulated samples are normalised to an integrated luminosity of 35.9 pb^{-1}.



(a) (b)

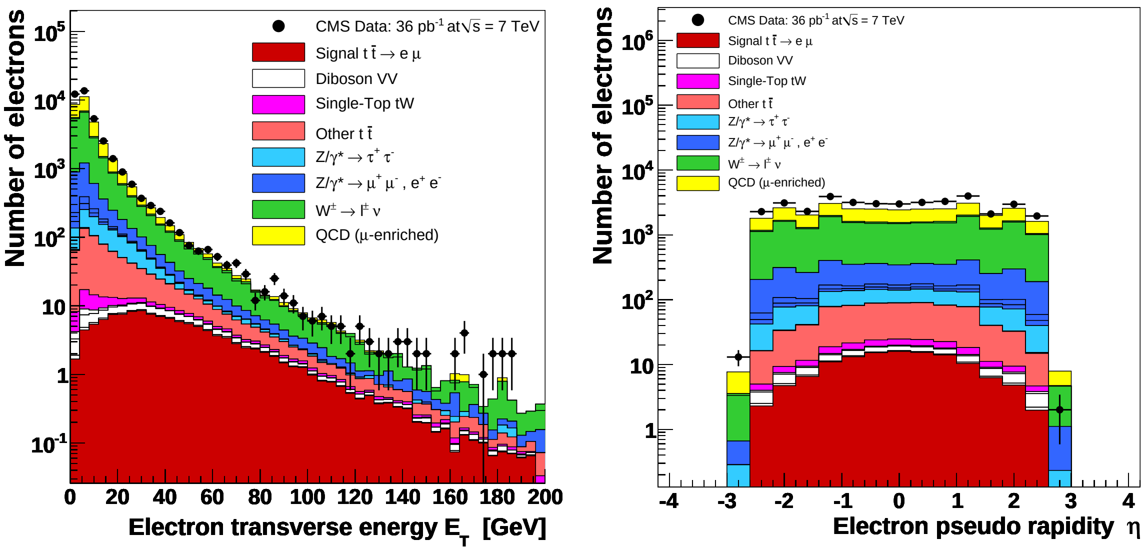
(c) (d)

Figure 6.2: (a) Muon relative combined isolation $I^{\mathrm{rel}}_{\mathrm{comb}}$ for reconstructed global and tracker muons, which pass the kinematic requirements $p_{\mathrm{T}} > 20$ GeV/c and $|\eta| < 2.4$ (**step0**). (b) Multiplicity of muons. (c) Transverse impact parameter of the muon track relative to the beam spot. (d) Number of muon track hits in the silicon tracker. Simulated samples are normalised to an integrated luminosity of 35.9 pb^{-1}.

In figure 6.1, the transverse momentum of the muon and its pseudorapidity are compared to the simulation after applying the kinematic requirements, as well as the selection of tracker and global muons (*step0*), because the QCD sample is already enriched with one muon of $p_{\mathrm{T}} > 15$ GeV/c. At this first stage of event selection, QCD events are the dominant source of non-isolated, high-p_{T} muons. The distributions for the relative combined isolation, as well as the multiplicity of muons, the number of tracker hits and the muon impact parameter d_0^{BS} before the identification and isolation requirements are shown in figure 6.2. All muon properties are compared after applying the kinematic requirements before the selection of isolated muons.

6.4 Electron Selection

Electron candidates selected for this analysis are required to have $p_{\mathrm{T}} > 20$ GeV/c and to be within the pseudorapidity region $|\eta| < 2.5$. The distribution of electron candidate kinematics before applying any identification or isolation requirement is shown in figure 6.3, where the presence of at least one well identified and isolated muon is required (*step1*), in order to reject non-genuine electrons from QCD background events. The main source of events with isolated muons are W+Jets events, in which the W bosons decays leptonically.



(a) (b)

Figure 6.3: (a) Electron transverse momentum p_{T} and (b) pseudorapidity η before applying any identification criteria after the full muon identification and isolation selection (**step1**). Simulated samples are normalised to an integrated luminosity of 35.9 pb^{-1}.

Similar to the muon selection, electron identification and isolation requirements are applied. Additionally, a **conversion veto requirement** is applied to the electron candidates, which is described later in this section. In order to reject mismeasured electrons, the E_{T} of the calorimeter cluster has to be greater than 15 GeV. In order to exclude muons, which are misreconstructed as electrons, electron candidates are rejected, if a tracker or global muon candidate is found within a cone of $\Delta R(\mathrm{e}, \mu) < 0.1$ around the electron candidate. The impact parameter of the GSF track associated to the electron relative to the beam spot is required to be less than 400 μm ($|d_0^{\mathrm{BS}}| < 0.04$ cm). Some of the electron quality requirements are defined separately for the ECAL barrel and endcap. The barrel and endcap regions are defined based on the supercluster pseudorapidity η_{SC} associated to the electron candidate: the region of $|\eta_{\mathrm{SC}}| < 1.479$ corresponds to the ECAL barrel, while $|\eta_{\mathrm{SC}}| > 1.479$ defines the ECAL endcap regions. The **electron identification** is based on a set of simple cuts applied to the shower shape ($\sigma_{\eta\eta}$) and track-to-cluster matching variables ($\Delta\eta_{\mathrm{in}}$, $\Delta\phi_{\mathrm{in}}$). In table 6.5, the different requirements on the electron identification variables are summarised. They are optimised in order to maximally separate signal events from W decays and QCD background events. The definition of several *operating points*, using simulated $\mathrm{W} \rightarrow \mathrm{e}\,\nu_{\mathrm{e}}$ and QCD events, is given in [35]. The working point, which is chosen for this analysis, corresponds to an efficiency of 90% on the simulated sample, referred to as "*WP90*". Neither the isolation selection nor the conversion rejection of WP90 are applied. The distributions of the electron identification variables before applying the selection requirements are shown in figure 6.4.

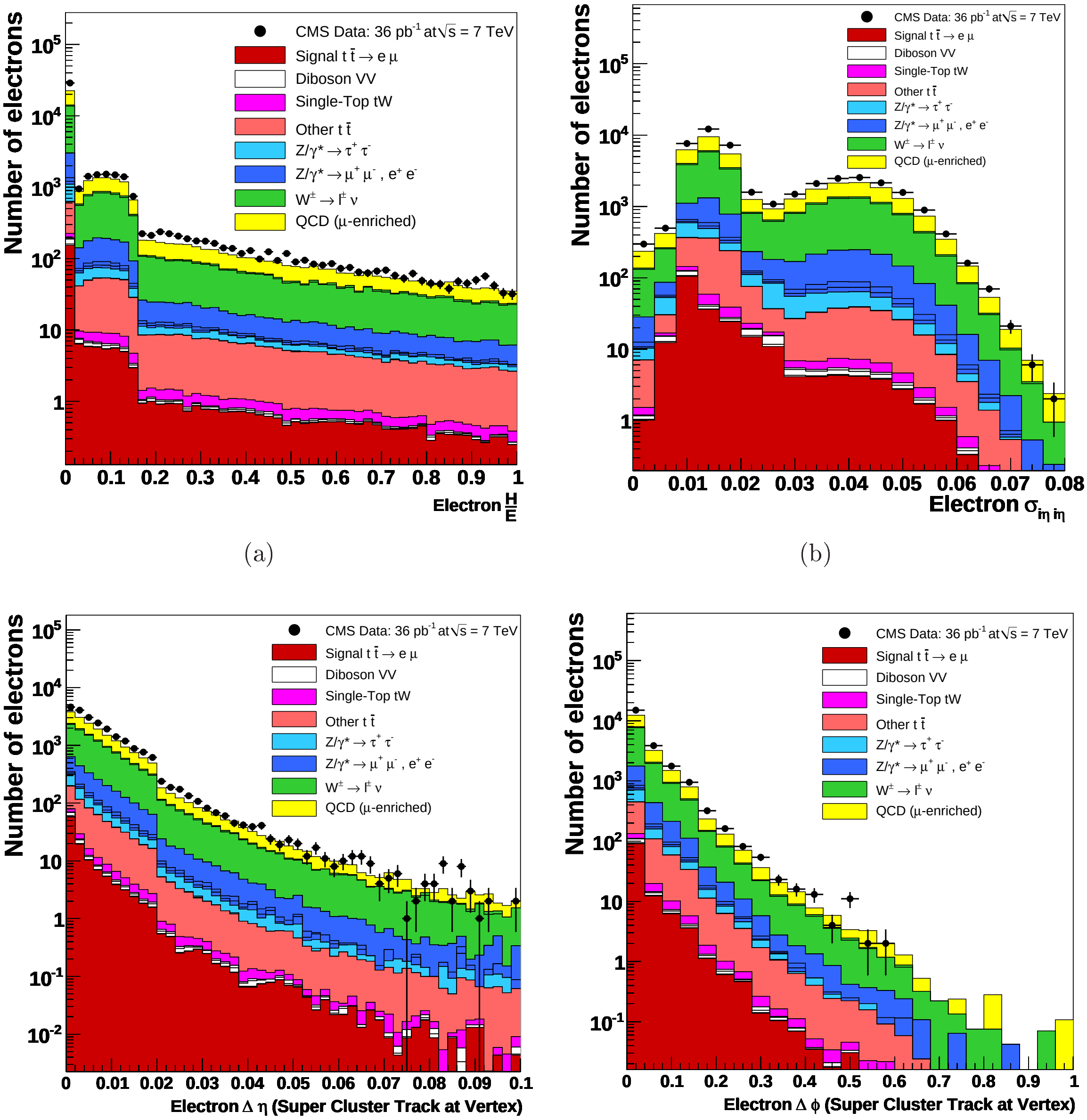
Electron ID variable (*WP90*)		ECAL barrel ($\lvert\eta_{\mathrm{SC}}\rvert < 1.479$)	ECAL endcap ($\lvert\eta_{\mathrm{SC}}\rvert > 1.479$)
$\sigma_{\eta\eta}$	$<$	0.01	0.03
$\Delta\eta_{\mathrm{in}}$	$<$	0.007	-
$\Delta\phi_{\mathrm{in}}$	$<$	0.8	0.7
H/E	$<$	0.12	0.05

Table 6.5: Electron identification variables. The requirements on shower shape ($\sigma_{\eta\eta}$) and track-to-cluster matching variables ($\Delta\eta_{\mathrm{in}}$, $\Delta\phi_{\mathrm{in}}$), well as the ratio of hadronic energy deposit H behind the seed cluster E, are defined separately for ECAL barrel ($|\eta_{\mathrm{SC}}| < 1.479$) and endcap ($|\eta_{\mathrm{SC}}| > 1.479$). Due to a recommendation of the electroweak working group, the requirement on the $\Delta\eta_{\mathrm{in}}$ variable in the endcaps is omitted in the software version, which is used for the reconstruction of the electron candidates (`CMSSW_3_8_6`), as described in [35].

The **conversion rejection** is based on a combination of information from finding a *conversion partner track* and from a requirement on the electron *track hit pattern* [45]. An electron candidate is rejected, if a partner track with opposite charge to the electron is found within $|\Delta\cot(\theta)| < 0.02$, where $\Delta\cot(\theta)$ is defined as:

$$\Delta\cot(\theta) \equiv \frac{1}{\tan(\theta_{\mathrm{trk}})} - \frac{1}{\tan(\theta_{\mathrm{trk}}^{\mathrm{elec}})} . \tag{6.2}$$

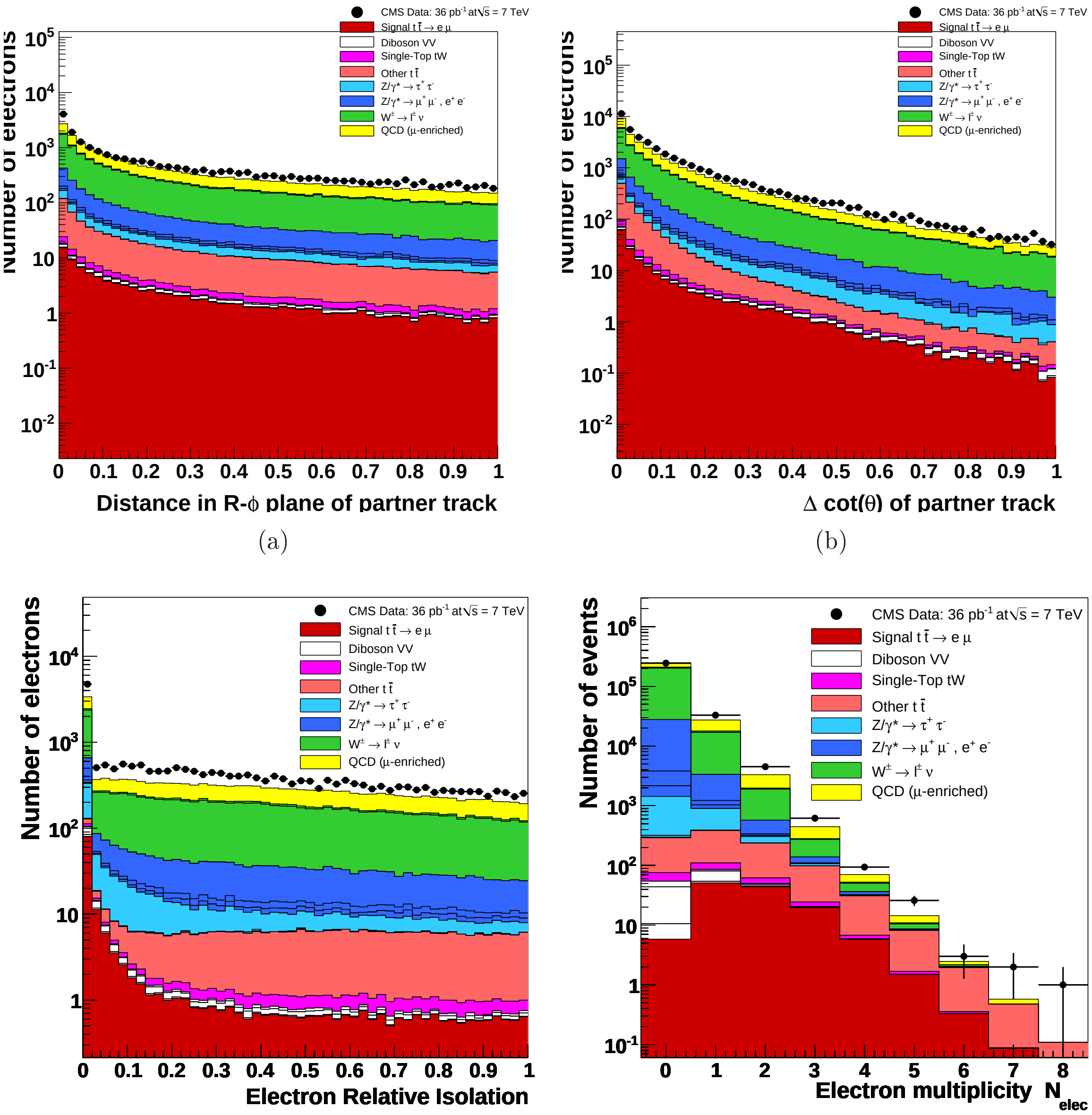
The electron candidate is also rejected, if the distance of closest approach d_{conv} of two tracks in the plane transverse to the beam is less than 200 μm ($|d_{\mathrm{conv}}| < 0.02$ cm). The corresponding distributions of $\Delta\cot(\theta)$ and d_{conv} are shown in figures 6.5 (a) and (b). Electron candidates from conversions occurring beyond the first few layers of the pixel tracker are expected to have more than one missing hit on the inwards propagated track trajectory and are rejected.



(a) (b)

(c) (d)

Figure 6.4: A simple and robust cut-based electron identification is applied on reconstructed electron candidates. A set of selection criteria based on shower shape ($\sigma_{\eta\eta}$) and track-to-cluster matching variables ($\Delta\eta_{\mathrm{in}}$, $\Delta\phi_{\mathrm{in}}$) is chosen based on the optimal selection efficiency of $\mathrm{W}^{\pm} \rightarrow \mathrm{e}^{\pm}\nu_{\mathrm{e}}$ events against QCD background events. The selected events have to pass the full muon selection requirements (**step1**). (a) Ratio of hadronic energy deposit H over the corresponding ECAL seed cluster energy E. (b) Shower shape variable $\sigma_{\eta\eta}$, defined in section 5.2.2. (c) and (d) Track-to-cluster matching variables. Simulated samples are normalised to an integrated luminosity of 35.9 pb^{-1}.



(a) (b)

(c) (d)

Figure 6.5: Electron conversion rejection is based on a veto requirement against conversion partner tracks and missing hits in the electron track hit pattern. All distributions are compared to simulation after the selection of at least one isolated, high-p_{T} muon (**step1**). (a) Distance of closest approach of two tracks in the plane transverse to the beam d_{conv} and (b) $\Delta\cot(\theta)$ defined in equation 6.2. (c) Electron relative combined isolation, which is defined separately for ECAL barrel and endcap regions. (d) Multiplicity of reconstructed electron candidates before applying identification or isolation requirements. Simulated samples are normalised to an integrated luminosity of 35.9 pb^{-1}.

The **electron isolation** is defined as the combined relative isolation given in equations 6.3 and 6.4, which is defined separately for barrel and endcap electrons:

$$I_{\text{comb}}^{\text{rel,barrel}} = \frac{I_{\text{Trk}} + max(0, I_{\text{ECAL}} - 1) + I_{\text{HCAL}}}{p_{\text{T}}(\text{e})} , \tag{6.3}$$

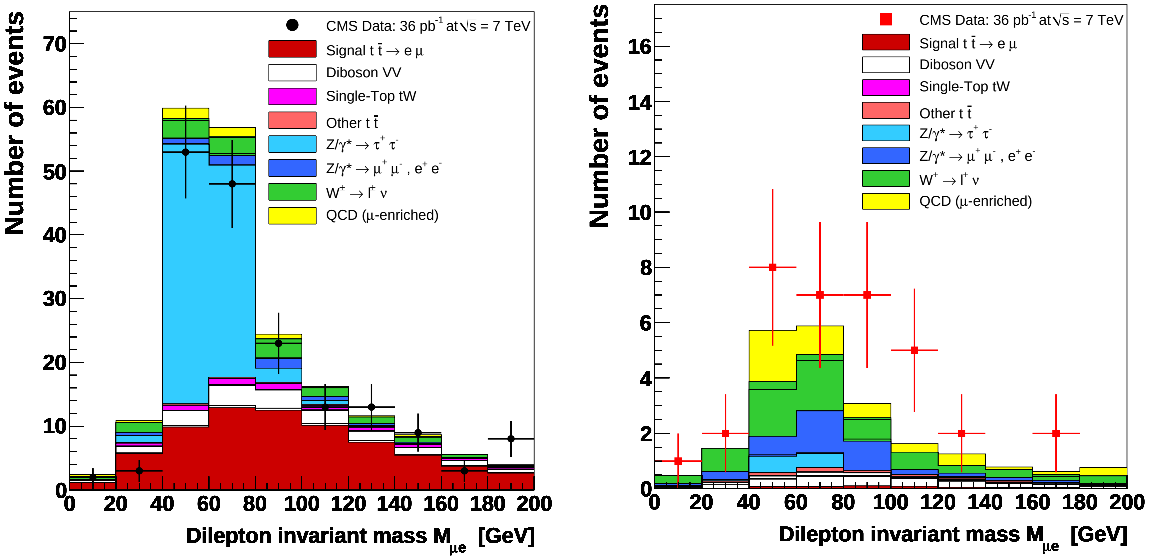
$$I_{\text{comb}}^{\text{rel,endcap}} = \frac{I_{\text{Trk}} + I_{\text{ECAL}} + I_{\text{HCAL}}}{p_{\text{T}}(\text{e})} . \tag{6.4}$$

The isolation is defined similarly as for muons: the values for the calorimeter isolation I_{ECAL} and I_{HCAL} are defined as the amount of transverse energy deposits inside a cone $\Delta R < 0.3$ around the electron candidate. The tracker isolation value I_{Trk} is defined as the sum of transverse momentum of the tracks in a cone of $\Delta R < 0.3$ around the electron. The track and energy deposits associated with the electron candidate are excluded. Isolated electrons are required to have $I_{\text{comb}} < 0.15$ both for the barrel and endcap region. The distribution of the relative combined isolation is shown in figure 6.5, as well as the multiplicity of reconstructed electron candidates before applying any identification or isolation requirements.

6.5 Dilepton Selection

Signal events are required to have at least one isolated, oppositely-charged muon-electron pair. A lepton pair is assigned to a particular decay channel using the two highest-p_{T} leptons in the event, which are fulfilling the isolation and identification criteria. The signature of signal events contains a lepton pair of unlike lepton flavour, therefore no events with invariant dilepton masses near the Z^0 boson mass have to be excluded. Events with a dilepton mass below 12 GeV/c^2 are excluded to reject $\text{Z}^0/\gamma^\star$ events with low masses, which were not simulated, as well as decay products from low-mass resonances.

In order to assure, that both leptons originate from the same, good primary vertex, a *lepton-vertex matching* is performed. Well identified primary vertices are selected according to the event selection described above. The relative difference of the z-position of any selected, isolated lepton to the good primary vertex has to be below 1 cm ($|z_\ell - z_{\text{vtx}}| < 1$ cm). Furthermore, all reconstructed leptons have to be matched to the same primary vertex. Otherwise, the event is rejected. The selection of events with at least one well identified and isolated muon and electron of opposite charge, which originate from a common, well identified vertex, is referred to as *full* $\mu^\pm e^\mp$ selection or *step2*. The distributions of the invariant muon-electron mass for oppositely- and like-charged lepton pairs is shown in figure 6.6. According to the simulation, the dominating remaining background contribution stems from $\text{Z}^0/\gamma^\star \rightarrow \tau^+\tau^-$ events. The contribution from $\text{t}\bar{\text{t}}$ quark pair signal events is already clearly visible and yields an equal number of events, as for the leading background contribution. A detailed discussion of the event yields after full $\mu^\pm\text{e}^\mp$ selection is given below.



