Measurement Of The Top Quark Pair Production Cross Section In 7 TeV pp Collisions

BY

IOANA MARIA ANGHEL
B.Sc. (Technical University of Cluj-Napoca) 2002

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Physics
in the Graduate College of the
University of Illinois at Chicago, 2011

Chicago, Illinois

Defense Committee:

Cecilia Gerber, Chair and Advisor
Mark Adams
David Hofman
Wai-Yee Keung
Christian Schwanenberger, University of Manchester / Fermilab

To my parents

ACKNOWLEDGMENTS

I would like to express my profound gratitude to my advisor, Prof. Cecilia Gerber. Her experience and the broad knowledge of physics not only guided me, but constantly inspired me all these years. Her support and consideration encouraged me to work hard to reach my goal. She represents the model of the person I would like to be.

I am hugely indebted to Dr. Francisco Yumiceva, for his advises and continuous help. I thank him for all the constructive discussions, for his patience and for sharing with me his own work experience.

I would like to thank all the people from the CMS Top group, those who helped pushing this analysis forward and shared their experience.

The success of my work performed for the commissioning of the CMS Silicon Strip Tracker was possible only with the support of Dr. Slawomir Tkaczyk and Dr. Didier Contardo and I thank them for that. I am in particular grateful to Dr. Domenico Giordano for being such an inspiring physicist and for all the opportunities that he offered me while working on the SST data quality monitoring.

I thank all the faculty members of UIC group at CMS, Prof. Mark Adams, Prof. Nikos Varelas and Prof. Richard Cavanaugh for their assistance. I also thank Dr. Victor Bazterra and Dr. Leonard Apanasevich for always being ready to help me.

ACKNOWLEDGMENTS (Continued)

This analysis, along with many others, represents the result of the work of all my dedicated collaborators that built and operated the CMS detector and I feel honored for being part of this amazing experiment.

I thank all my friends at UIC for this unforgettable experience.

I thank my parents and my brother for always being there for me, for their unconditional love and support.

Finally, I thank Cosmin for loving me so much.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

CHAPTER PAGE

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF TABLES (Continued)

TABLE **PAGE**

LIST OF TABLES (Continued)

TABLE **PAGE**

LIST OF TABLES (Continued)

TABLE **PAGE**

LIST OF FIGURES

LIST OF FIGURES (Continued)

FIGURE **PAGE**

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF ABBREVIATIONS

ALICE	**A** **L**arge **I**on **C**ollider **E**xperiment
APDs	**A**valanche **P**hoto**D**iode**s**
ATLAS	**A** **T**oroidal **LHC** **A**pparatu**S**
BSM	**B**eyond **S**tandard **M**odel
BR	**B**ranching **R**atio
CERN	European Organisation for Nuclear Research (French: **C**onseil **E**uropeen pour la **R**echerche **N**ucleaire)
CMS	**C**ompact **M**uon **S**olenoid
CSC	**C**athode **S**trip **C**hamber
EB	**E**lectromagnetic **B**arrel Calorimeter
DAQ	**D**ata **A**cquisition **S**ystem
DT	**D**rift **T**ube
ECAL	**E**lectromagnetic **CAL**orimeter
EE	**E**lectromagnetic **E**ndcaps Calorimeter
FED	**F**ront **E**nd **D**river

LIST OF ABBREVIATIONS (Continued)

FSR	**F**inal **S**tate **R**adiation
GSF	**G**aussian **S**um **F**ilter
HB	**H**adron **B**arrel Calorimeter
HCAL	**H**adron **CAL**orimeter
HE	**H**adron **E**ndcap Calorimeter
HF	**H**adron **F**orward Calorimeter
HLT	**H**igh **L**evel **T**rigger
HO	**H**adron **O**uter Calorimeter
HPD	**H**ybrid **P**hotodiode
ISR	**I**nitial **S**tate **R**adiation
JES	**J**et **E**nergy **S**cale
L1T	**L**evel **1** **T**rigger
LEP	**L**arge **E**lectron−**P**ositron **C**ollider
LHC	**L**arge **H**adron **C**ollider
LHCb	**L**arge **H**adron **C**ollider **b**eauty
LHCf	**L**arge **H**adron **C**ollider **f**orward
LO	**L**eading **O**rder
MC	**M**onte **C**arlo

LIST OF ABBREVIATIONS (Continued)