(a)

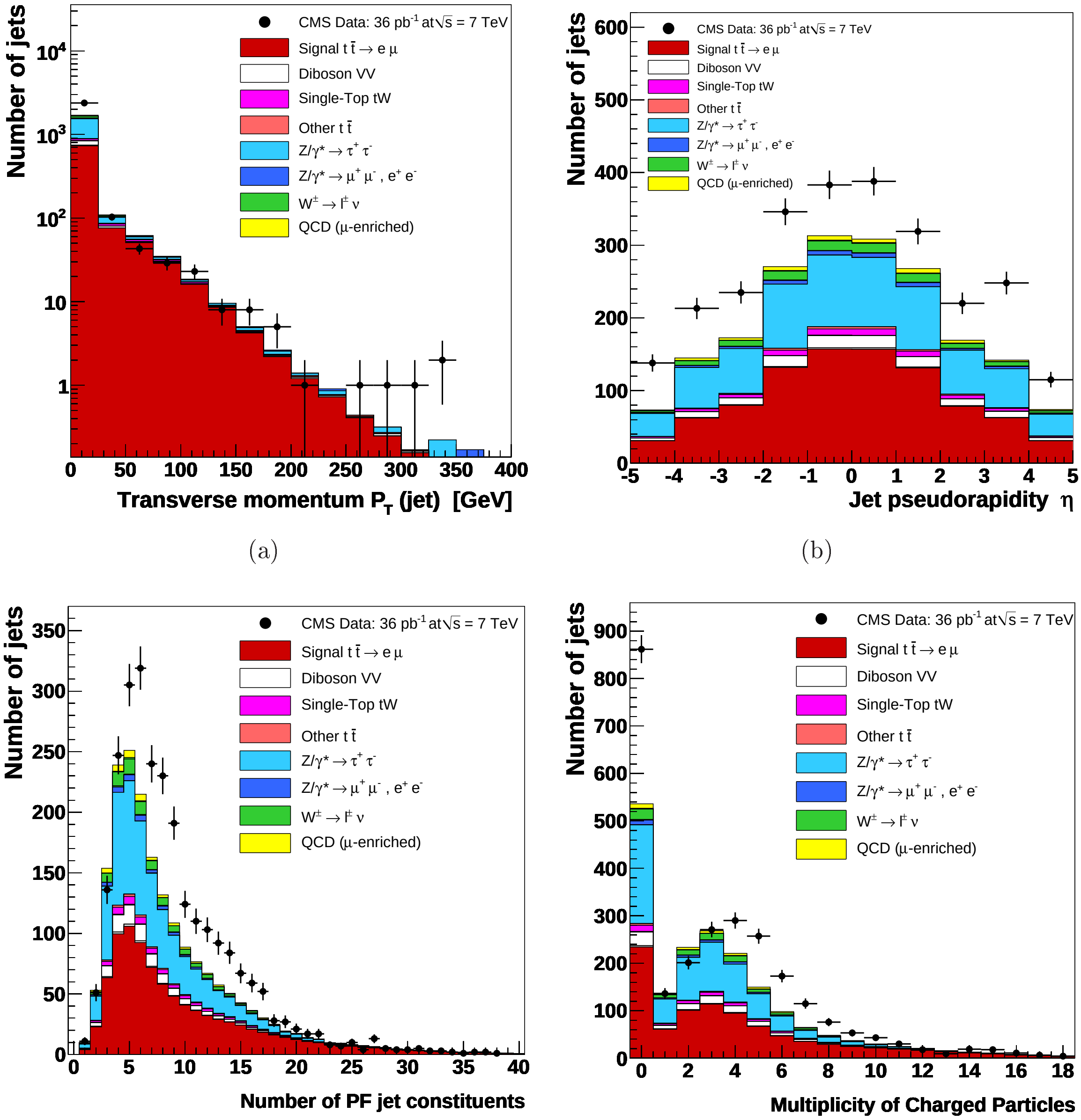
(b)

Figure 6.6: (a) Invariant mass for **oppositely-charged** events passing the full $\mu^{\pm}e^{\mp}$ selection (**step2**). (b) Invariant mass for **equally-charged** $\mu^{\pm}e^{\pm}$ events passing the full muon and electron identification and isolation selection. Simulated samples are normalised to an integrated luminosity of 35.9 pb^{-1}.

6.6 Jet Selection

Signal events are characterised by the presence of high-energetic, hadronic jets corresponding to b jets, and softer jets from initial and final state radiation. Most of the background events are not expected to have as much hadronic activity, therefore the requirement of at least two high-p_{T} jets in the event selection suppresses the remaining background contributions significantly. Jets can be used using three different reconstruction methods, as described in section 5.2.3. All jets are reconstructed using the Anti-k_{T} cluster algorithm with a clustering parameter equal to 0.5 (*AK5 jets*). All jet algorithms were shown to give the same performance in terms of expected signal and background yields [46]. The expected jet energy scale (JES) uncertainty, which is one of the largest systematic uncertainties to the cross section measurement, is found to be 5% for JPT and PF jets and 10% for calo jets [47]. Hence, Particle-Flow jets are used in this analysis.

The momentum of the jets is corrected using a relative and an absolute scale correction (L2 and L3 corrections) both in data and simulation. Jets in collision data are further corrected by the *residual correction* (`Spring10DataV2`), derived on older simulation samples and on early collision data [47]. The event selection requires at least two PF jets with corrected p_{T} greater than 30 GeV/c, within the pseudorapidity region of $|\eta| < 2.5$. Jet candidates are removed from the jet collection, if they are overlapping with one of the selected leptons within a cone $\Delta R(\mathrm{jet}, \mu/\mathrm{e}) < 0.4$ around the jet. In order to suppress detector noise, a simple cut-based **jet identification** selection is applied, which is referred to as "*loose ID*", as described in [48]. In table 6.6, all requirements on the PF jet identification variables are summarised.



(a) (b)

(c) (d)

Figure 6.7: Distributions of reconstructed Particle-Flow jets for events passing the full $\mu^\pm e^\mp$ selection (**step2**). (a) Transverse momentum p_T of PF jets and (b) pseudorapidity η. PF jet identification variables: (c) Number of PF jet constituents and (d) multiplicity of charged particles inside the jet. Simulated samples are normalised to an integrated luminosity of 35.9 pb^{-1}.

PF jet ID variable		(*loose* ID)
Neutral hadronic fraction	$<$	0.99
Neutral e.m. fraction	$<$	0.99
Number of PF jet constituents	$>$	1
`if` $(\lvert\eta\rvert < 2.4)$		
Charged hadronic fraction	$>$	0
Number of charged particles	$>$	0
Charged e.m. fraction	$<$	0.99

Table 6.6: Cut-based PF jet identification [48].

The distribution of PF jets reconstructed by the Anti-k_T algorithm with clustering parameter equal to 0.5 are compared to simulation after the full $\mu^{\pm}\mathrm{e}^{\mp}$ selection, shown in figure 6.7. Figure 6.7 (a) shows the transverse momentum of PF jets, which shows a deviation between data and simulation of approximately 25%, which originates exclusively from low-p_T jets with $p_\mathrm{T} < 20$ GeV/c, and therefore may originate from pileup events, which are not modelled in the simulation. Despite this deviation, the shape of the η-distribution, shown in (b) is described reasonably by the simulation. In figures (c) and (d), two jet ID variables are shown. Both, the number of PF jet constituents, as well as the multiplicity of charged particles inside a jet are lower in simulation compared to data, due to not properly simulated pileup events.

For the measurement of the top quark pair production cross section, events with at least two well identified, high-p_T jets are selected. In order to confirm the top quark-like topology of the selected events, the multiplicity of b-tagged jets is shown in figures 6.8 (c), (d) and 6.10 (c), (d). Two b-tag algorithm are used: the simple secondary vertex high-efficiency (SSVHE) tagger with a discriminator value greater than 1.74 (*medium* working point) and the track-counting high efficiency (TCHE) tagger with a discriminator value greater than 1.7 (*loose* working point) [37]. The corresponding efficiencies of the different working points are described in section 5.2.4.

6.7 Missing Transverse Energy Selection

Missing transverse energy (MET) $\not{E}_\mathrm{T}$ is a characteristic feature of top quark pair decays. It is natural in signal events, due to the presence of two neutrinos from the leptonically decaying W bosons. Requiring a significant amount of reconstructed MET in the event allows for the suppression of the remaining like-flavour Drell-Yan and QCD background events after the full $\mu^{\pm}\mathrm{e}^{\mp}$ selection, which do not contain a natural source of MET. Similar to the different jet reconstruction algorithms, there are three types of reconstructed MET available in CMS [38]: calorimeter-based $\not{E}_\mathrm{T}$ corrected for jet and muon response (*calo* MET), calorimeter-based $\not{E}_\mathrm{T}$ corrected using tracker response (*tcMET*) and $\not{E}_\mathrm{T}$ reconstructed from Particle-Flow objects (*PF MET*).

After requiring at least two high-p_T PF jets in the event, the contributions from QCD and like-flavour DY background events is found to be negligible. Therefore, no further requirement on missing transverse energy is applied in the $e^{\pm}\mu^{\mp}$ final state. Nevertheless, the PF MET distributions for collision data and simulated events are compared after several selection steps in the following section, because it is a characteristic feature of selected signal events.

6.8 Event Yields

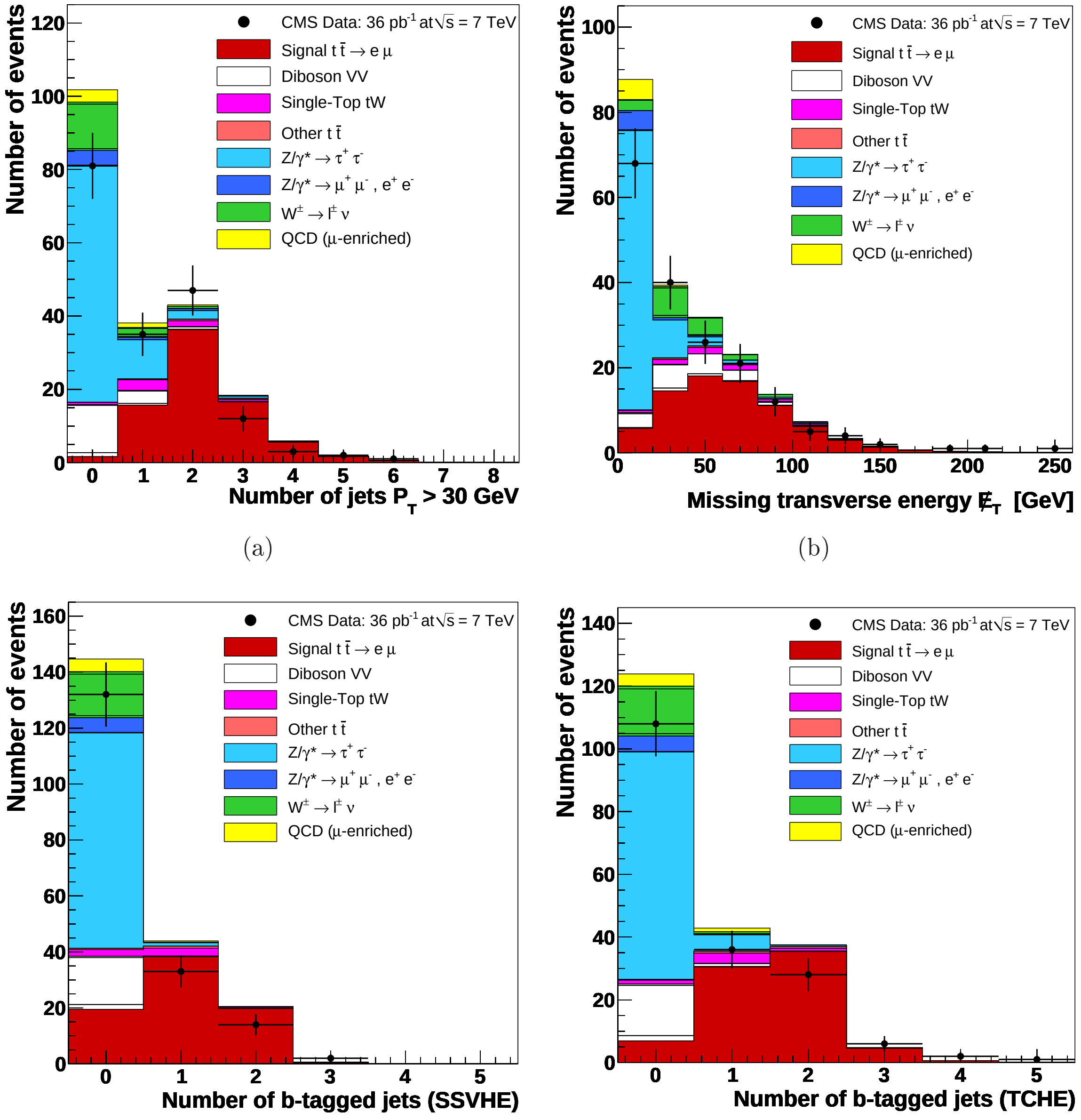
The numbers of selected events after each selection step are summarised in table 6.7. Since the QCD sample is already preselected on generator level for at least one muon with p_T greater than 15 GeV/c, the comparison between events from collision data and simulated events starts after applying the full muon identification and isolation requirements (*step1*). There is good agreement between selected events in data and the sum of simulated events after requiring one isolated muon within 1.1%. The distributions of the main discriminating variables are compared for events passing the dilepton selection after applying the lepton isolation and identification requirements (*step2*). Furthermore, the multiplicity of b-tagged jets is shown for two different b tagging algorithms.

Sample	1 iso. muon	2 iso. leptons	$\geq$ 2 jets	$\geq$ 1 b-tag
$t\bar{t}$ signal (μe)	127	78.0	58.0	53.6
$t\bar{t}$ other	771	1.3	1.0	0.9
single-top tW	62	5.7	1.9	1.6
Diboson VV	91	19.2	0.8	0.2
$Z^0/\gamma^\star \rightarrow \tau^+\tau^-$	1 739	78.2	2.5	0.7
$Z^0/\gamma^\star \rightarrow e^+e^-$, $\mu^+\mu^-$	28 952	5.5	0.5	0.1
W+Jets	196 227	16.6	0.5	0.1
QCD	51 468	5.2	0.4	0.2
Total sim. backgr.	279 310	131.7	7.6	3.8
Total simulated	279 437	209.7	65.6	57.4
Data	282 603	181	59	51

Table 6.7: Number of expected signal and background events, compared to the event yields in 35.9 pb^{-1} of data collected by the CMS experiment at $\sqrt{s} = 7$ TeV after different selection steps. Background estimates are taken from simulation only here.

Full $\mu^{\pm}e^{\mp}$ Selection

After selecting events requiring an oppositely-charged, well identified and isolated muon-electron pair, the top dilepton signal is already visible. According to the simulation, events from $Z^0/\gamma^\star \rightarrow \tau^+\tau^-$ processes are the dominating background contribution at this selection step, which is shown in the jet multiplicity and $\not{E}_T$ distributions in figures 6.8 (a) and (b). A deviation between events in data and simulation is also observed in events with three jets, dominated by $t\bar{t}$ quark pair events. It should be noted, that a deviation between the number of events in data and simulated is observed, where the dominating

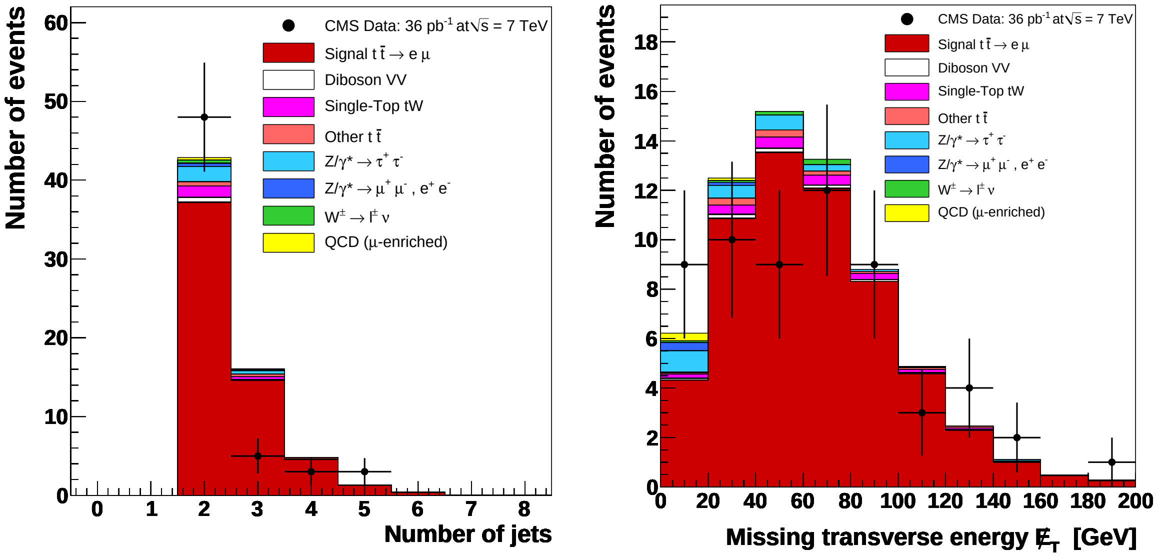


(a) (b)

(c) (d)

Figure 6.8: (a) Jet multiplicity for particle-flow jets with $p_T > 30$ GeV/c and (b) missing transverse energy $\not\!\!E_T$ for events passing the full $\mu^\pm e^\mp$ selection (**step2**). (c) Number of b-tagged jets for events passing the full $\mu^\pm e^\mp$ selection, tagged by the SSVHE algorithm *medium* working point and (d) tagged by the TCHE algorithm *loose* working point. Simulated samples are normalised to an integrated luminosity of 35.9 pb^{-1}.

contribution of simulated events stems from samples, which are generated with the D6T tune. There is good agreement between data events and simulated events in the invariant dilepton mass distributions within uncertainties, shown in figures 6.6 (a) and (b), both for same-sign and opposite-sign muon-electron pairs. Figures 6.8 (c) and (d) show good agreement between two-tag and one-tag events in the $\mu^{\pm}e^{\mp}$ final state, which shows, that the performance of the b-tagging algorithms is modeled well.



(a) (b)

Figure 6.9: (a) Jet multiplicity for Particle-Flow jets passing the dilepton-plus-jets selection (**step4**). (b) Distribution of missing transverse energy for events at the same stage of the event selection. Simulated samples are normalised to an integrated luminosity of 35.9 pb^{-1}.

Dilepton-Plus-Jets Selection

In addition to the full $\mu^{\pm}e^{\mp}$ selection, the presence of two high-energetic, well identified PF jets is required. The same distributions as shown for the full $\mu^{\pm}e^{\mp}$ selection are compared for events in data and simulation with at least two PF jets with $p_{\mathrm{T}} > 30$ GeV/c, $|\eta| < 2.5$ and a loose PF jet identification requirement, which are shown in figures 6.9 and 6.10. According to the simulation, the $\mathrm{t\bar{t}}$ quark pair signal is already dominating the composition of the selected events. After the selection of events with two high-p_{T} PF jets, without an additional requirement of missing transverse energy, a significance of the top signal of $S/\sqrt{S+B} = 6.7$ is obtained, which corresponds to a signal over background ratio of $S/B = 6.8$. The presence of an excess over the background expectations is clearly visible from the $\mu^{\pm}e^{\mp}$ decay channel alone. Figure 6.9 (a) shows, that almost all remaining background events are contained in the 2-jet bin of the jet multiplicity distribution. A deficit of selected events in the data relative to simulation with exactly three jets is obsevered, as shown in figure 6.9 (a), and needs to be investigated in more detail, if it persists in a larger data sample. Figure 6.9 (b) shows a reasonable agreement between data and simulation within statistical uncertainties in the full range of missing transverse energy. Figures 6.10 (a) and (b) show good agreement for the distributions of the SSVHE and the TCHE b-tagging discriminator. In figures 6.10 (c) and (d) the corresponding b-tagged jet multiplicities are compared for SSVHE *medium* and TCHE *loose* working points. An overall good agreement is shown for the TCHE algorithm, while the efficiency for the SSVHE algorithm is expected to be lower in data than in simulation. The TCHE and SSVHE b-tagging algorithms are studied in more detail in section 7.4.

6.9 Selected Equally-Charged Lepton Pairs

It is studied, if the contribution of background events in the final state with misidentified, non-genuine leptons can be estimated by a data-driven method. Those processes, which do not contain two real leptons of unlike lepton flavour, are QCD multijet events, production of W bosons with additional jets and $Z^0/\gamma^\star \rightarrow e^+e^-$, $\mu^+\mu^-$ events. The remaining background contributions, which are considered as *irreducible background* events, are estimated from simulated events, such as $Z^0/\gamma^\star \rightarrow \tau^+\tau^-$ events, single-top quark tW production, other $t\bar{t}$ quark pair decays and diboson events (WW, WZ and ZZ).

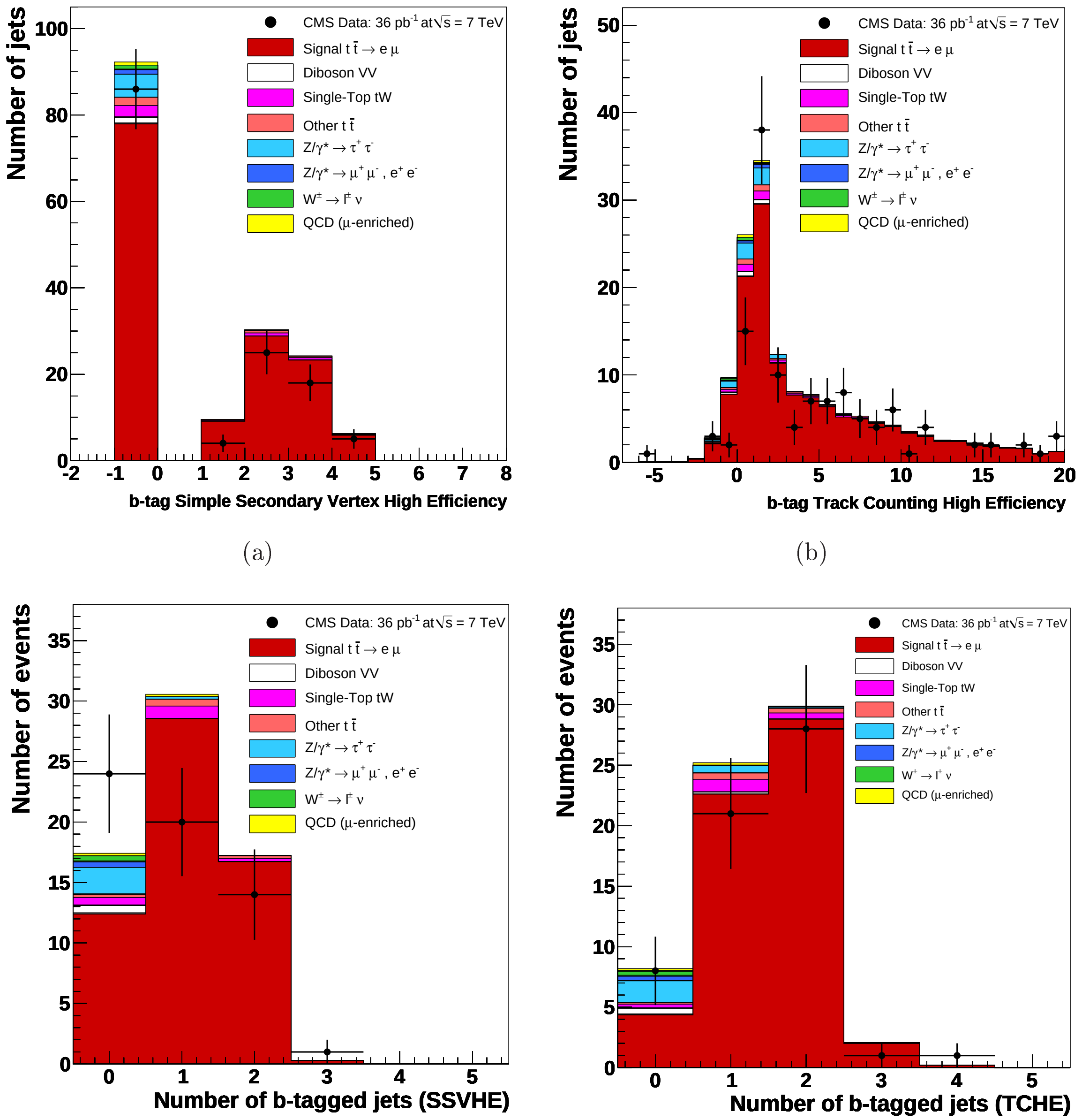
In order to estimate the contribution from events with at least one misidentified lepton, the number of events in the final state with **equally-charged** muon-electron pairs is measured in the data. Equally-charged lepton pairs are referred to as "*wrong-charge*" (WC) events $N_{\mathrm{WC}}^{\mathrm{BKG}}$, in contrast to **oppositely-charged** lepton pairs from signal events and irreducible background events, which are referred to as "*right-charge*" (RC) events. The contribution of background events with right-charge lepton pairs $N_{\mathrm{RC}}^{\mathrm{BKG}}$ is estimated by multiplying the number of wrong-charge events by the ratio of RC over WC events $R_{\mathrm{RC/WC}}^{\mathrm{Sim}}$, which is derived from simulation, given in equation 6.5:

$$N_{\mathrm{RC}}^{\mathrm{BKG}} = N_{\mathrm{WC}}^{\mathrm{BKG}} \times R_{\mathrm{RC/WC}}^{\mathrm{Sim}} . \tag{6.5}$$

After applying the requirement of two isolated leptons, we obtain an almost singal-free background sample for equally-charged muon-electron pairs, shown in figure 6.6 (b). We obtain a ratio of RC over WC events for events with an isolated muon-electron pair of $R_{\mathrm{RC/WC}}^{\mathrm{Sim}} \approx 1.7$. After applying the dilepton-plus-jets selection requirements, **6 WC events** are observed in data. The expected number of remaining WC events from other background processes, not estimated by the *RC/WC*-method, such as diboson or $Z^0/\gamma^\star \rightarrow \tau^+\tau^-$ events, is derived from simulation to be 1.0 events and is subtracted. The obtained number of WC background events is $N_{\mathrm{WC}}^{\mathrm{BKG}} = 6-1.0 = 5.0$. The number of RC background events is calculated as $N_{\mathrm{RC}}^{\mathrm{BKG}} = 5.0 \times 1.7 \approx 8.5$. This value is compared to the number of $N_{\mathrm{RC}}^{\mathrm{Sim}} = 1.5$ events, which are expected from simulation and is translated into a correction factor $\mathcal{C}_{\mathrm{RC}}^{\mathrm{BKG}}$ to the expected number of background events from misidentified leptons:

$$\mathcal{C}_{\mathrm{RC}}^{\mathrm{BKG}} = \frac{N_{\mathrm{RC}}^{\mathrm{BKG}}}{N_{\mathrm{RC}}^{\mathrm{Sim}}} \approx 5.7 . \tag{6.6}$$

Until this discrepancy is studied in further detail, the *RC/WC*-method is *not* considered as an estimation for the background contribution in the muon-electron final state. The number of background events obtained from simulated events is compared to another data-driven method, described in section 7.1.