NLO	**N**ext **L**eading **O**rder
NNLO	**N**ext to **N**ext **L**eading **O**rder
QCD	**Q**uantum **C**hromo **D**ynamics
QRF	**R**adio **F**requency **Q**uadrupole
PDF	**P**arton **D**istribution **F**unction
PF	**P**article **F**low
PS	**P**roton **S**ynchrotron
PSB	**P**roton **S**ynchrotron **B**ooster
PV	**P**rimary **V**ertex
RPC	**R**esistive **P**late **C**hamber
SF	**S**cale Factor
SL	**S**uper **L**ayer
SM	**S**tandard **M**odel
SPS	**S**uper **P**roton **S**ynchrotron
SST	**S**ilicon **S**trip **T**racker
SSV	**S**imple **S**econdary **V**ertex
SSVHEM	**S**imple **S**econdary **V**ertex **H**igh **E**fficiency **M**edium
TEC	**T**racker **E**nd **C**ap

LIST OF ABBREVIATIONS (Continued)

TPG	**T**rigger **P**rimitive **G**enerator
TID	**T**racker **I**nner **D**isk
TIB	**T**racker **I**nner **B**arrel
TOB	**T**racker **O**uter **B**arrel
TriDAS	**Tri**gger and **D**ata **A**cquisition **S**ystem
TOTEM	**TOT**al **E**lastic and diffractive cross section **M**easurement
VPTs	**V**acuum **P**hoto**T**riodes

SUMMARY

In this thesis we present the measurement of the top quark pair production cross section in pp collisions at a center of mass energy of 7 TeV, using 36 pb^{-1} of data collected with the CMS experiment at the LHC. The selection of the data sample is based on the final state with one isolated, high transverse momentum muon, missing transverse energy and at least three hadronic jets. The $t\bar{t}$ content of the selected sample is enhanced by requiring the presence of at least one jet consistent with b-quark hadronization. The measured cross section:

$$\sigma_{t\bar{t}} = 153.7 \pm 9.0(\text{stat.})^{+28.2}_{-24.4}(\text{syst.}) \pm 6.1(\text{lumi}) \text{ pb}$$

is in good agreement with higher-order QCD calculations for a top quark of mass m_t = 173 GeV/c^2 (1).

CHAPTER 1

INTRODUCTION

"The important thing in science is not so much to obtain new facts as to discover new ways of thinking about them."

William Lawrence Bragg

The existence of the sixth member of the family of quarks, the top quark, had been suggested ever since 1977, following the discovery of its electroweak partner, the bottom quark (2). Its direct observation by the CDF and DØ experiments at the Fermilab Tevatron Collider in 1995 was a great success of the Standard Model (SM) of Elementary Particle and Fields. Figure 1 shows the development of limits and measurements on the top quark mass from indirect and direct studies at e^+e^- and hadron colliders (3), (4). The top quark was discovered with a mass of exactly the value that was predicted from global fits to electroweak precision data.

With a mass comparable to the mass of a gold atom, the top quark occupies a unique position within the SM. The very large mass of the top quark arises from its couplings to the symmetry breaking sector of the SM, and this property makes it a unique probe of physics at the natural electroweak scale.