(a) (b) (c) (d)

Figure 6.10: (a) Distribution of the simple secondary vertex high efficiency (SSVHE) b-tagging algorithm for the *medium* working point and (b) of the track counting high efficiency (TCHE) b-tagging algorithm for the *loose* working point after dilepton-plus-jets selection (**step4**). (c) Number of b-tagged jets for events passing the dilepton-plus-jets selection, tagged by the SSVHE algorithm *medium* working point and (d) tagged by the TCHE algorithm *loose* working point. Simulated samples are normalised to an integrated luminosity of 35.9 pb^{-1}.

Chapter 7

Cross Section Measurements

7.1 Systematic Uncertainties

7.1.1 Experimental Uncertainties

Tag&Probe Method

The general idea of the *tag&probe* method is the determination of a certain event property (e.g. single-muon trigger efficiency) and its absolute performance in data. Therefore, a **clean physics event signature** (e.g. $Z^0 \to \mu^+\mu^-$) is chosen, on which the property is tested, without using this property in the event selection (*tag* events). The event property is tested on the selected event sample (*probe* events) and the corresponding efficiency is determined by calculating the ratio of events, which fulfil both the tag and probe requirements $N^{\mathrm{tag\&probe}}$ over the total number of tagged events N^{tag}. In general, the probed event property (single-muon trigger efficiency) is not necessarily identical with the tag requirement (two reconstructed, oppositely-charged, high-p_{T} muons). In order to ensure, that the selected events are recorded, the tagged events are also required to have fired the same single muon trigger as the probed events, which is *not* a necessary condition for the tag&probe method.

Muon Trigger Efficiency from $Z^0 \to \mu^+\mu^-$ Events

The single-muon trigger efficiency is measured in dimuon events with an invariant mass near the Z^0 mass peak, also passing the muon identification and isolation requirements. The selected, reconstructed muons are matched to muon candidates from the trigger object collection. The selected events are required to have at least one pair of oppositely-charged, isolated and identified muons, which fulfil the same object selection criteria as described above. The selected muons are required to have the same kinematic properties as muons in the top dilepton selection. Additionally, an invariant mass of the muon pair between 76 and 106 GeV/c^2 ($M_{Z^0}\pm$ 15 GeV/c^2), referred to as the **Z^0 window**, is required. The reconstructed muons are matched to muon candidates from the high-level trigger (HLT) object collection within a matching radius of $\Delta R(\mu^{\mathrm{RECO}}, \mu^{\mathrm{HLT}}) < 0.1$ in $(\eta - \phi)$ space. Events are rejected, if the distance between the two trigger objects or reconstructed muons is less than twice the matching radius or both muon candidates are matched to the same trigger object. At least one successful match between a triggered and a reconstructed muon is required. The number of events with *at least one* positive match is referred to as the number of *tag* events T. Events with *two matches* are referred to as the number of *tag&probe* events T+P, in which *both* muons are entered. The trigger efficiency ε is calculated from the number of muons, in which *exactly* two muon candidates are successfully matched to HLT muons (T+P events), divided by the number of muons, in which *at least* one muon candidate is matched to an HLT muon candidate (T *and* T+P events). The trigger efficiency is defined as:

$$\varepsilon = \frac{\text{T+P}}{\text{T} + \text{T+P}} = \frac{N^{\text{tag\&probe}}}{N^{\text{tag}}} \; . \tag{7.1}$$

In figures 7.1 (a) and (b) the transverse momentum p_{T} and the invariant mass $M_{\mu\mu}$ of the reconstructed muon pair are shown. The three different colours represent different triggers, which have been used for data taking, listed in table 6.3. Simulated events are selected using only the `HLT_Mu9` trigger [40]. Both figures show good agreement between selected events in data and simulation. Figures 7.1 (c) and (d) show the distributions of the measured single-muon trigger efficiencies as functions of the transverse momentum p_{T} of the muon and its pseudorapidity η for simulated and data events, triggered by the individual trigger paths. The trigger efficiencies obtained from simulation are higher than those obtained from data. For the comparison of the trigger efficiencies in data with simulation, only the `HLT_Mu9` trigger is shown here, because it is the only muon trigger path, which is used for the events selection on simulated data. The trigger efficiencies are approximately constant functions of the muon p_{T} , with slightly higher values for higher muon momenta. For the `HLT_Mu15_v1` and `HLT_Mu11` triggers, also a nearly constant response as a function of the muon pseudorapidity η is observed, while the efficiency for the `HLT_Mu9` triggered data sample is decreased significantly in the central region. This corresponds to the smallest amount of recorded luminosity (3.2 pb^{-1}) and therefore can be treated as negligible.

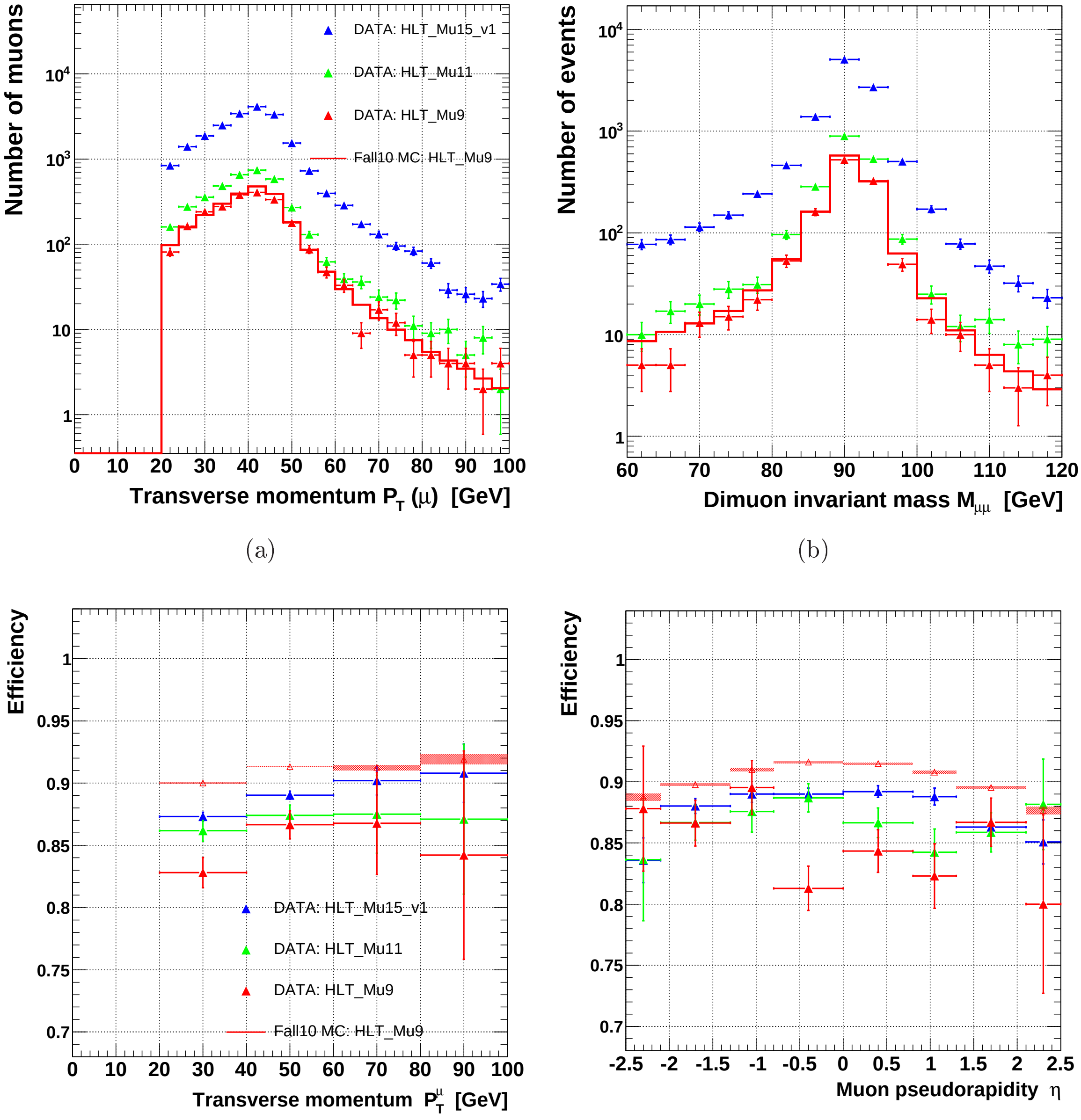
Muon Trigger Efficiency Correction

The total **single-muon trigger efficiency** for muons with $p_{\text{T}} > 20$ GeV/c and $|\eta| < 2.4$ is measured as:

$$\varepsilon_{\text{trig}}^{\text{Data}}(\mu) \;=\; 87.76 \pm 0.38\% \; . \tag{7.2}$$

The total trigger efficiency is obtained from the weighted sum of the three single-trigger efficiencies, which are weighted by the fraction of the integrated luminosity of the corresponding run period. Please note, that for the measurement of the trigger efficiencies in data, disjoint run periods have been used, as described in table 6.3. Therefore, it can be explained, that the lowest threshold trigger `HLT_Mu9` is found to have the lowest efficiency, whereas a higher efficiency is expected on the full 2010 data sample. From simulated events, the trigger efficiency yields:

$$\varepsilon_{\text{trig}}^{\text{Sim}}(\mu) \;=\; 90.71 \pm 0.04\% \; . \tag{7.3}$$



(a) (b)

(c) (d)

Figure 7.1: Muon trigger efficiencies measured with $Z^0 \to \mu^+\mu^-$ events in data and simulation. An invariant mass between 76 and 106 GeV/c^2 around the Z^0 mass peak is selected for the muon efficiency measurement based on the *tag&probe* method. (a) Muon transverse momentum p_T for selected muons, which pass all muon quality requirements. (b) Dimuon invariant mass spectrum for muon candidates with $p_T > 20$ GeV/c and $|\eta| < 2.4$. Events from data are selected by different muon trigger paths, shown in table 6.3 and compared to simulated $Z^0 \to \mu^+\mu^-$ events, selected by the same muon triggers. (c) Efficiency as a function of the transverse momentum p_T and (d) as a function of the pseudorapidity η of the muons with $p_T > 20$ GeV/c and $|\eta| < 2.4$. Simulated events are selected using only the `HLT_Mu9` trigger [40].

In table 7.1 the measured trigger efficiencies are summarised and compared to simulation. The total selection efficiency $\varepsilon_{\mathrm{tot}}$ used for the cross section calculation in the following chapter, which is obtained from simulation, is corrected for the deviation between efficiency in data and simulation by applying a **correction factor**. The correction factor $\mathcal{C}_{\mathrm{trig}}$ for single-muon triggered events is defined as the ratio of the trigger efficiency measured in data over the efficiency obtained from simulation, resulting in a correction factor of (1 - 3.25%), which is given in equation 7.4:

$$\mathcal{C}^{\mu}_{\mathrm{trig}} = \frac{1-(1-\varepsilon^{\mathrm{Data}}_{\mathrm{trig}}(\mu))}{1-(1-\varepsilon^{\mathrm{Sim}}_{\mathrm{trig}}(\mu))} = \frac{\varepsilon^{\mathrm{Data}}_{\mathrm{trig}}(\mu)}{\varepsilon^{\mathrm{Sim}}_{\mathrm{trig}}(\mu)} = 0.9675 \ . \tag{7.4}$$

Sample	Trigger	Efficiency [%]	L_{int} [pb^{-1}]
Simulation	`HLT_Mu9`	90.71 ± 0.04	
Data	`HLT_Mu9`	84.64 ± 0.81	3.2
Data	`HLT_Mu11`	86.82 ± 0.58	5.0
Data	`HLT_Mu15_v1`	88.29 ± 0.23	27.7
Data	combined	87.76 ± 0.38	35.9

Table 7.1: Single-muon trigger efficiencies in data and simulation obtained from dimuon events with an invariant mass close to the Z^0 mass peak (tag&probe method).

Lepton Selection Efficiency Corrections

First, the total selection efficiency of the event selection is estimated using simulated events. Then, several *correction factors* $\mathcal{C}$ are applied to the total efficiency, in order to account for differences between events in data and simulated events. As a first approximation, the total selection efficiency is assumed to be a product of the individual selection steps, relative to the previous selection steps, such as trigger selection and lepton selections:

$$\epsilon_{\mathrm{tot}} = \epsilon_{\mathrm{trig}} \times \epsilon_{\mathrm{reco}} \times \epsilon_{\mathrm{id}} \times \epsilon_{\mathrm{iso}} \ . \tag{7.5}$$

Lepton efficiencies are measured using the tag&probe method on the Z-mass window, as described above. The **single-muon trigger efficiency** was measured in data events to be on average 87.76 ± 0.38%. The corresponding correction factor on the single-muon trigger efficiency was calculated to be $\mathcal{C}^{\mu}_{\mathrm{trig}} = 0.9675$. The **single-electron trigger efficiency** is also measured using the tag&probe method. The resulting value is measured to be 96.0%, taken from [49]. The corresponding efficiency based on simulated events is 97.0%. The resulting scale factor on the single-electron trigger efficiency is calculated to be $\mathcal{C}^{\mathrm{e}}_{\mathrm{trig}} = 0.990$.

Top quark pair events decaying into a muon-electron pair are selected by requiring either a single-muon or single-electron trigger, as described in section 6.2. The combined trigger efficiency of two uncorrelated single-object triggers, is given by the efficiency of one trigger, adding the product of its inefficiency times the efficiency of the other trigger. Therefore, the **combined muon-electron trigger efficiency** is determined, using the following equation:

$$\epsilon^{\mathrm{e}\mu}_{\mathrm{trig}} \;\equiv\; \epsilon^{\mathrm{e}}_{\mathrm{trig}} + (1-\epsilon^{\mathrm{e}}_{\mathrm{trig}})\;\epsilon^{\mu}_{\mathrm{trig}} \;=\; \epsilon^{\mu}_{\mathrm{trig}} + (1-\epsilon^{\mu}_{\mathrm{trig}})\;\epsilon^{\mathrm{e}}_{\mathrm{trig}}\ . \tag{7.6}$$

Corresponding to equation 7.6, the muon-electron trigger efficiency, based on single-lepton efficiencies measured in data, is found to be 99.6 ± 0.1%, compared to an efficiency of 99.7% derived from simulated events [49]. The total combined trigger efficiency scale factor, which also accounts for the effect of *pre-firing*, is $\mathcal{C}^{\mu\mathrm{e}}_{\mathrm{trig}} = 0.994$, as given in [49].
In the following, a summary on lepton reconstruction, identification and isolation efficiency correction factors is given. The values are taken from [49]. The efficiencies are determined by using the tag&probe method on the Z-mass window, as described above. In order to account for the possible differences in the efficiencies derived from Z and $\mathrm{t\bar{t}}$ events, Z events are selected with an additional requirement of at least two additional jets. The differences, compared to the standard Z selection, give a measure for the systematic uncertainty of the lepton identification an isolation selections. The total difference is quoted to be within 4%, which leads to the assignment of a systematic error of 2% per lepton, given in [49].
The correction factors $\mathcal{C}^{\mu}_{\mathrm{reco/id/iso}}$ and $\mathcal{C}^{\mathrm{e}}_{\mathrm{reco/id/iso}}$ for *each lepton* are given first, before they are combined with the combined trigger efficiency $\mathcal{C}^{\mu\mathrm{e}}_{\mathrm{trig}}$, which is given above. A summary of all lepton efficiency correction factors is given in table 7.2.
The combination of the **single-electron** efficiency correction factors, accounting for reconstruction, identification, isolation and electron energy scale, yields $\mathcal{C}^{\mathrm{e}}_{\mathrm{reco/id/iso}} = 0.961$. The corresponding combined **single-muon** efficiency correction factor is stated as $\mathcal{C}^{\mu}_{\mathrm{reco/id/iso}} = 0.992$. This gives a correction factor, accounting for reconstruction, identification and isolation, for each muon-electron event of $\mathcal{C}^{\mu\mathrm{e}}_{\mathrm{reco/id/iso}} = 0.953$. The overall correction factor $\mathcal{C}^{\mu\mathrm{e}}_{\mathrm{tot}}$, including the trigger efficiency corrections, which is applied to each dilepton event, is given in the following equation:

$$\mathcal{C}^{\mu\mathrm{e}}_{\mathrm{tot}} \;=\; \mathcal{C}^{\mu\mathrm{e}}_{\mathrm{trig}} \times \mathcal{C}^{\mu\mathrm{e}}_{\mathrm{reco/id/iso}} \;=\; 0.947\ . \tag{7.7}$$

Lepton efficiency ϵ	Correction factor $\mathcal{C}$
$\epsilon^{\mu\mathrm{e}}_{\mathrm{trig}}$	$\mathcal{C}^{\mu\mathrm{e}}_{\mathrm{trig}} = 0.994$
$\epsilon^{\mathrm{e}}_{\mathrm{reco}}\ /\ \epsilon^{\mathrm{e}}_{\mathrm{id}}\ /\ \epsilon^{\mathrm{e}}_{\mathrm{iso}}$	$\mathcal{C}^{\mathrm{e}}_{\mathrm{reco/id/iso}} = 0.961$
$\epsilon^{\mu}_{\mathrm{reco}}\ /\ \epsilon^{\mu}_{\mathrm{id}}\ /\ \epsilon^{\mu}_{\mathrm{iso}}$	$\mathcal{C}^{\mu}_{\mathrm{reco/id/iso}} = 0.992$
	$\Rightarrow\ \mathcal{C}^{\mu\mathrm{e}}_{\mathrm{reco/id/iso}} = 0.953$
$\epsilon^{\mu\mathrm{e}}_{\mathrm{tot}}$	$\mathcal{C}^{\mu\mathrm{e}}_{\mathrm{tot}} = 0.947$

Table 7.2: Summary of all lepton selection efficiency correction factors. The various efficiencies ϵ are measured in data events using the tag&probe method on the Z-mass window. The value of the efficiency obtained from data is compared to the efficiency calculated from simulated events. Several *correction factors* $\mathcal{C}$ are applied to the efficiencies, which are assumed to factorise, according to equation 7.6, in order to account for differences between events in data and simulated events. Except for the single-muon trigger efficiency, determined in this analysis, all values are taken from [49].

Jet Energy Scale Uncertainties

The uncertainties arising from the **energy scale of jets and missing energy** (JES), as well as energy resolution effects, are reported in [49]. To estimate the systematic effect of the JES, jets and the hadronic part of MET are varied up and down simultaneously by 5%. The rate of $\mathrm{t\bar{t}}$ quark pair events with at least two jets is compared, using two POWHEG samples, one showered with PYTHIA and the other with HERWIG, as described in [49]. The resulting variation of selected signal events is found to be approximately ±3.0%. The presence of pileup events in proton-proton collisions mainly effects the missing transverse energy selection. The systematic effect on the event selection, due to pileup is not taken into consideration, because no $\not{E}_\mathrm{T}$ selection is applied in the muon-electron channel.

B-Tagging Efficiency

The uncertainty on the selection of events, which contain at least one jet *identified as b jet*, is also reported in [49]. The **systematic uncertainty** on the selection of b-tagged jets is found to be 5.0%, based on the relative variation of the number of b-tagged events, due to a variation of the b-tagging efficiency of 10%.

Background Estimation

In order to calculate the $\mathrm{t\bar{t}}$ quark pair production cross section, the number of observed events after full event selection has to be corrected by the number of expected background events in the final state. The number of background events is obtained from simulated samples. According to the simulation, the remaining background events are mainly composed of associated single-top quark and W boson (tW) production, diboson production, as well as $\mathrm{Z}^0/\gamma^\star \rightarrow \tau^+\tau^-$ events. All of these background processes are considered as *irreducible background processes* containing a real μe pair and therefore, an estimation of the number of background events in the final state is obtained from simulation. The amount of background events stemming from other decay modes of $\mathrm{t\bar{t}}$ quark pairs, $\mathrm{Z}^0/\gamma^\star \rightarrow e^+e^-$, $\mu^+\mu^-$, QCD or W boson events, which contain at least one or more misidentified lepton candidates ("*fake*" leptons) is also obtained from simulation, because the total amount of background is expected to be small, compared to the number of signal events: the signal-over-background ratio is almost seven ($S/B = 6.8$), after the full event selection, without any requirement of MET or b-tagged jets. For each of the irreducible background contributions, an uncertainty of 30% is assigned to the total number of expected events in the final state, as proposed in [49].

The number of background events containing at least one or more misidentified leptons is compared to the result of a *data-driven background estimation*, described in [49]. The number of background events estimated from simulation yields 2.4 events, compared to $\left(1.5^{+1.6}_{-1.5}\right)$ events estimated from data. In order to account for the larger uncertainty of the modelling of misidentified leptons in the simulated samples, an increased systematic error of 50% is assigned to the total number of expected fake lepton events in the final state. In table 7.3, all background contributions after full event selection are summarised. The statistical error on the background expectation is considered to be negligible, the systematic error on the sum of all background contributions is calculated by varying the amount of expected background events in the final state by 30% or 50%, which is found to have a relative error on the cross section of +5.4/-5.3%, which is the biggest contribution to the total systematic uncertainty.

Background process	Simulation	Uncertainty
Real μe pair		± 30%
Single-top tW	1.87	± 0.56
Diboson VV	0.81	± 0.24
$Z^0/\gamma^\star \to \tau^+\tau^-$	2.48	± 0.74
Fake leptons		± 50%
$t\bar{t}$ other	1.04	± 0.52
$Z^0/\gamma^\star \to e^+e^-$, $\mu^+\mu^-$	0.46	± 0.23
W+Jets	0.50	± 0.25
QCD	0.41	± 0.21
Total background	7.57	± 2.75

Table 7.3: The number of expected background events from simulation after dilepton-plus-jets selection and its associated systematic uncertainties. The contributions of different background processes are exclusively obtained from simulation. An uncertainty of 30% is assigned to the total number of expected events with a real μe pair in the final state [49]. An increased uncertainty of 50% is assigned to the total number of expected events steming from events containing at least one misidentified muon or electron, called *fake* lepton events.

7.1.2 Theoretical Uncertainties

Decay Branching Ratio

For the calculation of the branching ratio of $t\bar{t}$ quark pairs decaying into a muon-electron pair, given in chapter 3, the leading order value of the branching fraction of a W boson into a lepton-neutrino pair ($W \to \ell\nu$) of 1/9 has been used. According to [3], the current world average value for $BR(W \to \ell\nu)$ is $(10.80 \pm 0.09)\%$. Thus, a scaling factor of $(9 \times 0.108)^2$ is applied to the total branching ratio of 3.41%, which results into a corrected value of 3.22%, used for the determination of the inclusive $t\bar{t}$ quark pair production cross section. The relative uncertainty on the cross section measurement is 1.7%.

Signal Uncertainties

The total selection efficiency is estimated using a simulated $t\bar{t}$ quark pair signal sample and is further corrected by applying several scale factors, which are derived from events in data, as described above. In the following, different systematic uncertainties of the parameters, which have been used for the generation of the $t\bar{t}$ quark pair signal sample, such as the **mass of the top quark** or the **energy scale** of the hard scattering process (Q^2 scale), are discussed.

For the generation of the signal sample, a central value of the **top quark mass** of 172.5 GeV/c^2 is used. In order to study the influence of the top quark mass parameter on the event selection efficiency, the top quark mass is varied up and down by 6 GeV/c^2. The result is found to be an increase of the acceptance as a function of mass of 0.8% per GeV/c^2, as described in [17]. The uncertainty of the top quark mass is stated to be 2 GeV/c^2 [3]. Hence, a systematic uncertainty of 1.6% is assigned, due to the uncertainty of the top quark mass. In order to study the influence of the **Q^2 scale** parameter, the value of the energy scale has been increased and decreased by a factor of two. The result of the variation of the Q^2 scale has been found to be significant and a systematic error of 2% is assigned, as proposed in [49]. A systematic uncertainty of 2% is included, as proposed in [49], due to the uncertainty of the **decay model**. Other systematic effects have also been studied, such as the influence of the *shower model* (ISR/FSR), the *jet-parton matching*, which is relevant for the use of the samples generated with MadGraph, as well as the influence of different *PDF*s. The resulting systematic effect of this sources is found to be below 1% and is considered to be negligible.

7.1.3 Luminosity Uncertainty

The measurement of the luminosity has been described in section 4.2.7. The absolute luminosity is determined from the **beam size**, measured by Van Der Meer scans, and the **beam current**. The dominant contribution to the systematic uncertainty of 2.5% arises from fill variations. The total systematic error of the luminosity measurement is 3.6%, given in [28].

Source	Uncertainty [%]
Lepton Selection	±4.0
Jet energy scale	±3.0
Background	+5.4/-5.3
Branching ratio	±1.7
Top quark mass	±1.6
Q^2 scale	±2.0
Decay model	±2.0
Total (w/o lumi.)	±8.2
Luminosity	±3.6
Total (w/o b-tag)	±9.0
B-tagging	±5.0
Total (incl. b-tag)	±10.3

Table 7.4: Summary of all sources of systematic uncertainties.

7.2 Inclusive Cross Section

The inclusive cross section for top quark pair production is measured in a simple *counting experiment*. The number of 59 selected **signal events** N_{Sig} is used for the calculation of the cross section. A robust, cut-based event selection is applied both to data and simulated events. The number of $\mathrm{t\bar{t}}$ quark pair candidate events is derived from the number of selected (observed) events N_{Obs}, by subtracting the expected number of **background events** N_{BKG} after the complete event selection:

$$N_{\mathrm{Sig}} \;=\; N_{\mathrm{Obs}} - N_{\mathrm{BKG}} \;.$$

The contributions from background events have been discussed above. The number of muon-electron signal candidate events is obtained from both a muon-triggered and an electron-triggered data sample, as described in section 6.1. In order to avoid multiple counting of selected events, events selected by one trigger are vetoed in the other data sample. By requiring either the single-muon or single-electron trigger first, the identical number of 59 candidate events is observed. The number of events, observed after full event selection, by using *only* the muon-triggered or electron-triggered data sample is given in table 7.5. The single-muon and single-electron trigger efficiencies have been measured in data to be $\epsilon^{\mu}_{\mathrm{trig}} = 0.88$ and $\epsilon^{\mathrm{e}}_{\mathrm{trig}} = 0.96$, as shown in the previous section. The number of observed events is 52 events from the only muon-triggered data sample and 58 events from the only electron-triggered data sample, which is in good agreement for an expected combined trigger efficiency of approximately 100%. In the following, only the total number of 59 observed signal candidate events is used.

Data sample	Observed events N_{Obs}	Total selection efficiency ϵ_{tot} [%]
μ-only triggered	52	28.2 ± 0.4
e-only triggered	58	29.3 ± 0.4
μe-triggered	59	30.0 ± 0.4

Table 7.5: The number of observed signal candidate events after the dilepton-plus-jets selection requirements, using three different trigger requirements.

After requiring one oppositely-charged, isolated muon-electron pair and two additional high-p_{T} Particle-Flow jets, 59 $\mathrm{t\bar{t}}$ candidate events are observed. The total sum of expected background events from simulation is 7.6 events. The expected contributions of signal and background processes after several sequential selection steps have been summarised in table 6.7. The number of remaining background events is remarkably small, whereas the muon-electron decay channel yields a branching ratio, which is twice as large compared to the dimuon or dielectron decay channel. Figures 6.9 (a) and (b) show the jet multiplicity for PF jets and the distribution of missing transverse energy after the dilepton-plus-jets selection. The $\mathrm{t\bar{t}}$ cross section is calculated using equation 7.8:

$$\sigma_{\mathrm{t\bar{t}}} \;=\; \frac{N_{\mathrm{Sig}}}{\mathcal{A}\cdot\epsilon_{\mathrm{sel}}\cdot\mathrm{BR}\cdot\mathrm{L}} \;=\; \frac{N_{\mathrm{Obs}} - N_{\mathrm{BKG}}}{\mathcal{C}\cdot\epsilon_{\mathrm{tot}}\cdot\mathrm{BR}\cdot\mathrm{L}}\;, \tag{7.8}$$

where $\epsilon_{\mathrm{tot}} = \mathcal{A}\times\epsilon_{\mathrm{sel}}$ is the total selection efficiency for signal events, which is derived from simulation, BR is the branching ratio of $\mathrm{t\bar{t}}$ into a muon-electron pair and L is the

total integrated luminosity. Several correction (scaling) factors $\mathcal{C}$ are applied to the total selection efficiency ϵ_{tot} of the muon, electron and jet selection, which is estimated from simulated $\text{t}\bar{\text{t}}$ events to be 30.0%. A combined scale factor without trigger correction per muon-electron event of $\mathcal{C}^{\mu\text{e}}_{\text{reco/id/iso}} = 0.953$ is taken from [49], which accounts for the efficiencies for lepton reconstruction, identification, isolation and charge-misidentification, as described in the previous section. Combining this scale factor with the trigger efficiency correction factor of $\mathcal{C}^{\mu\text{e}}_{\text{trig}} = 0.994$, an combined correction factor of $\mathcal{C}^{\mu\text{e}}_{\text{tot}} = 0.947$ is applied to the total selection efficiency.
According to the branching ratio of W $(\rightarrow \tau) \rightarrow (\mu, e)$ taken from the particle data group [3], the total branching ratio of $\text{t}\bar{\text{t}}$ pairs into the muon-electron final state, including decays via τ leptons, is calculated to be BR = 3.22%, as described in section 7.1. As described in section 6.1, the total integrated luminosity of the used data sample is 35.9 pb^{-1}, which has a relative uncertainty of 4%. Based on 59 observed signal events, using an estiomation of N_{BKG} from simulated events, the **cross section of top quark pair production** in proton-proton collions at $\sqrt{s} = 7$ TeV is calculated as:

$$\boxed{\sigma_{\text{t}\bar{\text{t}}} \quad = \quad (\ 156.2 \ \pm\ 24.8\,(\text{stat.}) \ \pm\ 14.1\,(\text{sys.})\ )\ \text{pb}}$$

The signal-over-background ratio is $N_{\text{Sig}}/N_{\text{BKG}} = 6.8$, which corresponds to a **significance** of the top signal of:

$$\frac{N_{\text{Sig}}}{\sqrt{N_{\text{Sig}} + N_{\text{BKG}}}} \ = \ 6.7\ .$$

7.3 Kinematic Event Reconstruction

In order to validate the **top-like topology** of the selected events, a kinematic reconstruction algorithm is applied to the final event sample, both from data and simulation. The method is similar but not identical to the MWT method, which is used at the DØ experiment in order to measure the top quark mass in the dilepton final state, described in [50] and [51], and additionally provides a discriminator for the determination of the **top quark mass** from dileptonically decaying top quark pairs. A detailed description of the kinematic event reconstruction is given in [52] and [53]. A brief summary of the kinematic reconstruction method is given below.
It is studied, if this method can be used in order to further **reject background events** and improve the significance of the signal sample. Since it is expected, that the kinematic event reconstruction provides a solution for $\text{t}\bar{\text{t}}$ signal events with a higher probability, it is compared to the results found by the application of b-tagging to the final event sample, described in section 7.4. The cross section is calculated for the number of events, in which a kinematic solution is found and the distribution for the resulting top quark mass is shown in figure 7.3. The resulting top quark mass may not be used for an independent measurement, because the weighting of the event solutions depends on a predefined top quark mass, which was used for the generation of a neutrino reference spectrum, shown in figure 7.2 [53].

Method of Kinematic Reconstruction

The dileptonic top quark pair decay comprises two neutrinos, one neutrino ν_ℓ and one anti-neutrino $\bar{\nu}_{\ell'}$, from the leptonic decays of the W bosons. They cannot be fully reconstructed, because they leave the detector without interaction. Thus, only the transverse projection of the *sum of their momenta* is measured as $\not\!\!E_{\mathrm{T}}$. The **missing transverse energy** is measured in x- and y-components, given in equation 7.9:

$$\not\!\!E_{\mathrm{T}} \;=\; \begin{pmatrix} \not\!\!E_x \\ \not\!\!E_y \end{pmatrix} = \begin{pmatrix} p_x(\nu) + p_x(\bar{\nu}) \\ p_y(\nu) + p_y(\bar{\nu}) \end{pmatrix} . \tag{7.9}$$

Furthermore, for massless neutrinos, the **energy of the neutrino and anti-neutrino** are related as follows to their momenta:

$$E^2(\nu) \;=\; p_x^2(\nu) \;+\; p_y^2(\nu) \;+\; p_z^2(\nu) \;, \tag{7.10}$$

$$E^2(\bar{\nu}) \;=\; p_x^2(\bar{\nu}) \;+\; p_y^2(\bar{\nu}) \;+\; p_z^2(\bar{\nu}) \;. \tag{7.11}$$

The masses of the **top and anti-top quarks**, as well as the two **W bosons**, are related to the energies and momenta of their decay products, as described in the following equations:

$$\begin{aligned} m^2(\mathrm{t}) \;=\; & \;\;(\;\; E(\mathrm{b}) + E(\bar{\ell}) + E(\nu) \;\;)^2 \\ - & \;\;(\;\; p_x(\mathrm{b}) + p_x(\bar{\ell}) + p_x(\nu) \;\;)^2 \\ - & \;\;(\;\; p_y(\mathrm{b}) + p_y(\bar{\ell}) + p_y(\nu) \;\;)^2 \\ - & \;\;(\;\; p_z(\mathrm{b}) + p_z(\bar{\ell}) + p_z(\nu) \;\;)^2 \end{aligned} \tag{7.12}$$

$$\begin{aligned} m^2(\bar{\mathrm{t}}) \;=\; & \;\;(\;\; E(\bar{\mathrm{b}}) + E(\ell) + E(\bar{\nu}) \;\;)^2 \\ - & \;\;(\;\; p_x(\bar{\mathrm{b}}) + p_x(\ell) + p_x(\bar{\nu}) \;\;)^2 \\ - & \;\;(\;\; p_y(\bar{\mathrm{b}}) + p_y(\ell) + p_y(\bar{\nu}) \;\;)^2 \\ - & \;\;(\;\; p_z(\bar{\mathrm{b}}) + p_z(\ell) + p_z(\bar{\nu}) \;\;)^2 \end{aligned} \tag{7.13}$$

$$\begin{aligned} m^2(\mathrm{W}^+) \;=\; & \;\;(\;\; E(\bar{\ell}) + E(\nu) \;\;)^2 \\ - & \;\;(\;\; p_x(\bar{\ell}) + p_x(\nu) \;\;)^2 \\ - & \;\;(\;\; p_y(\bar{\ell}) + p_y(\nu) \;\;)^2 \\ - & \;\;(\;\; p_z(\bar{\ell}) + p_z(\nu) \;\;)^2 \end{aligned} \tag{7.14}$$

$$\begin{aligned} m^2(\mathrm{W}^-) \;=\; & \;\;(\;\; E(\ell) + E(\bar{\nu}) \;\;)^2 \\ - & \;\;(\;\; p_x(\ell) + p_x(\bar{\nu}) \;\;)^2 \\ - & \;\;(\;\; p_y(\ell) + p_y(\bar{\nu}) \;\;)^2 \\ - & \;\;(\;\; p_z(\ell) + p_z(\bar{\nu}) \;\;)^2 \end{aligned} \tag{7.15}$$

The four-momenta of the two leptons $p(\ell)$, $p(\bar{\ell})$ and the to b jets $p(\mathrm{b})$, $p(\bar{\mathrm{b}})$ are measured in the experiment. The system of six non-linear equations 7.10 to 7.15 contains six individual neutrino momenta, which are unknown. Therefore, the system of equations is *underconstrained* and cannot be solved analytically nor by a kinematic fitting procedure, as described in [52].

By measuring the missing transverse energy $\not\!\!E_{\mathrm{T}}$ in the event, two constraints on the system of kinematics are given by the x- and y-components of $\not\!\!E_{\mathrm{T}}$ in equation 7.9: $\not\!\!E_x$ and $\not\!\!E_y$. Furthermore, the mass of the W bosons $m(\mathrm{W}^+)$ and $m(\mathrm{W}^-)$ has been measured to very high precision and can also be used as two additional constraints to equations 7.14 and 7.15. Finally, the masses of the top and the anti-top quark are expected to be identical. Therefore, a fifth constraint is introduced by equalising equations 7.12 and 7.13: $m(\mathrm{t}) = m(\bar{\mathrm{t}})$. In summary, the presence of six unknown neutrino momenta is reduced by introducing **five boundary conditions**:

$$\not{E}_x,\, \not{E}_y,\, m(\mathrm{W}^+),\, m(\mathrm{W}^-) \text{ and } m(\mathrm{t}) = m(\bar{\mathrm{t}}) \ .$$

However, the system of kinematic equations still remains underconstrained. The set of equations can further be transformed into a single equation, which corresponds to the roots of a polynomial of fourth order $\mathcal{P}_4(m_\mathrm{t}, p_x^i(\bar{\nu}))$, with $i = 0,1,2,3,4$, in one component of the neutrino momenta and the mass of the top quark as a free parameter. Without loss of generality, the x-component of the anti-neutrino $p_x(\bar{\nu})$ has been chosen [53]. The resulting equation $\mathcal{P}_4(m_\mathrm{t}, p_x^i(\bar{\nu})) = 0$ is solvable for a **fixed top quark mass** with a fourfold ambiguity. The value of the top quark mass parameter, which is used as input for resolving $\mathcal{P}_4 = 0$, is varied between 100 and 300 GeV/c^2 in 1 GeV/c^2 steps. For a polynomial equation of fourth order, up to four solutions may exist for each event. By varying the input top quark mass parameter 200 times, maximally 800 solution for the kinematics of each $\mathrm{t}\bar{\mathrm{t}}$ event can be found and therefore, up to 800 reconstructed top quark mass values for each event. Hence, in order to select the most probable value of the top quark mass, a **weighting factor** is assigned to each kinematic event solution and only the solution corresponding to the highest weight is selected.

Up to this point, the method of finding solutions to the kinematic equations, described in this analysis, is similar to MWT method, which is described in [50] and [51]. The main difference consists in the **weighting of the solutions**, which are found by the kinematic event reconstruction method. As described above, up to 800 solution for the kinematics of the event may exist. In order to determine the most probable solution, a comparison of the reconstructed neutrino energies $E(\nu)$ and $E(\bar{\nu})$ to a **generated neutrino energy spectrum** is performed. Therefore, 10 000 dileptonically decaying $\mathrm{t}\bar{\mathrm{t}}$ events have been generated using Pythia. The two-dimensional distribution of the generated neutrino and anti-neutrino energy is shown in figure 7.2. The weighting factor is derived by comparing the reconstructed neutrino energies to the generated reference spectrum. Hence, the best agreement to the simulated neutrino energy spectrum gives the highest weight and the corresponding solution of the kinematic equations is chosen.

Top Quark Properties and Cross Section Measurement

The properties of the reconstructed top quark, such as the reconstructed top quark mass and the transverse momentum of the top quark, and those of the top-antitop quark pair, such as the invariant mass of the $\mathrm{t}\bar{\mathrm{t}}$ quark pair and the p_T of the top-antitop quark pair, are shown in figure 7.3. Figure 7.3 (a) shows, that the reconstructed top quark mass has a clearly visible maximum around the nominal value of the top quark mass, as expected from simulation. It should be noted, that events, for which **no solution** is found by the kinematic event reconstruction, are entered in the first bin at a **negative value** of the top quark mass. Therefore, the figure 7.3 (a) shows, that the probability for rejecting background events is *not* significantly higher than for signal events. Nevertheless, an

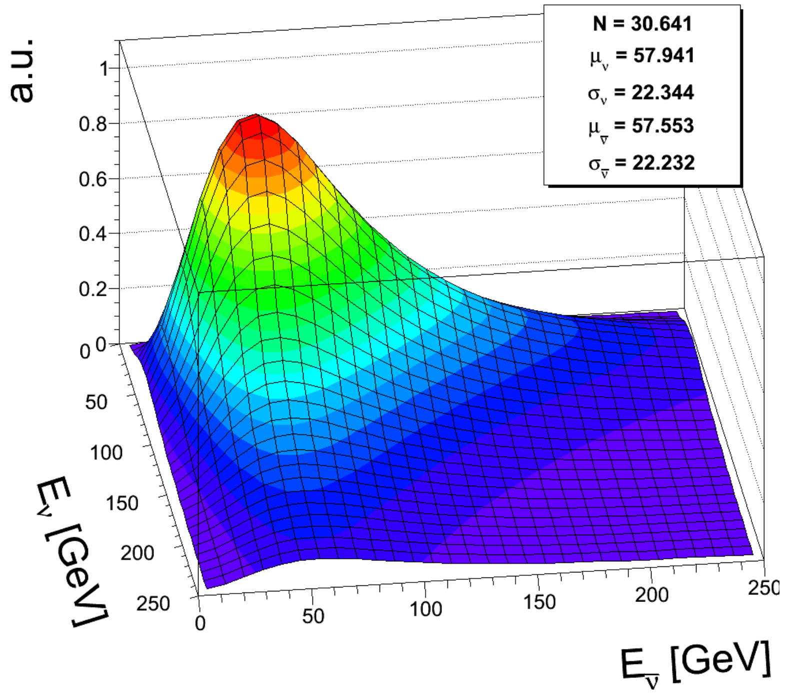
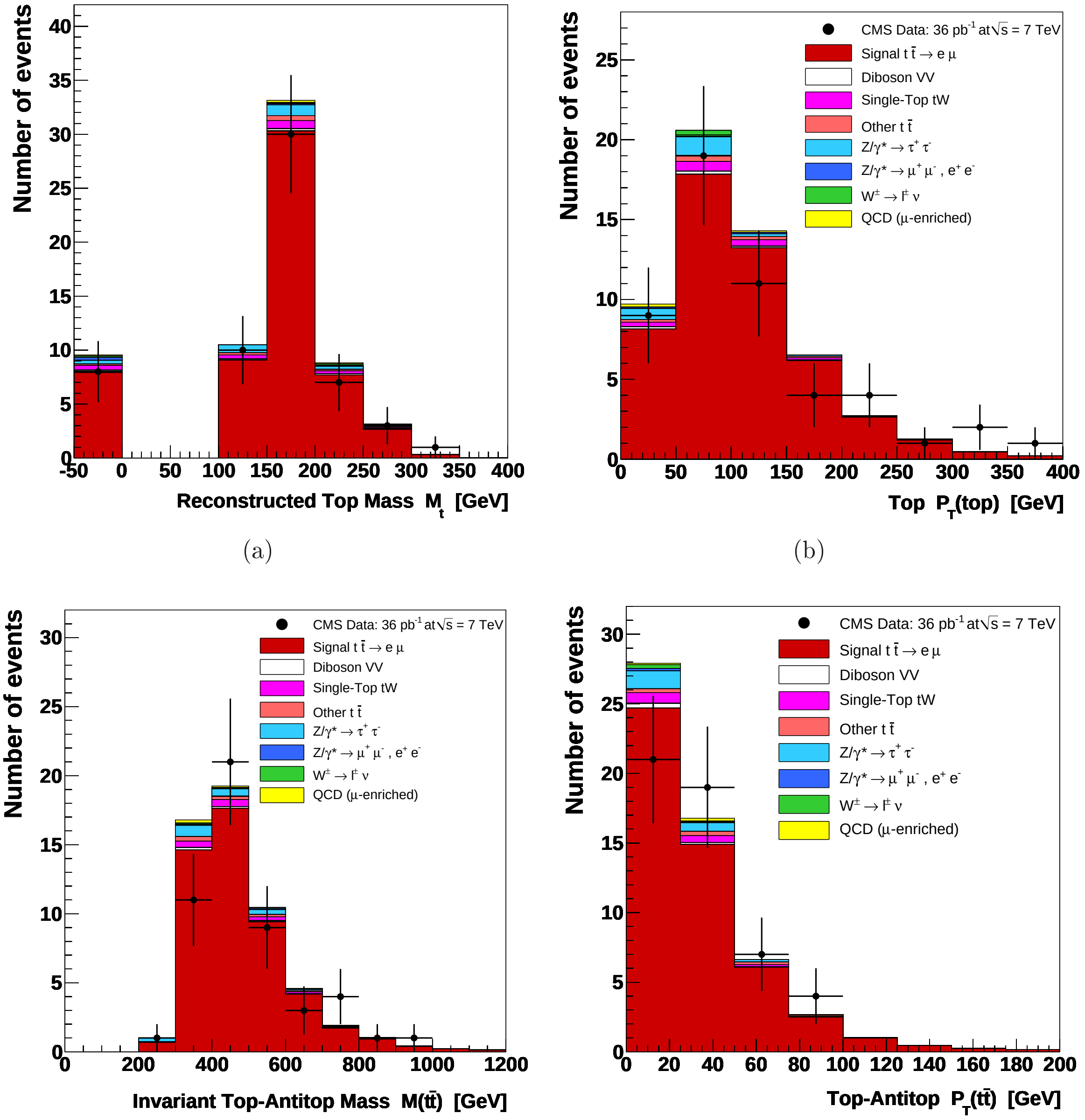


Figure 7.2: Two-dimensional distribution of the generated neutrino and anti-neutrino energy, used for the kinematic event reconstruction [53].

excellent agreement between the reconstructed events in data and those obtained from simulation is found in the distributions for (a) the single-top quark and (c) the invariant $t\bar{t}$ quark pair masses. There is also good agreement between the data and simulated events for the distributions of (b) the transverse momentum of the single-top quark and (d) the p_T of the $t\bar{t}$ quark pair system, which are compatible within the statistical uncertainties. Based on 59 $t\bar{t}$ quark pair candidate events observed in data, a solution for the kinematic event reconstruction is found for 51 of them. The number of background events, which is obtained from simulation, for which no solution is found, is 1.6 events. The significance of the signal decreases from 6.7 to 6.3. The total selection efficiency on simulated $t\bar{t}$ signal events, including the kinematic event reconstruction, is obtained from simulation to be 25.9%. Based on 51 events in data and 5.9 remaining, simulated background events, the $t\bar{t}$ cross section, using only events, which are successfully reconstructed, is calculated to be:

$$\boxed{\sigma_{t\bar{t}}^{\mathrm{kin}} \quad = \quad (\ 158.6 \pm 26.6(\mathrm{stat.}) \pm 14.3(\mathrm{sys.})\ )\ \ \mathrm{pb}}$$



(a) (b)

(c) (d)

Figure 7.3: Properties of the reconstructed top quark and top-antitop quark pair, found by the kinematic event reconstruction method applied to data and simulated events after dilepton-plus-jets selection. It should be noted, that events, for which **no solution** is found by the kinematic event reconstruction, are entered in the first bin at a **negative value** of the top quark mass. (a) Distribution of the single-top quark mass and (b) the transverse momentum p_T of the single-top quark. (c) Top-antitop quark invariant mass distribution and (d) transverse momentum of the $t\bar{t}$ system. Simulated samples are normalised to an integrated luminosity of 35.9 pb^{-1}. Events, for which no kinematic event reconstruction is found, are entered with a negative value.