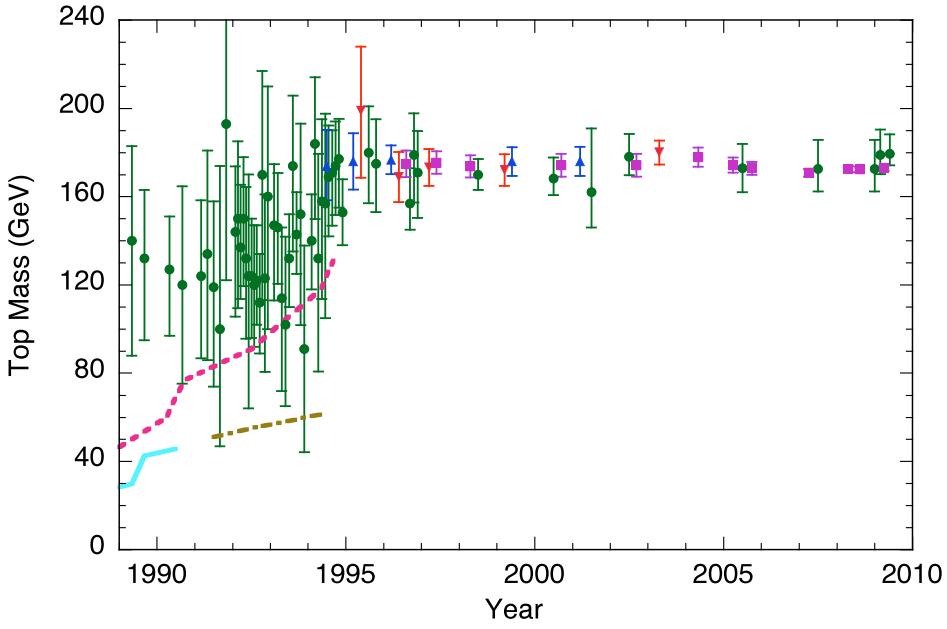

Figure 1. History of the limits on or measurements of the top quark mass: the full circles are the indirect bounds on the top quark mass from precision electroweak data; the squares represent world average direct measurement of the top quark mass; the triangles are the published CDF and DØ measurements; the dash-dotted and dashed lines represent the lower bounds from $p\bar{p}$ and e^+e^- colliders.

At hadron colliders, top quarks are predominantly produced as top-antitop ($t\bar{t}$) pairs. At the LHC, the $t\bar{t}$ production mechanism is dominated by gluon fusion, in contrast to the Tevatron, where the $t\bar{t}$ pairs were produced mainly through quark-antiquark annihilation. LHC has been designed to be a top quark factory, with about one top quark pair produced per second at design luminosities. Such a large sample of top quarks will eventually enable detailed studies of the top quark properties, and searches for new physics in top decays, leading to a deeper understanding of electroweak symmetry breaking and the origin of

mass. In particular, the measurement of the $t\bar{t}$ cross section at the LHC can provide not only important tests of our understanding of perturbative QCD, but it can also be used in searches for new physics.

In this thesis we present a measurement of the $t\bar{t}$ pair production cross section in the muon+jets channel using 7 TeV center-of-mass energy data with an integrated luminosity of 36 pb^{-1}. The theoretical aspects regarding the top quark, its decay and production are presented in Chapter 2. The experiment used to collect the data is described in Chapter 3. The methods used to identify the particles emerging from the pp collisions are described in Chapter 4. The selection of the $t\bar{t}$ events is presented in Chapter 5 and the estimation of the backgrounds is shown in Chapters 6 and 7. The $t\bar{t}$ production cross section extraction is described in Chapter 8.

CHAPTER 2

THEORETICAL INTRODUCTION

2.1 Standard Model

The SM (5), (6), (7), (8), (9) is a non-abelian Yang-Mills gauge theory based on a set of fields and the symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$, where C, L and Y stand for color quantum number, left handed chirality (weak isospin) and hypercharge, respectively. The fields correspond to the two groups of spin 1/2 particles (fermions), quarks and leptons. Each group consists of six particles, which are related in pairs, or **generations**.

Table I and Table II show the three generations of leptons and quarks and their properties (3). For each fermion, an anti-fermion exists, having the same mass and opposite electric charge. The lightest and most stable particles make up the first generation, whereas the heavier and less stable particles belong to the second and third generations. All stable visible matter in the Universe is made from particles that belong to the first generation, because any heavier particles quickly decay to a more most stable level. The information regarding the strength of the flavor changing weak decays of the quarks is contained in a unitary matrix, the Cabbibo-Kobayashi-Maskawa (CKM) matrix (8). The probability of a transition from quark q_i into a quark q_j is proportional to the corresponding CKM matrix element squared, $\mid V_{ij} \mid^2$.

Generation	**Lepton Flavor**	**Charge (e^-)**	**Mass (MeV/c^2)**
I	electron (e)	-1	0.511
	e neutrino (ν_e)	0	$< 2.2 \times 10^{-6}$
II	muon (μ)	-1	106
	μ neutrino (ν_μ)	0	< 0.17
III	tau (τ)	-1	1.776×10^3
	τ neutrino (ν_τ)	0	< 15.5

TABLE I

LEPTONS IN THE STANDARD MODEL.

Generation	**Quark Flavor**	**Charge (e^-)**	**Mass (MeV/c^2)**
I	up (u)	+2/3	2.5 ± 0.8
	down (d)	-1/3	5.0 ± 0.9
II	charm (c)	+2/3	$(1.27 \pm 0.09) \times 10^3$
	strange (s)	-1/3	101 ± 29
III	top (t)	+2/3	$(173.1 \pm 1.3) \times 10^3$
	bottom (b)	-1/3	$(4.19 \pm 0.18) \times 10^3$

TABLE II

QUARKS IN THE STANDARD MODEL.