7.4 Performance of B-Tagging

The b-tagging efficiency is determined using the $t\bar{t}$-enriched data sample after requiring at least two high-p_{T} jets. The number of events after dilepton-plus-jets selection, which has at least one b-tagged jet, is compared to the total number of selected events, using two different b-tagging discriminators, provided by the *track counting* and the *simple secondary vertex* algorithms. The **b-tagging efficiency** is calculated using the following equation:

$$\epsilon_{\text{b-tag}} \;=\; \frac{N^{\text{Sig}}_{\text{b-tag}}}{N^{\text{Sig}}_{2\,\text{jets}}} \; . \qquad (7.16)$$

Track Counting High Efficiency

The event yields, for events with at least one TCHE b-tagged jet at the *loose* working point, are summarised in table 7.6. Based on 51 events with at least one b-tag observed in data and an expected number of 3.7 remaining background events, the **b-tagging efficiency** is measured to be:

$$\epsilon^{\text{TCHE}}_{\text{b-tag}} \;=\; (\,92.02 \pm 3.78\,)\;\% \; .$$

From simulated signal events, a b-tagging efficiency of (92.48 ± 0.23) % is obtained, which is in excellent agreement with the value obtained from data events. The $\mathrm{t\bar{t}}$ quark pair cross section is recalculated for events with at least one b-tagged jet. The multiplicity of b jets found by the TCHE algorithm at the *loose* working point is shown in figure 6.10 (d). Compared to the potential background rejection of the kinematic event reconstruction, described in section 7.3, most of the remaining background events are rejected by requiring one or more TCHE b-tagged jets. The **signal significance** is found to be 6.6, which is slightly lower, than the significance obtained after the dilepton-plus-jets selection, which is 6.7. Nevertheless, an optimal signal-over-background ratio of $N_{\text{Sig}}/N_{\text{BKG}} = 12.2$ is derived. Based on 51 observed $\mathrm{t\bar{t}}$ quark pair candidate events and a total selection efficiency of 27.76%, derived from simulation, the measured cross section for $\mathrm{t\bar{t}}$ quark pair events with at least one TCHE b-tagged jet is found to be:

$$\boxed{\sigma^{\text{TCHE}}_{\mathrm{t\bar{t}}} \quad = \quad (\;155.2 \;\pm\; 24.3\,(\text{stat.}) \;\pm\; 16.0\,(\text{sys.})\;)\;\; \text{pb}}$$

This value is compatible with the values calculated after the dilepton-plus-jets selection in section 7.2 and the kinematic event reconstruction 7.3.

Simple Secondary Vertex High Efficiency

In table 7.7, the event yields for events with at least one SSVHE b-tagged jet at the *medium* working point, are summarised. Based on 35 events with at least one b-tag observed in data and an expected number of 2.6 remaining background events, the **b-tagging efficiency** is measured to be:

$$\epsilon^{\text{SSVHE}}_{\text{b-tag}} = (\,63.04 \pm 6.73\,)\;\% \; .$$

TCHE *loose*	≥ 2 jets	0 b-tags	≥ 1 b-tag
$t\bar{t}$ signal (μe)	57.9	4.3	53.6
Total background	7.6	3.9	3.7
Total simulation	65.5	8.2	57.3
Total events in data	59	8	51
Signal events in data	**51.4**	4.1	**47.3**

Table 7.6: Event yields in data and simulation after dilepton-plus-jets selection and requiring at least one TCHE *loose* b-tagged PF jet.

From simulated signal events, a b-tagging efficiency of $(\,78.62 \pm 0.35\,)\,\%$ is obtained, which shows a significant deviation from the efficiency measured from data events. Hence, in order to correct the b-tagging efficiency in a data-driven way, a correction factor of $\mathcal{C}_{\mathrm{SSVHE}} = \frac{0.630}{0.786} \approx 0.8$ is applied to the total selection efficiency, which is derived from simulation.
As for the TCHE b-tagging algorithm, the $t\bar{t}$ quark pair cross section is calculated for events with at least one b-tagged jet. The multiplicity of b jets found by the SSVHE algorithm at *medium* working point is shown in figure 6.10 (d), which already shows a significantly lower tagging efficiency in data compared to simulated events. As for the TCHE algorithm, nearly all of the remaining background events are rejected by the requirement of one or more SSVHE b-tagged jet. The **signal significance** is 5.5, which is also lower than the significance obtained after the dilepton-plus-jets selection, which is 6.7. A very pure signal-over-background ratio of $N_{\mathrm{Sig}}/N_{\mathrm{BKG}} = 12.7$ is derived, which is slightly higher, than for TCHE tagged events. Based on 35 observed $t\bar{t}$ quark pair candidate events and a *corrected* total selection efficiency of $(\,\mathcal{C}_{\mathrm{SSVHE}} \times 23.60\%\,) = 18.92\%$, the measured cross section for $t\bar{t}$ quark pairevents with at least one SSVHE b-tagged jet is found to be:

$$\boxed{\sigma_{t\bar{t}}^{\mathrm{SSVHE}} \quad = \quad (\ 156.4\ \pm\ 29.5(\mathrm{stat.})\ \pm\ 16.1(\mathrm{sys.})\ )\ \mathrm{pb}}$$

which is also found to be compatible with the values calculated after the dilepton-plus-jets selection in section 7.2 and the kinematic event reconstruction 7.3, as well as the value obtained from TCHE b-tagged events.

SSVHE *medium*	≥ 2 jets	0 b-tags	≥ 1 b-tag
$t\bar{t}$ signal (μe)	57.9	12.3	45.6
Total background	7.6	5.0	2.6
Total simulation	65.5	17.3	48.2
Total events in data	59	24	35
Signal events in data	**51.4**	19.0	**32.4**

Table 7.7: Event yields in data and simulation after dilepton-plus-jets selection and requiring at least one SSVHE *medium* b-tagged PF jet.

7.5 Comparison to other Results

As introduced in section 3.2, the measurement of the inclusive top quark pair cross section, has been performed also by the CMS and ATLAS collaborations, based on the data taking period in 2010 with an integrated luminosity of about 36 pb^{-1}. The measured values presented above are compared to the measurement performed by the top quark working group of the CMS collaboration, which is described in [19]. Figure 7.4 shows a summary of the measurements in all dileptonic and semileptonic decay channels, with and without the use of b-tagging with an integrated luminosity of about 3 and 36 pb^{-1}. The combined result, which is obtained in all three dileptonic decay channels (ee, μe and $\mu\mu$) without the requirement of a b-tagged jet is $\sigma_{\mathrm{t\bar{t}}}^{\mathrm{dilep.}} = (168 \pm 24)$ pb. This value is compatible within less than one standard deviation with all measurements performed in this thesis in the muon-electron decay channel. The *arithmetic average* of the different top quark pair cross section values, obtained after the standard dilepton-plus-jets selection, kinematic event reconstruction, as well as after the requirement of two different b jet identification algorithms, is $\langle\sigma_{\mathrm{t\bar{t}}}\rangle = (157 \pm 30)$ pb.

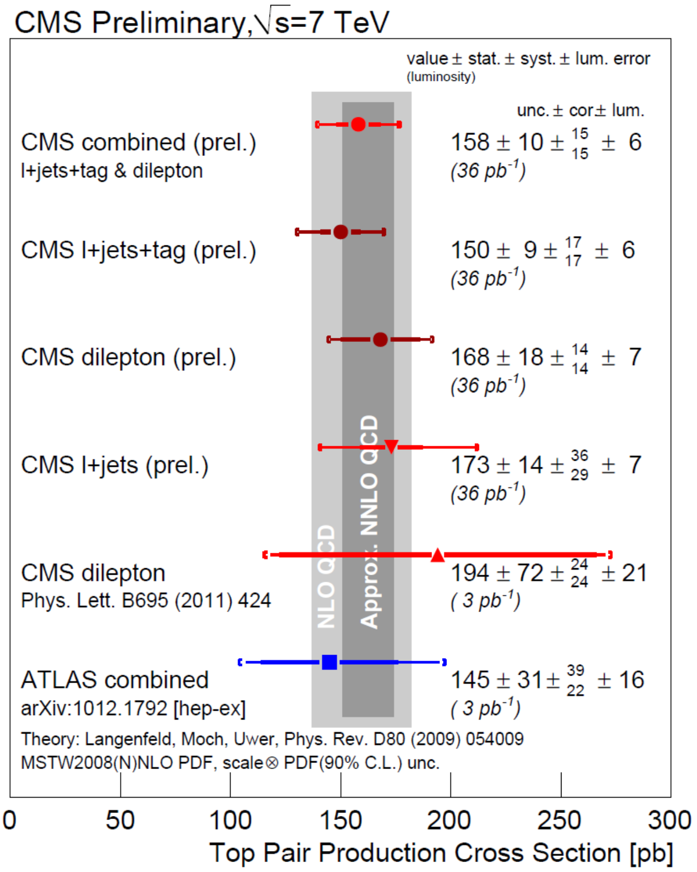


Figure 7.4: The measurement of the inclusive top quark pair cross section, performed by the CMS and ATLAS collaborations, based on the data taking period in 2010 with an integrated luminosity of about 3 and 36 pb^{-1} in all dileptonic and semileptonic decay channels [19].

7.6 Differential Cross Sections

In this section, the first results from the measurement of differential cross sections based on the data from the complete 2010 data taking periods are presented. For the first time in CMS, the cross section of the production of top quark pairs is measured differentially as a function of the kinematic observables of the final state objects, such as the transverse momentum p_{T} of the two leptons and the invariant mass of the lepton pair. Based on the solution of the kinematic event reconstruction presented above, the cross section is also calculated differentially as a function of the kinematic properties of the reconstructed top-antitop quark pair. The differential $\mathrm{t\bar{t}}$ quark pair production cross section as a function of the observable X is calculated using equation 7.17:

$$\frac{d\sigma_{\mathrm{t\bar{t}}}}{dX} = \frac{\Delta N_{\mathrm{Obs}} - \Delta N_{\mathrm{BKG}}}{\Delta\epsilon_{\mathrm{tot}} \cdot \Delta X \cdot \mathrm{BR} \cdot \mathrm{L}} , \tag{7.17}$$

where $\Delta\epsilon_{\mathrm{tot}}$ is the total selection efficiency for signal events, which is derived from simulation, calculated individually in each bin of the observable X. The number of signal events is also calculated for each bin of the observable X, where ΔX is the width of each individual bin and ΔN_{Obs} and ΔN_{BKG} are the number of observed events in the data and simulated background events counted in each bin. As already introduced, BR is the branching ratio of a $\mathrm{t\bar{t}}$ quark pair decaying into a muon-electron pair and L is the total integrated luminosity.

Migration Effects: Purity and Stability

In order to measure the $\mathrm{t\bar{t}}$ quark pair cross section differentially, a binning ΔX is imposed for each observable X. Therefore, bin-to-bin migration effects are studied, which are formulated in terms of **purity** and **stability** in each bin. Purity and stability are (similarly to the selection efficiency) determined from simulation, by comparing the distributions of generator-level objects (generated particles) to those of the corresponding reconstructed object. Therefore, a geometrical matching of $\Delta R < 0.1$ in $(\eta - \phi)$ space is performed between generated and reconstructed objects. In figures 7.5 - 7.8 (c), the two-dimensional distributions of each kinematic variable are shown. Based on the correlation distributions, the purity and stability are calculated for each bin i, as defined in the following equations:

$$\mathrm{Purity}(i) = \frac{\text{Number of events generated AND reconstructed in bin } i}{\text{Number of ALL events reconstructed in bin } i} , \tag{7.18}$$

$$\mathrm{Stability}(i) = \frac{\text{Number of events generated AND reconstructed in bin } i}{\text{Number of ALL reconstructed events generated in bin } i} . \tag{7.19}$$

Hence, the **purity** corresponds to the number of events on the diagonal, divided by the *integral* of all events in the corresponding horizontal bin of reconstructed objects. Similarly, the **stability** is defined as the number of events on the diagonal, divided by the *integral* of all events in the corresponding vertical bin of generated objects, which have been reconstructed in *any* bin.

Lepton Transverse Momentum

In case of the lepton transverse momenta, the correlation between reconstructed and generated leptons is clearly visible, hence, a high purity and stability of greater than 80% is found for most of the bins in muon and electron p_{T} , as it is shown in figures 7.5 and 7.6 (d). Also the selection efficiency shows a constant distribution in all bins except for the first bin, where a lower value is obtained due to the lepton selection requirement of $p_{\mathrm{T}} > 20$ GeV/c. The purity and stability is calculated from the correlation between generated (GEN) and reconstructed (RECO) lepton momenta, shown in figures 7.5 and 7.6 (c). The corresponding p_{T} -distributions after applying the dilepton-plus-jets selection requirements are shown in figures 7.5 and 7.6 (a). A good agreement between events from data and simulation is obtained over the full range of the lepton momenta. Using equation 7.17, the corresponing **differential $\mathrm{t\bar{t}}$ quark pair cross sections** as a function of the lepton transverse momenta $\frac{d\sigma_{\mathrm{t\bar{t}}}}{dp_{\mathrm{T}}(\mu)}$ and $\frac{d\sigma_{\mathrm{t\bar{t}}}}{dp_{\mathrm{T}}(\mathrm{e})}$ are shown in figures 7.5 and 7.6 (b). The distribution calculated from 2010 data events (dots) is compared to the cross section obtained from the simulated signal sample (line), which is normalised to the NLO $\mathrm{t\bar{t}}$ quark pair cross section at $\sqrt{s} = 7$ TeV of 157.5 pb. An almost perfect agreement is found between the measured cross section from data event and simulation in each bin of the muon and electron momenta within the statistical uncertainties. Only the *statistical uncertainties* are taken into account in figures 7.5 and 7.6 (b), as well as for all distributions of the differential cross section in the following.

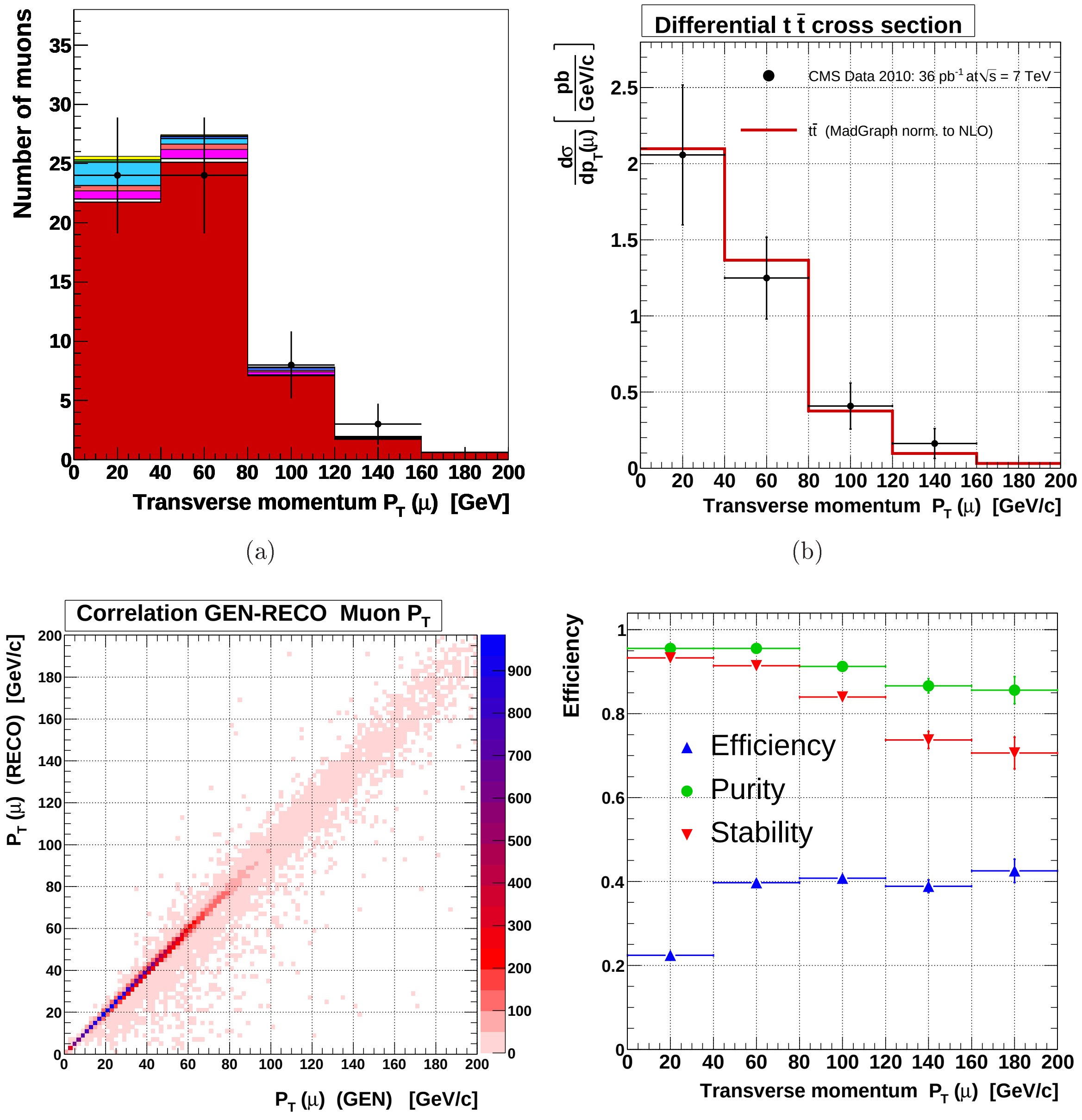
Dilepton Invariant Mass

The differential $\mathrm{t\bar{t}}$ quark pair cross section as a function of the dilepton invariant mass $\frac{d\sigma_{\mathrm{t\bar{t}}}}{dM_{\mu\mathrm{e}}}$ is shown in figure 7.7 (b). Also for the dilepton mass, an overall good agreement is found within the statistical uncertainties in each bin. In case of the invariant muon-electron mass, the reconstructed and generated dilepton masses are reasonably correlated, which is shown in figure 7.7 (c). As in the case of the lepton momenta, a purity and stability of above 80% is found in most of the bins in $M_{\mu\mathrm{e}}$, as it is shown in figure 7.7 (d). The selection efficiency increases significantly for higher dilepton masses. The corresponding distribution of the dilepton mass after applying the dilepton-plus-jets selection requirements is shown in figure 7.7 (a), which shows a good agreement between events from data and simulated events.

Top Quark Properties

The method for the kinematic event reconstruction, described above, which provides a solution for the four-momenta of the top-antitop quark pair, is used to study the kinematic properties of the single top quark and the top-antitop quark pair. The value of the top quark mass, which is used as input for solving the kinematic equations, is varied between 100 and 300 GeV/c^2 in 1 GeV/c^2 steps. In particular, the $\mathrm{t\bar{t}}$ quark pair invariant mass distribution $M_{\mathrm{t\bar{t}}}$ and the transverse momentum of the reconstructed top quark p_{T} (top) are analysed. First, the distribution of the reconstructed **$\mathrm{t\bar{t}}$ quark pair invariant mass** after the dilepton-plus-jets selection is shown in figure 7.8 (a). The comparison between selected events from data and simulated events yields good agreement within the statistical

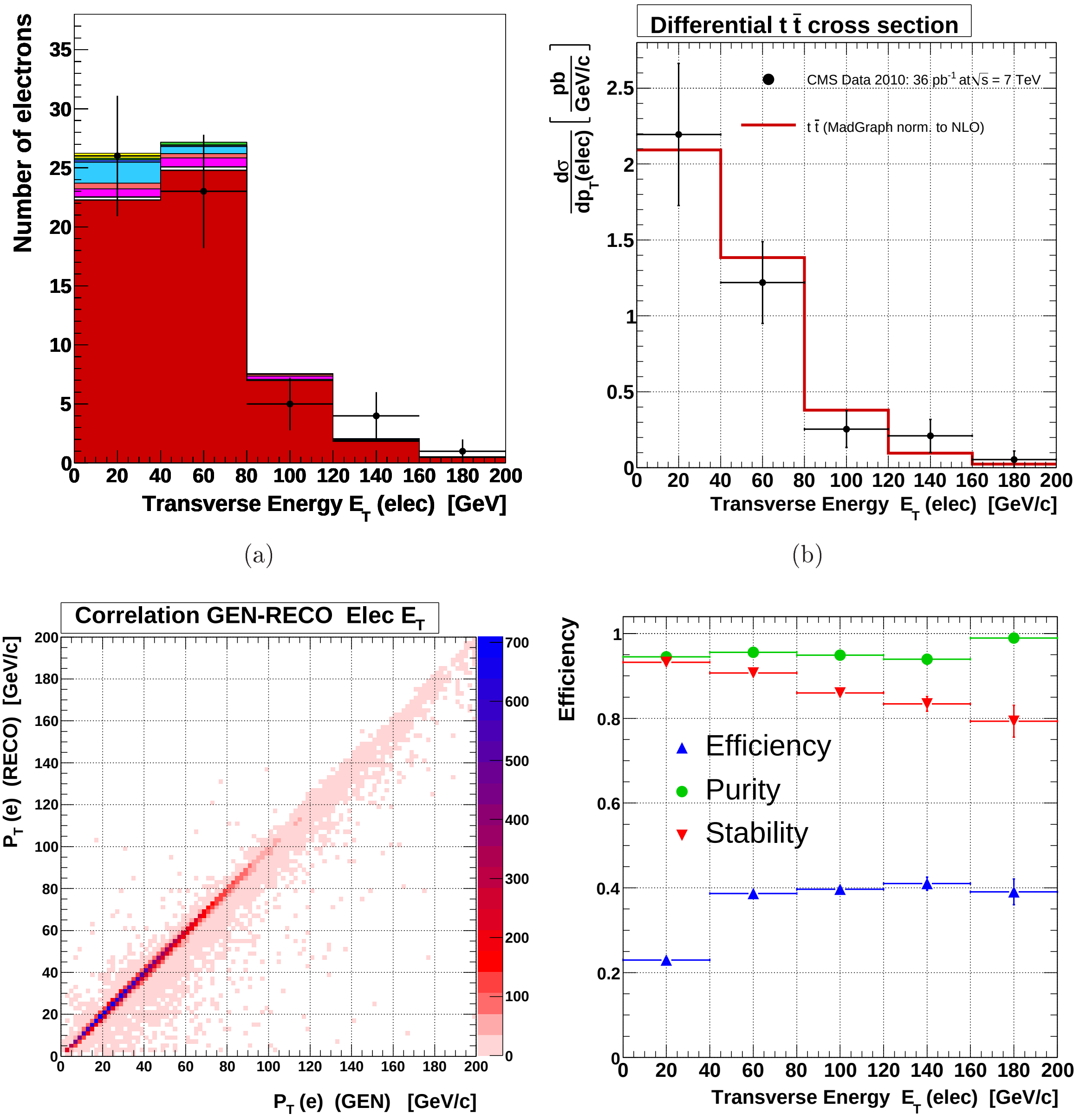
uncertainties. The distribution of the correlation between the reconstructed and generated $\mathrm{t\bar{t}}$ quark pair masses, which are shown in 7.8 (c), show only a low correlation, which results into a relatively low and also varying distribution of the purity and stability in the range of $M_{\mathrm{t\bar{t}}}$, which is shown in figure 7.8 (d). Nevertheless, the total selection efficiency of $\mathrm{t\bar{t}}$ quark pairs shows a constant behaviour over the full range of reconstructed masses. In order to improve the purity and stability, a wider binning should be chosen. The differential $\mathrm{t\bar{t}}$ quark pair cross section as a function of the reconstructed top-antitop quark invariant mass $\frac{d\sigma_{\mathrm{t\bar{t}}}}{dM_{\mathrm{t\bar{t}}}}$ is shown in figure 7.8 (b). A good agreement is found within the statistical uncertainties in each bin of $M_{\mathrm{t\bar{t}}}$.
Secondly, the distribution of the reconstructed **single top quark transverse momentum** is analysed. Figure 7.9 (a) shows the distribution of the reconstructed p_{T} (top), which shows a statistical fluctuation of about 2σ between $100 < p_{\mathrm{T}}(\mathrm{top}) < 200$ GeV/c. Nevertheless, the overall distribution of the reconstructed top quark p_{T} shows reasonable agreement compared to the simulation within the statistical uncertainties. The correlation between the generated top quark momentum and the reconstructed p_{T} (top), which is shown in figure 7.9 (c), has to be studied in further detail in order to provide high purities and stabilities over the full range of p_{T} (top). For low transverse momenta of the reconstructed top quark, a purity of about 70% is obtained, which is shown in figure 7.9 (d). The total selection efficiency slighty increases towards higher values of p_{T} (top). The differential $\mathrm{t\bar{t}}$ quark pair cross section as a function of the reconstructed single top quark transverse momentum $\frac{d\sigma_{\mathrm{t\bar{t}}}}{dp_{\mathrm{T}}(\mathrm{top})}$ is shown in figure 7.8 (b), which shows reasonable agreement between data and simulated events within statistical uncertainties.



(a) (b)

(c) (d)

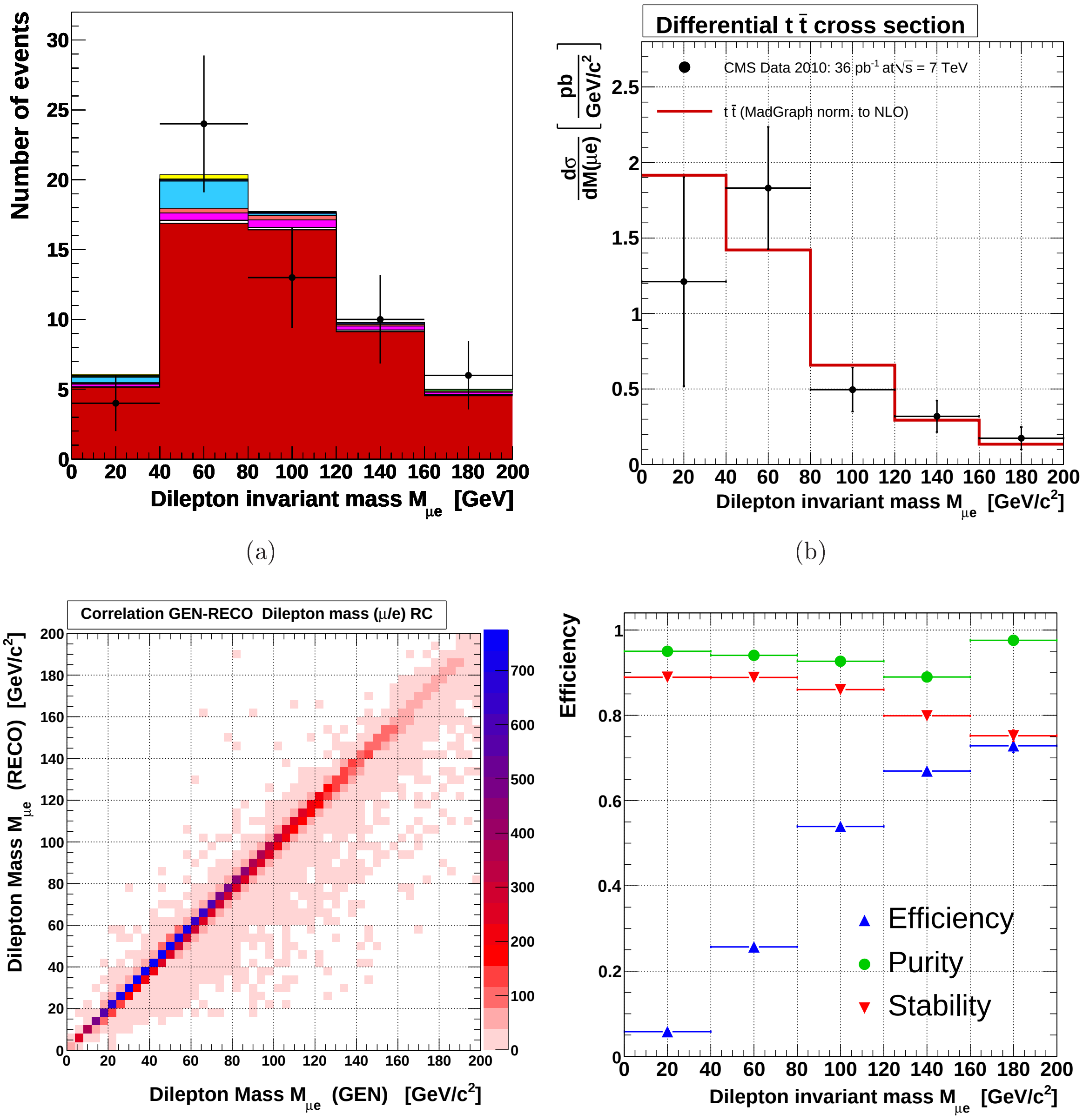
Figure 7.5: (a) **Muon transverse momentum** p_{T} (μ) after dilepton-plus-jets selection. Simulated samples are normalised to an integrated luminosity of 35.9 pb^{-1}. (b) Differential $\mathrm{t}\bar{\mathrm{t}}$ quark pair cross section as a function of the muon transverse momentum $\frac{d\sigma_{\mathrm{t}\bar{\mathrm{t}}}}{dp_{\mathrm{T}}(\mu)}$. The distribution calculated from 2010 data events (dots) is compared to the cross section obtained from the simulated signal sample (line), which is normalised to the NLO $\mathrm{t}\bar{\mathrm{t}}$ quark pair cross section at $\sqrt{s} = 7$ TeV of 157.5 pb. (c) Distribution of the correlation between generated muon p_{T} (GEN) and reconstructed p_{T} (RECO) and (d) the corresponding purities and stabilities for each individual bin, derived from (c). The total selection efficiency, used for the calculation of the differential cross section in each bin of p_{T} (μ), is shown in (d).



(a) (b)

(c) (d)

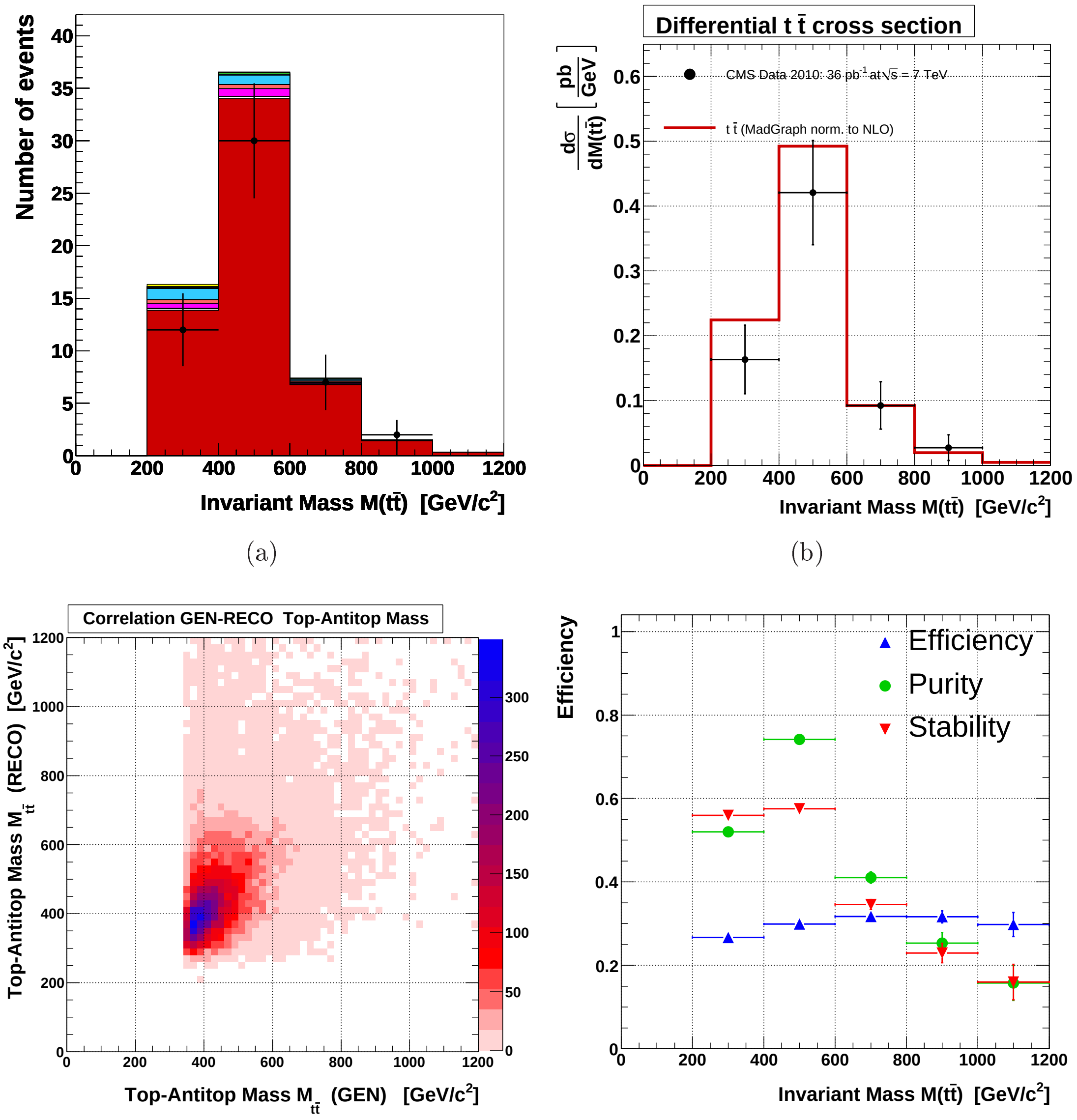
Figure 7.6: (a) **Electron transverse momentum** $p_{\mathrm{T}}\,(e)$ after dilepton-plus-jets selection. Simulated samples are normalised to an integrated luminosity of 35.9 pb^{-1}. (b) Differential $\mathrm{t\bar{t}}$ quark pair cross section as a function of the electron transverse momentum $\frac{d\sigma_{\mathrm{t\bar{t}}}}{dp_{\mathrm{T}}(\mathrm{e})}$. The distribution calculated from 2010 data events (dots) is compared to the cross section obtained from the simulated signal sample (line), which is normalised to the NLO $\mathrm{t\bar{t}}$ quark pair cross section at $\sqrt{s} = 7$ TeV of 157.5 pb. (c) Distribution of the correlation between generated electron p_{T} (GEN) and reconstructed p_{T} (RECO) and (d) the corresponding purities and stabilities for each individual bin, derived from (c). The total selection efficiency, used for the calculation of the differential cross section in each bin of $p_{\mathrm{T}}\,(e)$, is shown in (d).



(a) (b)

(c) (d)

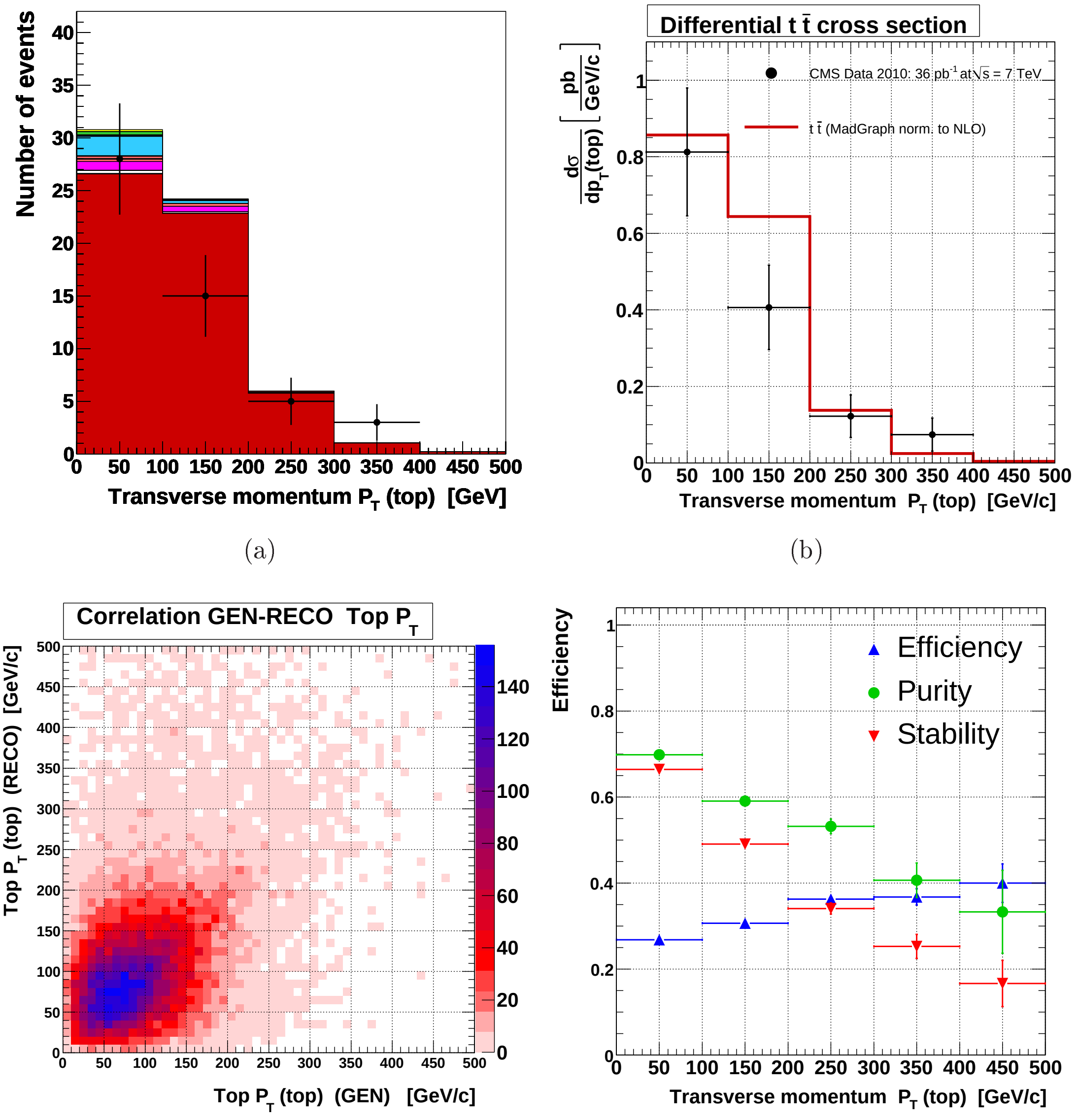
Figure 7.7: (a) **Muon-electron invariant mass** $M_{\mu e}$ after dilepton-plus-jets selection. Simulated samples are normalised to an integrated luminosity of 35.9 pb^{-1}. (b) Differential $t\bar{t}$ quark pair cross section as a function of the dilepton invariant mass $\frac{d\sigma_{t\bar{t}}}{dM_{\mu e}}$. The distribution calculated from 2010 data events (dots) is compared to the cross section obtained from the simulated signal sample (line), which is normalised to the NLO $t\bar{t}$ quark pair cross section at $\sqrt{s} = 7$ TeV of 157.5 pb. (c) Distribution of the correlation between generated dilepton mass (GEN) and reconstructed $M_{\mu e}$ (RECO) and (d) the corresponding purities and stabilities for each individual bin, derived from (c). The total selection efficiency, used for the calculation of the differential cross section in each bin of $M_{\mu e}$, is shown in (d).



(a) (b)

(c) (d)

Figure 7.8: The method for the kinematic event reconstruction, is used for the reconstruction of the invariant mass of the top-antitop quark pair. The value of the top quark mass, which is used as input for solving the kinematic equations, is varied between 100 and 300 GeV/c^2 in 1 GeV/c^2 steps. (a) Reconstructed **top-antitop quark pair invariant mass** $M_{t\bar{t}}$ after dilepton-plus-jets selection. Simulated samples are normalised to an integrated luminosity of 35.9 pb^{-1}. (b) Differential $t\bar{t}$ quark pair cross section as a function of the $t\bar{t}$ quark pair invariant mass $\frac{d\sigma_{t\bar{t}}}{dM_{t\bar{t}}}$. The distribution calculated from 2010 data events (dots) is compared to the cross section obtained from the simulated signal sample (line), which is normalised to the NLO $t\bar{t}$ quark pair cross section at $\sqrt{s} = 7$ TeV of 157.5 pb. (c) Distribution of the correlation between generated $t\bar{t}$ quark pair mass (GEN) and reconstructed $M_{t\bar{t}}$ (RECO) and (d) the corresponding purities and stabilities for each individual bin, derived from (c). The total selection efficiency, used for the calculation of the differential cross section in each bin of $M_{t\bar{t}}$, is shown in (d).



(a) (b)

(c) (d)

Figure 7.9: The method for the kinematic event reconstruction, is used for the reconstruction of the transverse momentum of the single top quark. (a) Reconstructed **single top quark transverse momentum** p_{T} (top) after dilepton-plus-jets selection. Simulated samples are normalised to an integrated luminosity of 35.9 pb^{-1}. (b) Differential $t\bar{t}$ quark pair cross section as a function of the reconstructed top quark transverse momentum $\frac{d\sigma_{t\bar{t}}}{dp_{\mathrm{T}}(\mathrm{top})}$. The distribution calculated from 2010 data events (dots) is compared to the cross section obtained from the simulated signal sample (line), which is normalised to the NLO $t\bar{t}$ quark pair cross section at $\sqrt{s} = 7$ TeV of 157.5 pb. (c) Distribution of the correlation between generated top quark p_{T} (GEN) and reconstructed p_{T} (top) (RECO) and (d) the corresponding purities and stabilities for each individual bin, derived from (c). The total selection efficiency, used for the calculation of the differential cross section in each bin of p_{T} (top), is shown in (d).

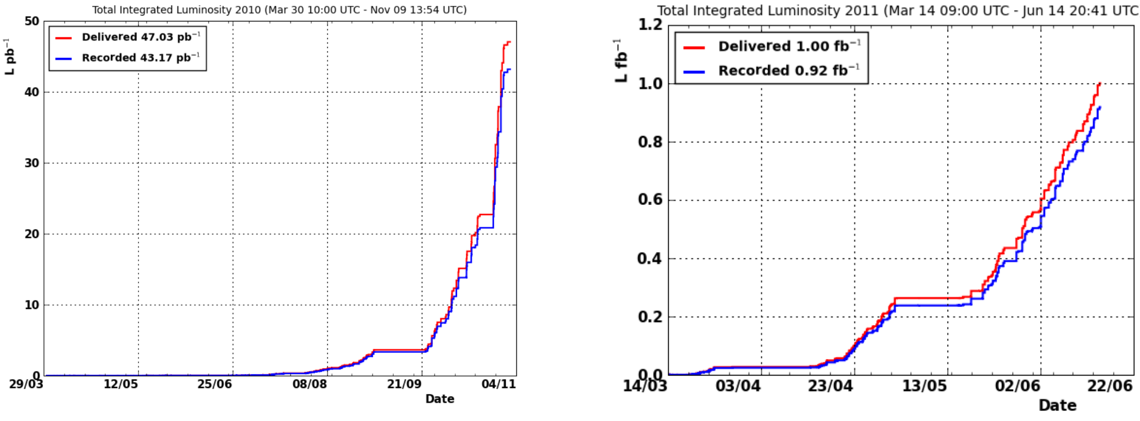
Chapter 8

Outlook

I have an appointment with eternity and I don't want to be late.
Star Trek VII - Generations

2011 Data Taking Period

An outlook is given of the results from the analysis of the data recorded in the first part of 2011. Figure 8.1 (b) shows the total integrated luminosity delivered by the LHC in the first part of 2011, reaching almost 1 fb^{-1}, which is by a factor of 20 higher than for the complete data taking period in 2010, shown for comparison in figure 8.1 (a). The inclusive top-antitop quark pair production cross section is measured in the muon-electron decay channel using an integrated luminosity of 191 pb^{-1}. Differential $\mathrm{t\bar{t}}$ quark pair cross sections are measured also in the μe final state as a function of the kinematic observables of the final state leptons. An almost identical event selection as for the analysis of the data recorded in 2010 is applied to the 2011 data sample, as well as to the simulated samples. The measurement of the top quark pair production cross section using all three dileptonic decay channels is described in [54].



(a) (b)

Figure 8.1: (a) Total integrated luminosity $L = \int dt\, \mathcal{L}$ for the complete data taking period in in 2010 compared to (b) the first data taking period in 2011.

Data Samples and Event Selection

A comprehensive description of the data samples used for the analysis of the data samples recorded in 2011 is given in [54]. The event selection is almost similar to the selection of the 2010 data samples, which has been described in chapter 6. In this outlook, only the *muon-electron triggered* data sample is used. The collision data sets used in this analysis are listed in table 8.1. These samples are reprocessed with CMS software version `CMSSW_4_1_4` using a list of *certified, good runs* given in [55]. The selected runs, which are used as input to the event selection, correspond to an integrated luminosity of 191 pb^{-1}. The simulated data samples used in this analysis correspond to the *Spring11* production series of Monte Carlo samples and comprise the same set of signal and background processes as described in chapter 6. Systematic uncertainties are not yet addressed in this analysis of the 2011 data sample.

Sample	Run range
`/MuEG/Run2011A-PromptReco-v1/AOD`	see [55]
`/MuEG/Run2011A-PromptReco-v2/AOD`	see [55]

Table 8.1: Collision data samples selected by muon-electron triggers of the 2011 data taking period, which are used in this analysis. The corresponding run ranges are given in [55].