There are four fundamental forces in the Universe: gravitational, electromagnetic, weak and strong interactions. The SM includes all of them, except the gravity. The inclusion of gravitation into a consistent theory with the other forces is not yet possible. Since the magnitude of the gravitation is many orders of magnitude smaller compared to the other three forces, it can be neglected in most particle physics calculations. Within the SM, the forces are mediated by one or more gauge bosons with integer spin, following Bose-Einstein statistics: the strong force is mediated by gluons, the electromagnetic force is mediated by photons, while the weak force mediators are the W^+, W^- and Z bosons. All quarks and leptons can interact electromagnetically and via the weak interaction, but only quarks interact via the strong interaction. The fundamental interactions included in the SM together with their bosons and the particles they act upon are listed in Table III (3). In addition, the yet unobserved Higgs boson is thought to be responsible for the fact that the particles have mass.

2.2 Top Quark Production

At the LHC, the top quark is produced mainly in pairs via the strong interaction, but is also produced singly via the charged-current weak interactions (single top production).

2.2.1 Top Quark Production via the Strong Interaction

The underlying theoretical framework for the calculation of hadronic cross sections is the Quantum Chromodynamics (QCD)-the "improved" parton model which regards a high energy hadron as a composition of quasi-free partons (quarks and gluons) sharing the

Force	**Mediator**	**Charge** (e^-)	**Mass** (MeV/c^2)	particles that are affected
Strong	8 gluons (g)	0	0	quarks, gluons
Electromagnetic	photon (γ)	0	0	quarks, charged leptons
Weak	$W^{\pm}$	± 1	$(80.399 \pm 0.023) \times 10^3$	quarks, leptons
	Z^0	0	$(91.188 \pm 0.002) \times 10^3$	$W^{\pm}$, Z

TABLE III

FORCES BETWEEN FERMIONS AND THEIR MEDIATORS.

hadron momentum. Each parton i carries a momentum fraction x_i, as shown in Figure 2.

Based on the factorization theorem, cross sections can be calculated as a convolution of parton distribution functions (PDF) $f(x, \mu^2)$ for colliding hadrons (p,p) and the factorized hard parton-parton cross section $\widehat{\sigma_{ij}}$:

$$\sigma(pp \to t\bar{t}) \quad = \quad \sum_{i,j=q,\bar{q},g} \int dx_i dx_j f_i(x_i, \mu^2) f_j(x_j, \mu^2) \cdot \widehat{\sigma}_{ij}(ij \to t\bar{t}; \widehat{s}, \mu^2). \qquad (2.1)$$

At the LHC, the hadrons are protons (pp). The variable $\widehat{s}$ denotes the square of the center of mass energy of the colliding partons. The sum in Equation 2.1 runs over all pairs of partons (i, j) contributing to the process. The PDF $f_i(x_i, \mu^2)$ describes the probability density for finding a parton i inside the hadron carrying a momentum fraction x_i. The PDFs as well as $\widehat{\sigma}_{ij}$ have a residual dependence on the factorization μ_F^2 and

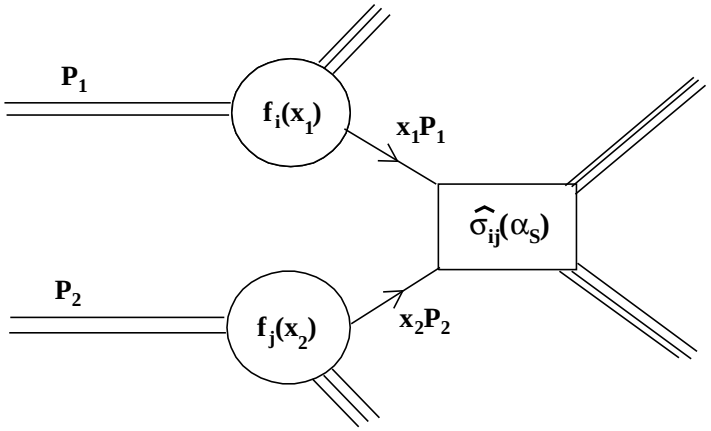

Figure 2. Parton model description of a hard scattering process.

renormalization scale μ_R^2. It is common to choose the same scale μ for both μ_F and μ_R. The factorization scheme serves as a method to systematically eliminate collinear divergencies from $\widehat{\sigma}_{ij}$ and absorb them into the PDFs, which are extracted from global fits to experimental data.

Figure 3 shows PDFs of the CTEQ collaboration (10) as a function of the longitudinal momentum fractions x at $\mu^2 = m_t^2$ for a top mass $m_t = 173$ GeV. For the $t\bar{t}$ production, $\widehat{s} \to (2m_t)^2$ and in this case $x_i = x_j = 2m_t/\sqrt{s}$. The gluons start to dominate in the x region below 0.15. At the LHC, for a center of mass energy of 7 TeV and $m_t = 173$ GeV, the threshold for producing $t\bar{t}$ is $x = 2m_t/\sqrt{s} = 0.05$, meaning that the gluon-gluon interactions dominate in top quarks production. The gluon-gluon interactions represent

about 85% of the processes which produce $t\bar{t}$ at the LHC, while the quark anti-quark interactions represent only 15%.