Preliminary Cross Section Results

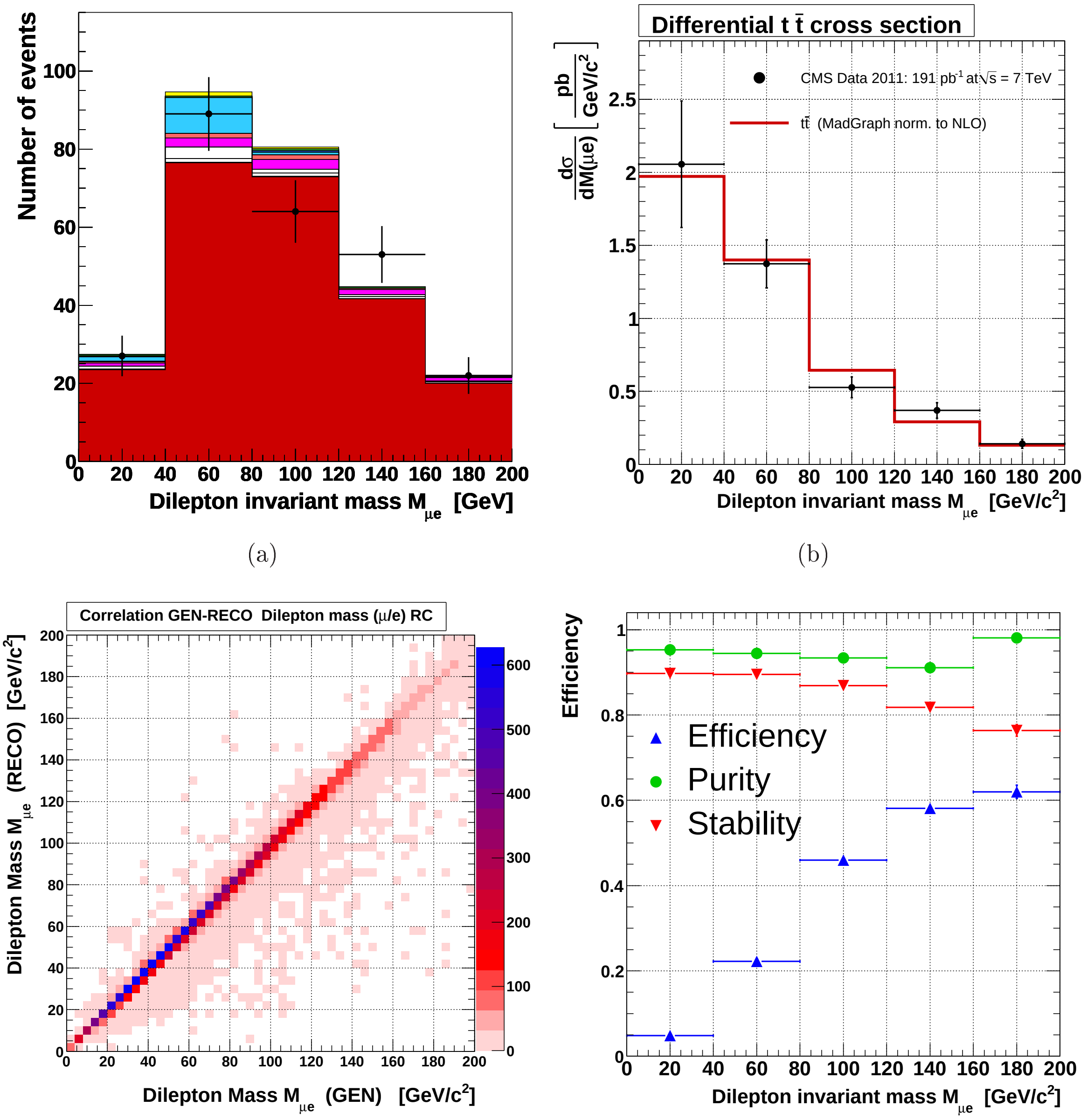
First results from the measurement of differential cross sections based on the data from the first part of the 2011 data taking periods are presented. The corresponding **inclusive $\text{t}\bar{\text{t}}$ quark pair cross section** in the muon-electron channel is $\sigma_{\text{t}\bar{\text{t}}}^{\mu e} = (\ 161\ \pm\ 11(\text{stat.}))$ pb, as reported in [54]. As described in section 7.6, the cross section of the production of top quark pairs is measured differentially as a function of the kinematic observables of the final state object, such as the transverse momentum p_{T} of the two leptons and the invariant mass of the lepton pair.

Dilepton Invariant Mass

The differential $\text{t}\bar{\text{t}}$ quark pair cross section as a function of the dilepton invariant mass $\frac{d\sigma_{\text{t}\bar{\text{t}}}}{dM_{\mu\text{e}}}$ is shown in figure 8.2 (b). The distribution calculated from the 2011 data events (dots) is compared to the cross section obtained from the simulated signal sample (line), which is normalised to the NLO $\text{t}\bar{\text{t}}$ quark pair cross section at $\sqrt{s} = 7$ TeV of 157.5 pb. For the 2011 data samples, an excellent agreement is found in each bin of $M_{\mu\text{e}}$. The reconstructed and generated muon-electron masses are also reasonably correlated, which is shown in figure 8.2 (c). The purity and stability are found to be almost identical compared to the 2010 data and are above 80% in most of the bins in $M_{\mu\text{e}}$, as it is shown in figure 8.2 (d). The selection efficiency increases significantly for higher dilepton masses. The corresponding distribution of the dilepton mass after applying the dilepton-plus-jets selection requirements is shown in figure 8.2 (a), which shows a good agreement between events from data and simulated events, as for the 2010 data samples.

Lepton Transverse Momentum

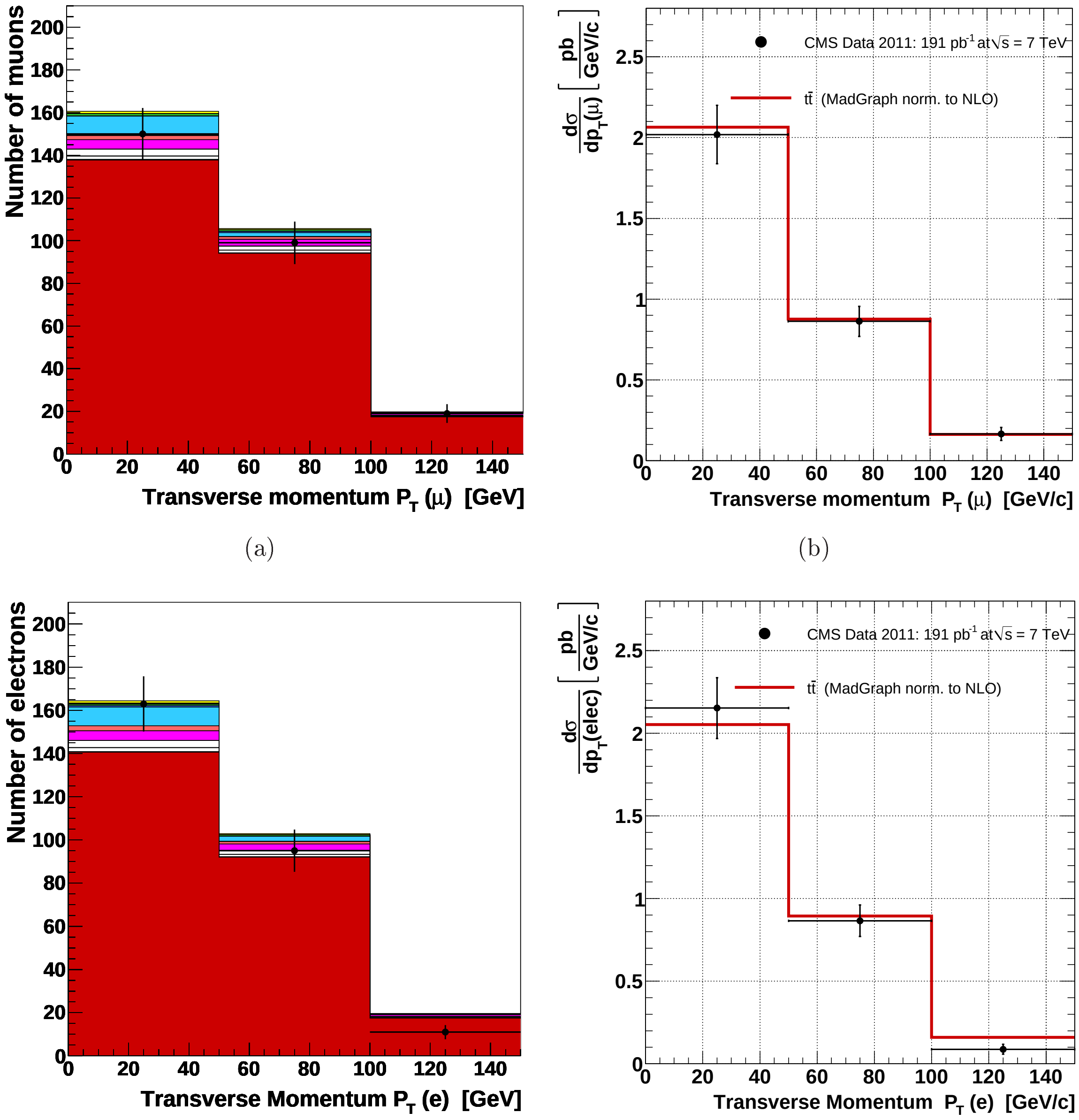
The distributions of the lepton momenta after applying the dilepton-plus-jets selection requirements are shown in figures 8.3 (a) and (c). A good agreement between events from data and simulation is obtained over the full range of the lepton momenta. The corresponding differential $\mathrm{t\bar{t}}$ quark pair cross sections as a function of the lepton transverse momenta $\frac{d\sigma_{\mathrm{t\bar{t}}}}{dp_{\mathrm{T}}(\mu)}$ and $\frac{d\sigma_{\mathrm{t\bar{t}}}}{dp_{\mathrm{T}}(\mathrm{e})}$ are shown in figures 8.3 (b) and (d). An almost perfect agreement is found between the measured cross section from data event and simulation in each bin of the muon and electron momenta within the statistical uncertainties. As mentioned above, only the *statistical uncertainties* are taken into account.



(a) (b)

(c) (d)

Figure 8.2: (a) **Muon-electron invariant mass** $M_{\mu e}$ after dilepton-plus-jets selection. Simulated samples are normalised to an integrated luminosity of 191 pb^{-1}. (b) Differential $\text{t}\bar{\text{t}}$ quark pair cross section as a function of the dilepton invariant mass $\frac{d\sigma_{\text{t}\bar{\text{t}}}}{dM_{\mu e}}$. The distribution calculated from 2011 data events (dots) is compared to the cross section obtained from the simulated signal sample (line), which is normalised to the NLO $\text{t}\bar{\text{t}}$ quark pair cross section at $\sqrt{s} = 7$ TeV of 157.5 pb. (c) Distribution of the correlation between generated dilepton mass (GEN) and reconstructed $M_{\mu e}$ (RECO) and (d) the corresponding purities and stabilities for each individual bin, derived from (c). The total selection efficiency, used for the calculation of the differential cross section in each bin of $M_{\mu e}$, is shown in (d).



(a) (b)

(c) (d)

Figure 8.3: (a) **Muon transverse momentum** p_{T} (μ) after dilepton-plus-jets selection. Simulated samples are normalised to an integrated luminosity of 191 pb^{-1}. (b) Differential $\mathrm{t\bar{t}}$ quark pair cross section as a function of the muon transverse momentum $\frac{d\sigma_{\mathrm{t\bar{t}}}}{dp_{\mathrm{T}}(\mu)}$. The distribution calculated from 2011 data events (dots) is compared to the cross section obtained from the simulated signal sample (line), which is normalised to the NLO $\mathrm{t\bar{t}}$ quark pair cross section at $\sqrt{s}$ = 7 TeV of 157.5 pb. (c) **Electron transverse momentum** p_{T} (e) after dilepton-plus-jets selection. (d) Differential $\mathrm{t\bar{t}}$ quark pair cross section as a function of the electron transverse momentum $\frac{d\sigma_{\mathrm{t\bar{t}}}}{dp_{\mathrm{T}}(\mathrm{e})}$.

Chapter 9

Summary and Conclusions

There are more things in heaven and earth, Horatio,
Than are dreamt of in your philosophy.
Hamlet Act I, Scene V

The measurement of the top quark pair production cross section in proton-proton collisions at $\sqrt{s} = 7$ TeV has been presented, using the muon-electron decay channel. Signal events are selected by requiring an oppositely-charged, well identified and isolated muon-electron pair and two additional high-p_{T} Particle-Flow jets, using a data sample, which corresponds to the full data set recorded in 2010 with an integrated luminosity of 35.9 pb^{-1}. Signal events are selected using a robust, cut-based event selection on an electron- or muon-triggered data sample. After requiring two additional high-p_{T} jets on top of the full dilepton selection, without any requirement of missing transverse energy or b-tagging, a significance of the top quark pair signal of almost 7σ is obtained.

Several correction factors are applied to the total selection efficiency, which is determined from simulated top quark pair events. The single-muon trigger efficiency is measured directly from the data using the *tag&probe* method. The remaining efficiencies for lepton identification, reconstruction and isolation, as well as the single-electron trigger efficiency are also corrected for the values, which have been derived from the tag&probe method. An estimation of the remaining background contributions from processes with two real, unlike-flavoured leptons, such as diboson production or $\mathrm{Z}^0/\gamma^\star \rightarrow \tau^+\tau^-$ events, is obtained from simulation, as well as processes with at least one misidentified muon or electron (*fake* leptons).

Based on 59 observed $\mathrm{t\bar{t}}$ quark pair candidate events, the measured cross section after applying the standard dilepton-plus-jets selection is $\sigma_{\mathrm{t\bar{t}}} = (156 \pm 25(\mathrm{stat.}) \pm 14(\mathrm{sys.}))$ pb. In order to validate the top quark-like topology of the selected events, a kinematic event reconstruction is applied to the final event sample. The cross section is recalculated for the number of events, in which a kinematic solution is found by the event reconstruction algorithm. Also a distribution for the resulting single top quark mass is obtained. Applying the kinematic event reconstruction algorithm to the selected event sample, the resulting cross section value is found to be $\sigma_{\mathrm{t\bar{t}}}^{\mathrm{kin}} = (159 \pm 27(\mathrm{stat.}) \pm 14(\mathrm{sys.}))$ pb.

As an alternative, also the requirement of at least one b-tagged jet is investigated on the $\mathrm{t\bar{t}}$ -enriched signal sample after the dilepton-plus-jets selection. Two different b-tagging discriminators are studied, provided by the track counting and the simple secondary vertex

algorithms. The *b-tagging efficiency* obtained from the data is found to be (92.5 ± 3.8)% for the TCHE algorithm at the *loose* working point and (63.0 ± 6.7)% for the SSVHE algorithm at the *medium* working point. Since the expected b-tagging efficiency for the SSVHE algorithm deviates significantly from the value measured in the data, it is corrected for the calculation of the total selection efficiency to the measured value. The cross section is calculated, requiring at least one b-tagged jet for both algorithms to be $\sigma_{\mathrm{t\bar{t}}}^{\mathrm{TCHE}} = (155 \pm 24(\mathrm{stat.}) \pm 16(\mathrm{sys.}))$ pb and $\sigma_{\mathrm{t\bar{t}}}^{\mathrm{SSVHE}} = (156 \pm 30(\mathrm{stat.}) \pm 16(\mathrm{sys.}))$ pb.
The *arithmetic average* of the different top quark pair cross section values, obtained after the standard dilepton-plus-jets selection, kinematic event reconstruction, as well as after the requirement of two different b jet identification algorithms, is $\langle\sigma_{\mathrm{t\bar{t}}}\rangle = (157 \pm 30)$ pb. All measured cross section values are found to be in good agreement with each other and with the NLO $\mathrm{t\bar{t}}$ quark pair cross section at $\sqrt{s} = 7$ TeV of 157.5 pb. A summary of the different inclusive top quark pair production cross sections is given in table 9.1.

Selection	Cross Section [pb]	Stat. uncer. [pb]	Sys. uncer. [pb]
≥ 2 jets	156.2	± 24.8	± 14.1
Kin. reco.	158.6	± 26.6	± 14.3
≥ 1 TCHEL	155.2	± 24.3	± 16.0
≥ 1 SSVHEM	156.4	± 29.5	± 16.1

Table 9.1: Summary of the different inclusive top quark pair cross section values obtained after full dilepton-plus-jets selection (≥ 2 jets). Optionally, a kinematic event reconstruction (Kin. reco.) and the requirement of at least one TCHE *loose* (≥ 1 TCHEL) and SSVHE *medium* (≥ 1 SSVHEM) b-tagged jet is applied.

The first results from the measurement of differential cross sections based on the data from the complete 2010 data taking period have been presented. For the first time in CMS, the cross section of the production of top quark pairs is measured differentially as a function of the kinematic observables of the final state object, such as the transverse momentum p_{T} of the muon and the electron, as well as the invariant mass of the lepton pair. Based on the solution of the kinematic event reconstruction, the cross section is also calculated differentially as a function of the kinematic properties of the reconstructed top-antitop quark pair. An excellent agreement is found between the measured cross sections from data events and simulation in each bin of the lepton momenta, the dilepton mass, as well as for the reconstructed top quark properties, within the statistical uncertainties. Finally, an outlook of the results from the analysis of the data recorded in the first part of 2011 has been given. First results from the measurement of differential cross sections were presented. Also for the data recorded in 2011, an excellent agreement is found between the measured cross sections from data events and simulation in each bin of the observables of the final state leptons.

Appendix A

Selected Candidate Events

A.1 Top Quark Candidate Events

Run	Lumi.	Event	$N_{\text{leptons}} = 2$ (RC μ e)	Dilepton mass $M_{\mu\text{e}}$ [GeV/c^2]	N_{jets} with $N_{\text{b-tags}} \geq 2$ ($p_{\text{T}} > 30$ GeV/c)
148829	280	273430518		176.97	2
148862	51	78710892		50.15	2
148029	457	357580841		124.43	2
147451	120	139902700		88.97	2
148862	154	247353758		124.64	2
143657	633	584500498		60.97	2
149291	751	759541584		24.57	2
149011	596	839574880	Figure A.1	123.87	2
148862	417	616863960		49.85	2
147926	357	380430462		42.91	2
149181	245	68403528		82.09	2
148862	357	532196916		98.58	2
147114	340	276016381		45.29	2
147926	493	531114775		149.28	2
149291	89	37887806		62.16	3
148864	225	267767817	Figure A.2	217.13	3
148864	284	334989072		42.89	3
148864	579	654374646		34.88	4
147757	240	203029597		189.82	5

Table A.1: Selected **top quark pair candidate** events after dilepton-plus-jets selection. Additionally, the selected events are required to have at least two THCE *medium* b-tagged jets. According to simulation, this constitutes an almost background-free top quark pair event sample. Two images created from the event display *fireworks* are shown in figures A.1 and A.2 below.

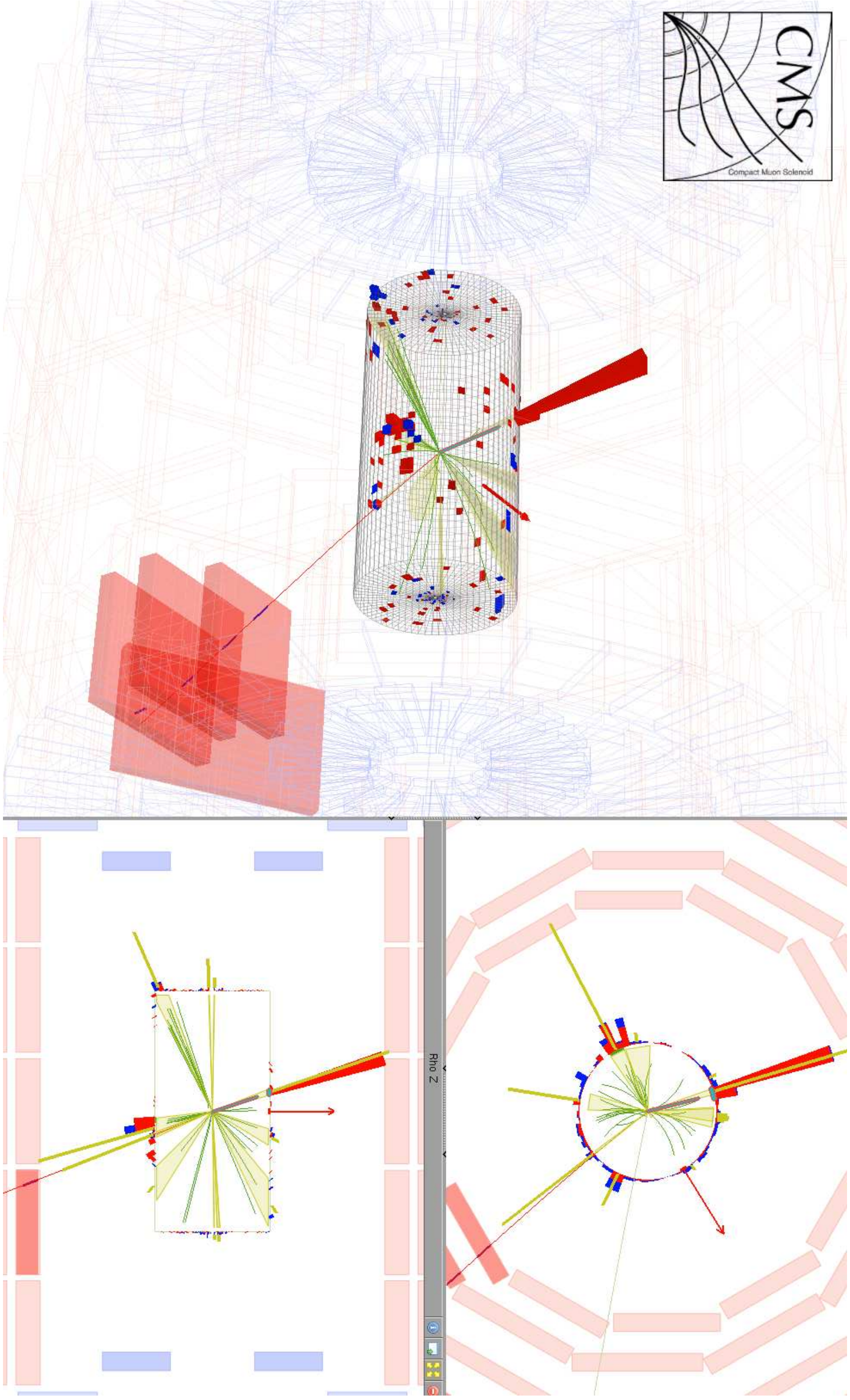


Figure A.1: Selected **top quark pair candidate** event (*Run:149011, Lumi:596, Event:839574880*) after dilepton-plus-jets selection given in table A.1. The selected event is required to have at least two TCHE *medium* tagged jets. The invariant mass of the muon-electron pair is $M_{\mu e} = 123.87$ GeV/c^2. *Lepton momenta*: $p_T(\mu) = 53.1$ GeV/c and $p_T(e) = 65.8$ GeV/c. $\not{E}_T = 36.0$ GeV.