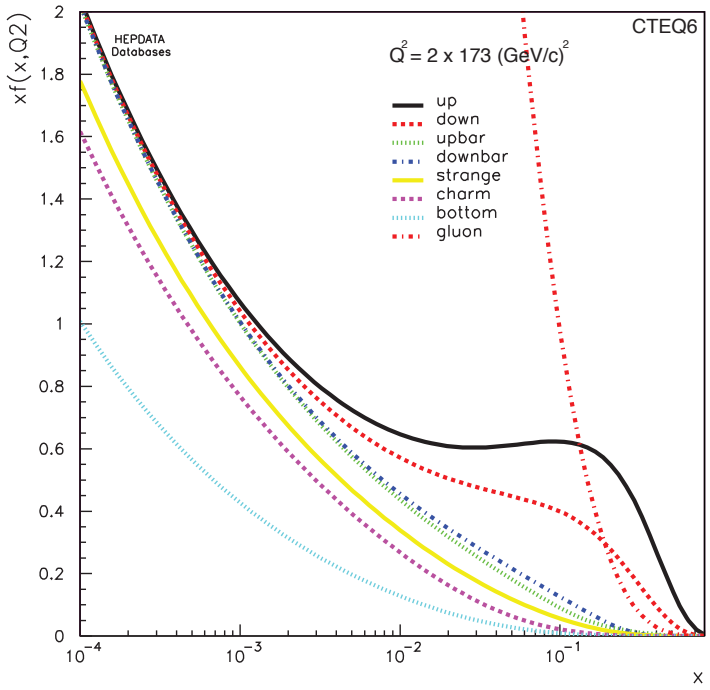

Figure 3. The quark, anti-quark and gluon momentum densities inside the proton as a function of the longitudinal momentum fractions x at $\mu^2 = m_t^2$ for top mass $m_t = 173$ GeV. The parametrization is CTEQ6.

The leading order (LO) processes for producing $t\bar{t}$, contributing with α_s^2 to the perturbation series, are quark-antiquark annihilation $q\bar{q} \to t\bar{t}$, and gluon-gluon fusion $gg \to t\bar{t}$. The corresponding Feynman diagrams for these processes are depicted in Figure 4.

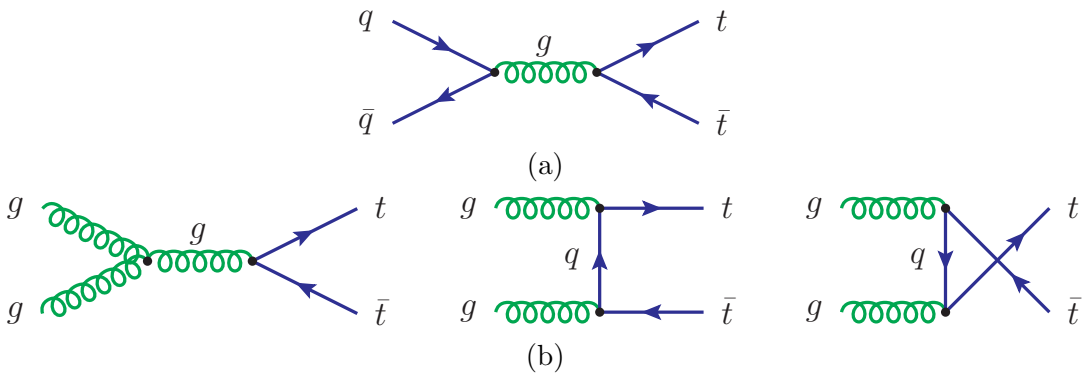

Figure 4. Feynman diagrams of the LO processes for $t\bar{t}$ production: (a) quark-antiquark annihilation $q\bar{q} \to t\bar{t}$ and (b) gluon-gluon fusion, $gg \to t\bar{t}$.

The next-to-leading order (NLO) calculation of the $t\bar{t}$ production cross section involves virtual contributions to the LO processes, gluon bremsstrahlung processes and flavor excitation (11). The predicted cross section for top quark pair production at NLO is $\sigma_{t\bar{t}} = 157^{+23}_{-24}$ pb (12). Deviations of the measured cross section from the theoretical prediction could indicate effects beyond QCD perturbation theory. Explanations might include substantial non-perturbative effects, new production mechanisms or additional top decay modes Beyond the Standard Model (BSM).

2.2.2 Top Quark Production via the Weak Interaction

Single top quarks can be produced via the electroweak interaction involving the Wtb vertex. There are three production modes which are distinguished by the virtuality of the W boson and their corresponding Feynman diagrams are shown in Figure 5.

The single top production modes are clasified as follows:

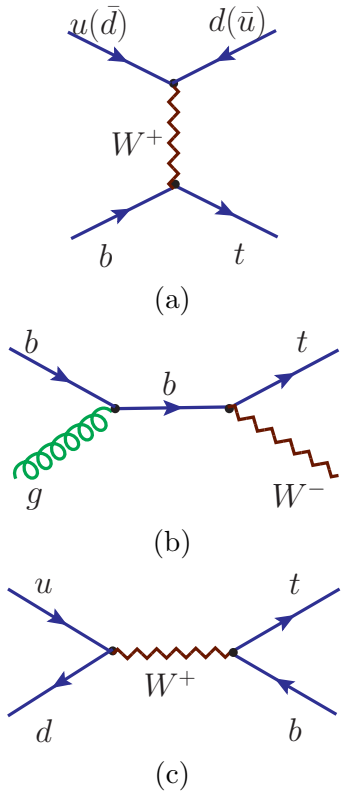

Figure 5. Representative Feynman diagrams for the three single top quark production modes: (a) t-channel, (b) associated production (c) s-channel. The diagrams for single anti-top quark production can be obtained by interchanging quarks and anti-quarks.

- **t-channel** - a virtual W boson strikes a b-quark inside the proton. Since the b-quark originates from a gluon splitting into a $b\bar{b}$, this mode is also known as Wg fusion. Production in the t-channel is the dominant source of single top quarks at the LHC, the theoretical cross section for this channel is $\sigma_t = 64.6^{+3.4}_{-3.2}$ pb (13).

- **tW-channel** - is also known as associated production. The top quark is produced in association with a real W boson. The initial b quark is a sea quark inside the proton.

The associated production at LHC surpasses the s channel, the single top production cross section for this channel was predicted to be $\sigma_{tW} = 10.6 \pm 0.8$ pb (14).

- **s-channel** - this production mode is of Drell-Yan type and it is also known as $t\bar{b}$ production. A W boson is produced by the fusion of two quarks belonging to an SU(2) isospin doublet. Production in the s-channel is pretty small at the LHC, $\sigma_s = 4.2 \pm 0.2$ pb (15).

2.3 Top Quark Decay

Within the SM top quarks decay via the weak interaction to a W boson and a quark q, where q = d, s, b. The rate of the decay is proportional to the CKM matrix element $\mid V_{tq} \mid^2$. The decay $t \to W^+b$ and its charged conjugate occurs almost 100% of the time, with $\mid V_{tb} \mid = 0.999100^{+0.000034}_{-0.000004}$ (8).

The $t\bar{t}$ decay channels are classified according to the decay of the two W bosons present in the $t\bar{t}$ pair decay. The decay modes for the W boson are shown in Table IV (3) together with their corresponding branching ratios (BR). At Born level, all three leptonic W decay modes have the same probability, but each of the two hadronic modes is three times more likely due to the color factor of three.

As shown in Figure 6, the $t\bar{t}$ decay is separated into the following final states:

- **dilepton** - both W bosons from the $t\bar{t}$ process decay leptonically (electron or muon) and the corresponding neutrinos. Due to the presence of the two leptons, this channel has the smallest background.

Decay Mode	**BR at Born level**	**BR**
$W^+ \to e^+\nu_e$	1/9	$(10.75 \pm 0.13)\%$
$W^+ \to \mu^+\nu_\mu$	1/9	$(10.57 \pm 0.15)\%$
$W^+ \to \tau^+\nu_\tau$	1/9	$(11.25 \pm 0.20)\%$
$W^+ \to u\bar{d}, c\bar{s}$	$2 \cdot 3 \cdot 1/9 = 6/9$	$(67.60 \pm 0.27)\%$

TABLE IV

BRANCHING RATIOS AT BORN LEVEL FOR ALL DECAY MODES OF THE W^+ BOSON. THE BR VALUES FOR THE DECAY MODES OF THE W^- BOSON ARE IDENTICAL.