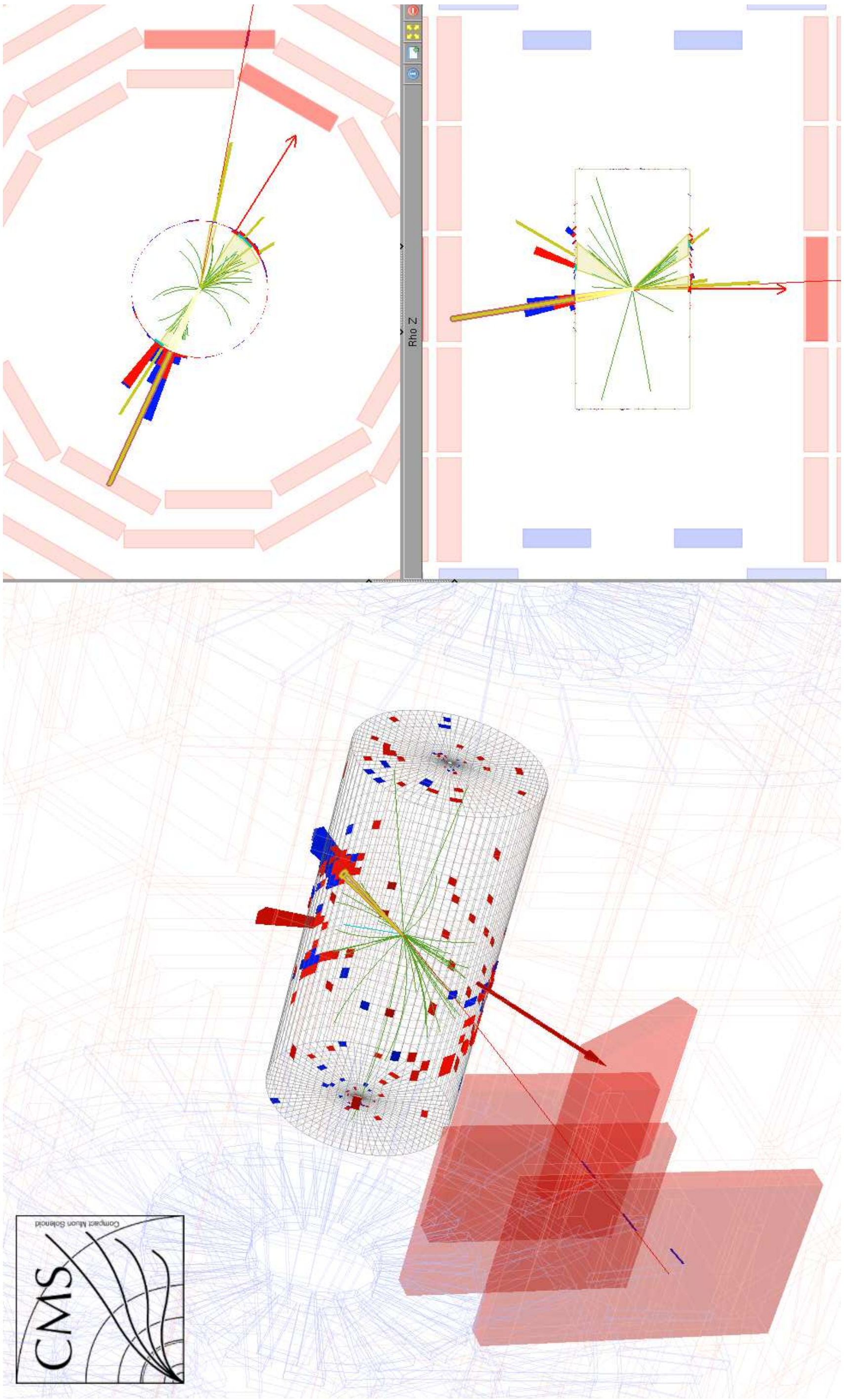


Figure A.2: Selected **top quark pair candidate** event (*Run:148864, Lumi:225, Event:267767817*) after dilepton-plus-jets selection given in table A.1. The selected events are required to have at least two TCHE *medium* tagged jets. . The invariant mass of the muon-electron pair is $M_{\mu e} = 217.13$ GeV/c^2. *Lepton momenta*: $p_{\mathrm{T}}\ (\mu) = 135.9$ GeV/c and $p_{\mathrm{T}}\ (e) = 87.7$ GeV/c. $\not\!\!E_{\mathrm{T}} = 197.9$ GeV. Momentum of leading-p_{T} jet: 252.1 GeV/c.

Run	Lumi.	Event	$N_{\mathrm{leptons}} = 2$ (RC μ e)	Dilepton mass $M_{\mu \mathrm{e}}$ [GeV/c^2]	N_{jets} ($p_{\mathrm{T}} > 30$ GeV/c)
149181	1340	1311956829		105.45	2
149063	84	98375451		207.21	2
149181	1706	1626022411		58.96	2
148829	155	153397876		47.46	2
148864	312	366351507		14.43	2
146644	1633	1255480064		73.51	2
147048	161	134625993		123.92	2
147115	2	878744		69.97	2
146807	178	183012929		156.10	2
146807	398	399257984		97.89	2
147454	68	60288052		83.56	2
149182	26	21483106		132.86	2
149181	754	723595683		106.46	2
148864	522	594577419		83.86	2
148862	82	131576801		75.40	2
149291	613	639394420		181.31	2
149291	683	700663787		188.81	2
149291	664	683687535		40.73	2
147219	102	69124299		127.27	2
147929	27	26117530		105.73	2
147390	103	91917863		106.80	2
148864	375	436534906		77.63	2
149291	362	387551026		119.19	2
147217	8	5192291		129.76	2
147927	4	2196303		75.22	2
149003	233	251972444		56.78	2
147390	701	564227635		183.28	2
149181	1194	1175707454		179.15	2
148031	251	213904114		49.80	2
147929	445	406944235		66.27	2
142971	235	222175234		75.91	2
142305	22	15915819		88.33	2
143962	227	244363110		133.62	2
147216	46	34569448		87.15	3
149003	176	170643045		61.40	3
148862	287	430865233		79.72	4
142038	702	367782938		28.85	4
149181	1345	1316743403		76.61	5
147926	190	174596074		65.26	5

Table A.2: Selected **top quark pair candidate** events fulfilling the dilepton-plus-jets selection.

A.2 Diboson WZ Candidate Event

Run	Lumi.	Event	$N_{\text{leptons}} = 3$	Dilepton mass $M_{\ell^+\ell^-}$ [GeV/c^2]	N_{jets} ($p_{\text{T}} > 30$ GeV/c)
147217	75	55188718	RC μ μ	90.66	2
147217	75	55188718	RC μ e	52.58	2
147217	75	55188718	WC μ e	135.65	2

Table A.3: Selected **diboson WZ candidate** event after dilepton-plus-jets selection. The invariant mass of the oppositely-charged muon pair is $M_{\ell^+\ell^-} = 90.66$ GeV/c^2. An image created from the event display *fireworks* is shown in figure A.3 below.

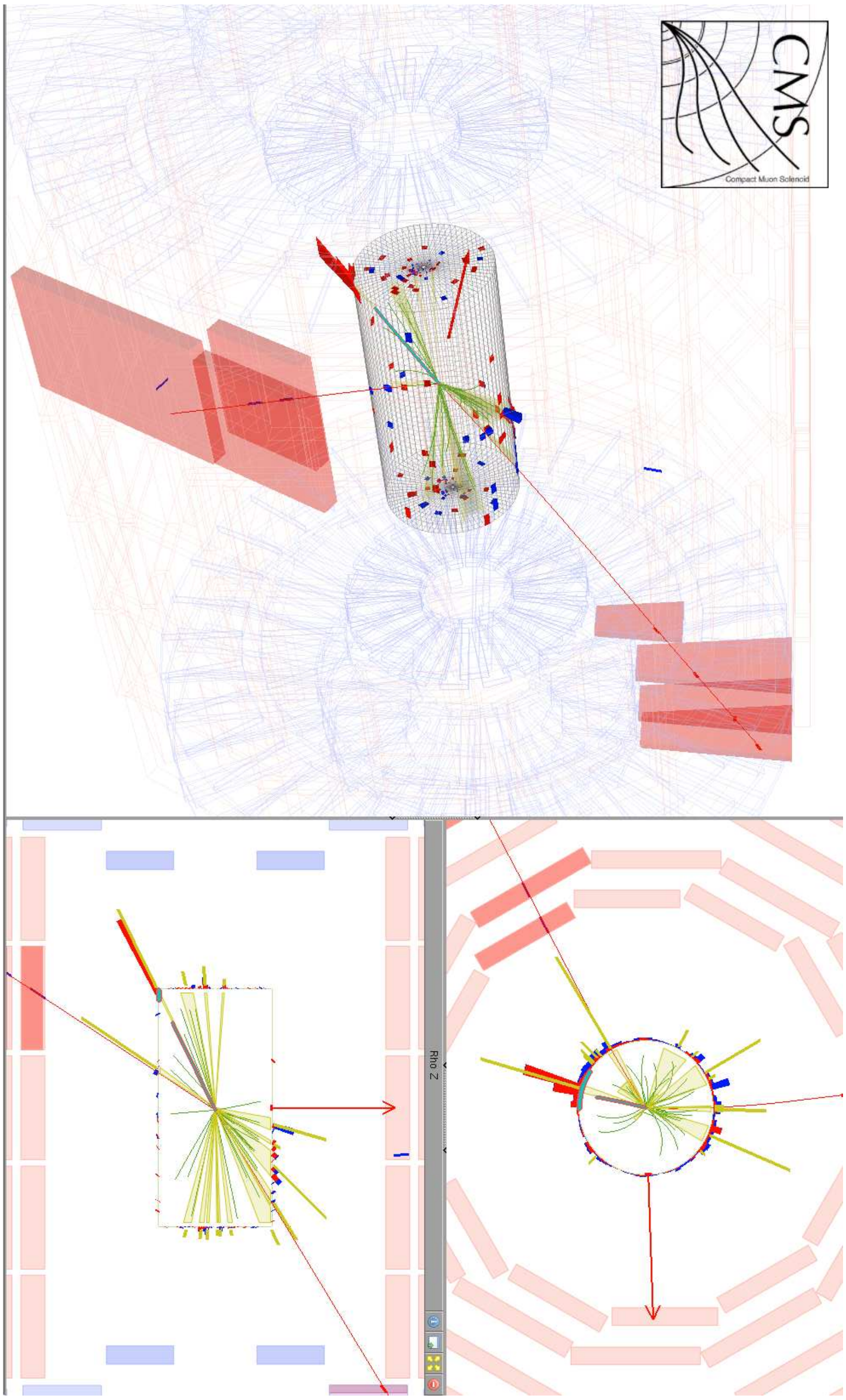


Figure A.3: Selected **diboson WZ candidate** event (*Run:147217, Lumi:75, Event:55188718*) after dilepton-plus-jets selection given in table A.3. The invariant mass of the oppositely-charged muon pair is $M_{\ell^+\ell^-} = 90.66$ GeV/c^2. *Lepton momenta*: $p_{\mathrm{T}}\,(\mu_1) = 22.4$ GeV/c, $p_{\mathrm{T}}\,(\mu_2) = 46.7$ GeV/c and $p_{\mathrm{T}}\,(e) = 47.5$ GeV/c. $\not\!\!E_{\mathrm{T}} = 68.4$ GeV.

Bibliography

[1] Michael E. Peskin and Daniel V. Schroeder. Introduction to quantum field theory. *Westview Press*, 1995.

[2] Donald H. Perkins. Introduction to high energy physics. *Westview Press*, 1995.

[3] K. Nakamura et al. (Particle Data Group). The Review of Particle Physics. *J. Phys.*, G37, 075021, 2010.

[4] ALEPH Collaboration, CDF Collaboration, DØ Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, SLD Collaboration, LEP Electroweak Working Group, Tevatron Electroweak Working Group, and SLD electroweak heavy flavour groups. Precision Electroweak Measurements and Constraints on the Standard Model. *CERN-PH-EP/2010-095*, 2010. arXiv:1012.2367v2.

[5] Siegfried Bethke. The 2009 World Average of α_s. *Eur.Phys.J.C64*, pages 689–703, 2009. doi: 10.1140/epjc/s10052-009-1173-1. arXiv:0908.1135v2.

[6] Torbjorn Sjostrand, Stephen Mrenna, and Peter Skands. PYTHIA 6.4 Physics and Manual. *JHEP 0605:026*, 2006. doi: LUTP06-13,FERMILAB-PUB-06-052-CD-T. arXiv:hep-ph/0603175v2.

[7] Anthony W. Thomas and Wolfram Weise. The Structure of the Nucleon. *WILEY-VCH Berlin*, 2001.

[8] The H1 and ZEUS Collaborations. Combined Measurement and QCD Analysis of the Inclusive ep Scattering Cross Sections at HERA. *JHEP*, 1001:109, 2010. doi: 10.1007/JHEP01(2010)109. arXiv:0911.0884v2[hep-ex].

[9] V. Gribov and L. Lipatov. *Sov.J.Nucl.Phys. 15*, page 438, 1972.

[10] G. Altarelli and G. Parisi. *Nucl.Phys. B126*, page 298, 1977.

[11] The CDF Collaboration. Observation of Top Quark Production in Pbar-P Collisions. *Phys.Rev.Lett.74*, pages 2626–2631, 1995. doi: 10.1103/PhysRevLett.74.2626. arXiv:hep-ex/9503002v2.

[12] S. Catani and et al. QCD. *CERN-TH/2000-131*, 2000. arXiv:hep-ph/0005025v1.

[13] Michael Spira. QCD Effects in Higgs Physics. *Fortsch.Phys. 46*, pages 203–284, 1998. arXiv:hep-ph/9705337v2.

[14] R. K. Ellis, W. J. Stirling, and B. R. Webber. QCD and Collider Physics. *Cambridge University Press*, 2003.

[15] CMS Collaboration. Combination of top pair production cross sections in pp collisions at 7 TeV and comparisons with theory. *CMS Physics Analysis Summary*, CMS-PAS-TOP-11-001, 2010.

[16] E. Shabalina on behalf of the CDF Collaboration and the DØ Collaboration. Top Quark Production Cross Section at the Tevatron. *FERMILAB-CONF-06-107-E*, 2006. arXiv:hep-ex/0605045v1.

[17] The CMS Collaboration. First Measurement of the Cross Section for Top-Quark Pair Production in Proton-Proton Collisions at $\sqrt{s} = 7$ TeV. *Phys.Lett.B*, 695: 424–443, 2011. doi: 10.1016/j.physletb.2010.11.058. arXiv:1010.5994v1.

[18] CMS TWiki. Standard Model Cross Sections. 2010. `https://twiki.cern.ch/twiki/bin/view/CMS/StandardModelCrossSections`.

[19] CMS Collaboration. Measurement of the t t-bar production cross section and the top quark mass in the dilepton channel in pp collisions at sqrt(s) = 7 TeV. *CERN-PH-EP-2011-055*, arXiv:1105.5661v1[hep-ex], 2011.

[20] Tevatron Electroweak Working Group, the CDF Collaboration, and the DØ Collaboration. Combination of CDF and D0 Results on the Mass of the Top Quark. *FERMILAB-TM-2466-E*, 2010. arXiv:1007.3178v1[hep-ex].

[21] Oliver Sim Brüning, Paul Collier, P Lebrun, Stephen Myers, Ranko Ostojic, John Poole, and Paul Proudlock. LHC Design Report. 2004.

[22] CMS Collaboration. CMS Physics Technical Design Report, Volume I: Detector Performance and Software. 2006.

[23] Christiane Lefèvre. The CERN accelerator complex. Complexe des accélérateurs du CERN. 2008. `http://cdsweb.cern.ch/record/1260465`.

[24] CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation (JINST)*, 3 S08004, 2008. doi: 10.1088/1748-0221/3/08/S08004.

[25] CMS Collaboration. The TriDAS Project Technical Design Report, Volume 1: The Trigger Systems. 2000.

[26] L. Tuura, A. Meyer, I. Segoni, and G. Della Ricca. CMS data quality monitoring: Systems and experiences. *J. Phys.*, Conf. Ser. 219 072020, 2010. doi: 10.1088/1742-6596/219/7/072020. 10.1088/1742-6596/219/7/072020.

[27] Markus Marienfeld, Dirk Dammann, and Achim Geiser. Online/Offline Data Quality Monitoring for Top-like Dilepton Events. *CMS Analysis Note*, 2010. CMS AN-2010/062.

[28] Nadia Adam, Nicola Bachetta, Seth Cooper, Giovanni Franzoni, Valerie Halyo, Adam Hunt, Yuichi Kubota, Daniel Marlow, Marco Zanetti, and Andrzej Zuranski. Absolute Calibration of Luminosity Measurement at CMS. *CMS Analysis Note*, 2010. CMS AN-2011/093.

[29] CMS Collaboration. Measurement of CMS Luminosity. *CMS Physics Analysis Summary*, CMS-PAS-EWK-10-004, 2010.

[30] Johan Alwall, Pavel Demin, Simon de Visscher, Rikkert Frederix, Michel Herquet, Fabio Maltoni, Tilman Plehn, David L. Rainwater, and Tim Stelzer. MadGraph/MadEvent v4: The New Web Generation. *JHEP0709:028*, 2007. doi: 10.1088/1126-6708/2007/09/028. arXiv:0706.2334v1.

[31] Z. Was. TAUOLA the library for tau lepton decay, and KKMC/KORALB/KORALZ/... status report. *Nucl.Phys.Proc.Suppl.*, 98:96–102, 2001. doi: 10.1016/S0920-5632(01)01200-2. arXiv:hep-ph/0011305v1.

[32] The CMS Collaboration. CMS detector. 2011. `http://cms.web.cern.ch/cms/Detector/Work/index.html`.

[33] CMS Collaboration. Performance of muon identification in pp collisions at $\sqrt{s} = 7$ TeV. *CMS Physics Analysis Summary*, 2010. CMS PAS MUO-10-002.

[34] S. Baffioni, C. Charlot, F. Ferri, D. Futyan, P. Meridiani, I. Puljak, C. Rovelli, R. Salerno, and Y. Sirois. Electron reconstruction in CMS. *CMS Note*, 2006. CMS NOTE-2006/040.

[35] CMS TWiki. Electron Identification Based on Simple Cuts. 2010. `https://twiki.cern.ch/twiki/bin/view/CMS/SimpleCutBasedEleID`.

[36] V. Chetluru, F. Pandolfi, P. Schieferdecker, and M. Zielinski. Jet Reconstruction Performance at CMS. *CMS Analysis Note*, 2009. CMS AN-2009/067.

[37] CMS Collaboration. Commissioning of b-jet identification with pp collisions at $\sqrt{s}$ = 7 TeV. *CMS Physics Analysis Summary*, CMS-PAS-BTV-10-001, 2010.

[38] CMS Collaboration. Missing Transverse Energy Performance in Minimum-Bias and Jet Events from Proton-Proton Collisions at $\sqrt{s} = 7$ TeV. *CMS Physics Analysis Summary*, CMS-PAS-JME-10-004, 2010.

[39] CMS Collaboration. Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. *CMS Physics Analysis Summary*, CMS-PAS-PFT-09-001, 2009.

[40] CMS TWiki. Pass6 reference analysis. 2010. `https://twiki.cern.ch/twiki/bin/view/CMS/TopDileptonRefAnalysis2010Pass6`.

[41] `cms-service-dqm.web.cern.ch`. Nov22 certification. 2010. `https://cms-service-dqm.web.cern.ch/cms-service-dqm/CAF/certification/Collisions10/7TeV/Reprocessing/Cert_136033-149442_7TeV_Nov4ReReco_Collisions10_JSON.txt`.

[42] CMS TWiki. FALL10 - 7 TeV Samples. 2010. `https://twiki.cern.ch/twiki/bin/view/CMS/TopSamplesAndXSections7TeV36SeriesFall10`.

[43] CMS TWiki. Generator Cross Sections. 2010. `https://twiki.cern.ch/twiki/pub/CMS/GeneratorMain/ShortXsec.pdf`.

[44] CMS TWiki. Physics Analysis Toolkit (PAT). 2010. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuidePAT`.

[45] CMS TWiki. Conversion Rejection. 2010. `https://twiki.cern.ch/twiki/bin/view/CMS/ConversionBackgroundRejection`.

[46] CMS Collaboration. Jet Performance in pp Collisions at 7 TeV. *CMS Physics Analysis Summary*, 2010. CMS-PAS-JME-10-003.

[47] CMS Collaboration. Jet Energy Corrections determination at 7 TeV. *CMS Physics Analysis Summary*, 2010. CMS-PAS-JME-10-010.

[48] CMS TWiki. Jet Identification. 2010. `https://twiki.cern.ch/twiki/bin/view/CMS/JetID`.

[49] D. Barge and et al. A measurement of top quark pair production cross section in dilepton final states in pp collisions at 7 TeV. *CMS Analysis Note*, CMS AN-2010/410, 2010.

[50] B. Abbott and et al. (The DØ Collaboration). Measurement of the Top Quark Mass Using Dilepton Events. *Phys. Rev. Lett.*, 80:2063–2068, 1998. arXiv:hep-ex/9706014v1.

[51] J. Andrea, A. Avetisyan, U. Heintz, M. Narain, and T. Speer. Measurement of the Top quark mass with fully-leptonic $\mathrm{t\bar{t}}$ decays in pp collisions at $\sqrt{s} = 7$ TeV with the full 2010 dataset. *CMS Analysis Note*, 2010. CMS AN-2010/369.

[52] Lars Sonnenschein. Algebraic approach to solve $t\bar{t}$ dilepton equations. *Phys.Rev.D*, 2005. doi: 10.1103/PhysRevD.72.095020. arXiv:hep-ph/0510100v1.

[53] Dirk Dammann. Production Cross Section Measurement of Top-Antitop Pairs in the Dimuon Decay Channel at $\sqrt{s} = 7$ TeV with the CMS Experiment. *DESY-THESIS-2011-016*, ISSN 1435-8085, 2011.

[54] M. Aldaya, W. Behrenhoff, D. Dammann, A. Geiser, J. Hauk, M. Marienfeld, and A.B. Meyer. Measurement of the Top Quark Pair Production Cross Section in the Di-Lepton Final State at $\sqrt{s} = 7$ TeV. *CMS Analysis Note*, 2011. CMS AN-2011/186.

[55] `cms-service-dqm.web.cern.ch`. May13 certification. 2011. `file:///afs/cern.ch/cms/CAF/CMSCOMM/COMM_DQM/certification/Collisions11/7TeV/Prompt/Cert_160404-163869_7TeV_PromptReco_Collisions11_JSON.txt`.

List of Figures

Danksagung

Wenn einer alleine träumt, so ist es nur ein Traum.
Wenn viele gemeinsam träumen, so ist es der Beginn einer neuen Wirklichkeit.
Hélder Pessoa Câmara

Ich möchte an dieser Stelle den Personen meinen Dank aussprechen, die maßgeblich zum Gelingen meiner Promotion beigetragen haben. Zu allererst möchte ich meinem Betreuer Achim Geiser danken. Seine Offenheit und Ehrlichkeit haben mir in vielen Momenten meiner Arbeit die Klarheit vermittelt, die nötig war, um das Ziel wieder zu fokussieren. Ebenso seine persönliche und herzliche Art werden mir in guter Erinnerung an die Zeit als wissenschaftliche Hilfskraft bleiben. Andreas Meyer begleitete meine bisherige "Karriere" am Deutschen Elektronen-Synchrotron im wahrsten Sinne des Wortes vom ersten Tag an. Andreas war es, der mich für das CMS-Experiment warb und mich bereits durch die Höhen und Tiefen meiner Diplomarbeit begleitete. Mein Aufenthalt am CERN im Spätsommer und Herbst 2008 verdanke ich ebenfalls seiner Initiative - diese Zeit wird mir in guter Erinnerung bleiben. Ich danke ihm auch für seine unvergleich kritisch und direkte, aber auch jederzeit ehrliche Art. Nicht zuletzt möchte ich auch meinen Kollegen und langjährigen Mitdoktoranden Dirk, Wolf, Johannes und Alex danken. Meine Art war sicherlich nicht immer die Einfachste und mit mir in einem Büro zu sitzen oftmals anstrengend, wenn ich mal wieder nicht leise denken konnte. Alex danke ich besonders für die hilfreichen Korrekturen in der finalen Phase der Doktorarbeit.

Diese Arbeit ist in besonderer Weise der Erinnerung an meinen Großvater Alfred gewidmet. Ihm habe ich maßgeblich die Entdeckung meiner Faszination an den Naturwissenschaften zu verdanken. Von ihm lernte ich, mit Begeisterung zu den Sternen am Firmament aufzuschauen und in Gedanken auf die Reise zu fernen Galaxien und in das Innerste der Materie zu gehen. Niemand hat mich mehr in dem Streben nach Wissen und Erkenntnis inspiriert als er. Meine gesamte Familie ist seit jeher ein wichtiger Rückhalt für mich gewesen und sie hat mich in vieler Hinsicht auf meinem Weg zur erfolgreichen Promotion unterstützt. Sie hat mir stets das Gefühl gegeben, dass ich in dem von mir Erreichten etwas Besonderes geleistet habe und dies hat mich stets mit tiefem, aufrichtigem Stolz erfüllt. Ich möchte meine Danksagung beschließen, indem ich meiner lieben Levke danke, die mein Leben veränderte, seitdem ersten Moment, an dem ich sie getroffen habe, sowie ich es niemals zu träumen gewagt hätte.