- **lepton+jets** - is also known as semileptonic channel. In this case, one W boson decays hadronically, while the other one decays leptonically into an electron or a muon. This channel provides the best combination of large statistics and a clear signature due to the presence of a lepton.
- **all hadronic** - both W bosons from the $t\bar{t}$ process decay hadronically. Experimentally, the channel suffers from large backgrounds from multijet production.

The W boson can also decay into a τ and a neutrino. The τ can decay further leptonically (electron or muon) or hadronically. The BRs for τ^- lepton leptonically decays are shown in Table V (3). The decaying characteristic of the τ lepton has a direct impact on the BRs of the $t\bar{t}$ decay channels. As an example, a $t\bar{t}$ process in which at least one of the W bosons decay to τ which further decays leptonically would be included in: dilepton

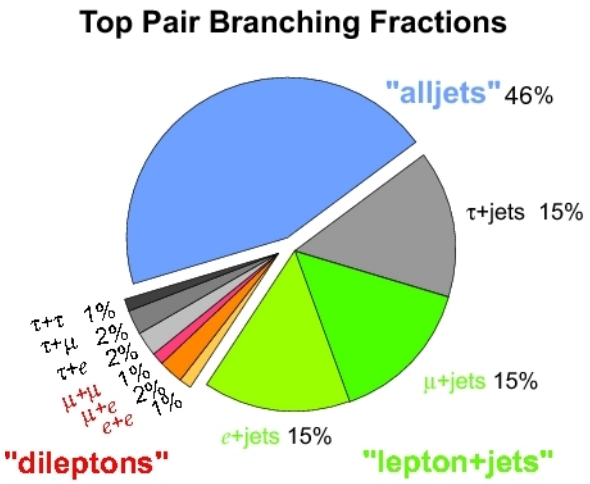

Figure 6. $t\bar{t}$ decay channles and the corresponding BR, according to the theoretical predictions at tree level.

channel of the $t\bar{t}$ decay, if the second W boson decays leptonically, or lepton+jets channel if the second W boson decays hadronically.

The BRs for the $t\bar{t}$ decay channels are shown in Table VI. The BR for the processes with at least one W boson decaying to a τ is quoted separately.

For the measurement performed in this thesis, we must determine the correct BR for the $t\bar{t} \to \mu + \text{jets}$ channel. For this channel, we must include the processes with W boson decaying to τ which decays further to $\mu^{-}\bar{\nu_{\mu}}\nu_{\tau}$, taking into account that the BR for $t\bar{t} \to \tau + \text{jets}$ is $14.52 \pm 0.09\%$ and the BR for the τ decay to muon is 17.36 ± 0.06 as shown in Table V. Therefore, the corrected BR for the $t\bar{t} \to \mu + \text{jets}$ decay channel is $\text{BR}_{t\bar{t}\to\mu+\text{jets}} = 17.04 \pm 0.11\%$.

τ^- **decay mode**	**BR**
$\tau^- \to e^- \bar{\nu_e} \nu_\tau$	$(17.84 \pm 0.06)\%$
$\tau^- \to \mu^- \bar{\nu_\mu} \nu_\tau$	$(17.36 \pm 0.06)\%$

TABLE V

BR VALUES FOR LEPTONICALLY DECAY MODES OF τ^-. THE BR VALUES FOR THE DECAY MODES OF τ^+ ARE IDENTICAL.

$t\bar{t}$ **decay channels**	$t\bar{t}$ **decay mode**	**BR**
ee	$t\bar{t} \to e^+ \nu_e e^- \bar{\nu_e} b\bar{b}$	$(1.14 \pm 0.02)\%$
$e\mu$	$t\bar{t} \to \mu^+ \nu_\mu \mu^- \bar{\nu_\mu} b\bar{b}$	$(1.14 \pm 0.02)\%$
$\mu\mu$	$t\bar{t} \to e^+ \nu_e \mu^- \bar{\nu_\mu} b\bar{b}$	$(2.28 \pm 0.04)\%$
e+jets	$t\bar{t} \to e^+ \nu_e q\bar{q'} b\bar{b}$	$(14.52 \pm 0.09)\%$
μ+jets	$t\bar{t} \to \mu^+ \nu_\mu q\bar{q'} b\bar{b}$	$(14.52 \pm 0.09)\%$
all jets	$t\bar{t} \to q\bar{q'} q\bar{q'} b\bar{b}$	$(46.19 \pm 0.46)\%$
	$t\bar{t} \to \tau$ final states	$(20.21 \pm 0.13)\%$

TABLE VI

$t\bar{t}$ DECAY CHANNELS AND THEIR BRs. THE BR FOR THE PROCESSES WITH AT LEAST ONE W BOSON DECAYING TO A τ IS QUOTED SEPARATELY.

2.4 The signature of the $t\bar{t}$ processes in the μ+jets channel

The $t\bar{t}$ decay channel considered for the measurement presented in this thesis is the the μ+jets channel. The signature of $t\bar{t}$ events in the muon+jets channel can be described as follows:

- One isolated muon originating from the decay of one of the W bosons. The muon has a large transverse momentum.
- One neutrino originating from the W boson. The neutrino is reconstructed as large missing transverse energy.
- Two b jets, as a result of hadronization of the b quarks.
- Two light jets form the hadronization of light flavor quarks (u, d, s, c) originating from the hadronically decaying W boson.
- Additional jets due to initial (ISR) and final state radiation (FSR).

A Feynman diagram of the $t\bar{t}$ process in the muon+jets t-channel is shown in Figure 7.

2.5 Backgrounds to the $t\bar{t}$ process

The objects present in the final state of the μ+jets $t\bar{t}$ events can also appear in the final states of other processes which represent the $t\bar{t}$ background. The sources of background are separated into instrumental background (QCD multijet production) and physics backgrounds (electroweak W boson production in association with jets, $Z/\gamma^* \rightarrow ll$

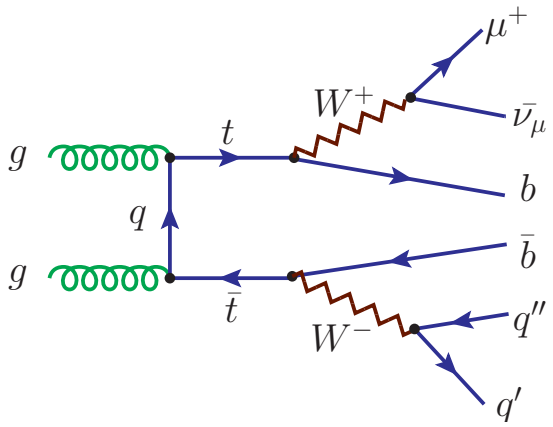

Figure 7. Example of Feynman diagram of $t\bar{t}$ production in the t-channel. The objects in the final state are also shown.

production in association with jets, single top quark production and diboson production). The instrumental backgrounds give rise to muons that appear isolated even though they did not originate from the decay of a W or a Z boson.

2.5.1 QCD multijet production

At the LHC, the multijet processes have the largest production cross section. The production mechanism for additional jets in the QCD events is the QCD bremsstrahlung, i.e. the gluon emission and gluon splitting to $q\bar{q}$. The cross section for QCD multijets production decreases with each additional jet due to the value of $\alpha_s < 1$ associated to each vertex.

The b or c quarks decay into a charged muon and a neutrino and the BR corresponding to this process is about 20%. The muon is usually found to be part of the structure of

the heavy flavor jet, in contrast to the $t\bar{t}$ processes, where the muon is rather isolated. In some cases though, the objects within the jet structure are too soft which leads to the misreconstruction of the jet. In this case, the muon originating from the semi-leptonic heavy quark decay can appear to be isolated. In addition, the neutrino originating from the heavy quark decay together with the misrecontruction of the jet will result in a significant missing transverse energy in the event. The number of QCD multijet processes interpreted as $t\bar{t}$ is quite significant, representing the second most important background for the $t\bar{t}$ production, after applying the full selection.

2.5.2 Electroweak W boson production in association with jets

At the LHC, the W boson production proceeds via the Drell-Yan process. There are three processes in the W+jets sample: W+bb, when one or two b-jets are produced along with the W boson, W+cc, when one or two jets are c-jets, and W+light jets, when the jets produced are neither b nor c-jets. At NLO, the relative contributions of these samples are different from the LO calculation, in particular W+bb and W+cc contributions are expected to be higher. The W+light jets sample is considerably reduced by using b-tagging techniques, but a large part of the W+bb sample is selected together with the $t\bar{t}$ sample, due to the presence of the b-jets in the final state.

The W boson production in association with jets is found to be the dominant background to the $t\bar{t}$ production in the muon+jets channel.

2.5.3 $Z/\gamma^* \to ll$ production in association with jets

The production $Z/\gamma^* \to ll$ in association with jets proceeds also via the Drell-Yan process, as the electroweak W boson production, but the production rate for $Z/\gamma^* \to ll$ is ten times smaller than the corresponding W production rate. Contrary to the W +jets though, the final state of the $Z/\gamma^* \to ll$ events involves two leptons and no neutrino. Therefore, the presence of $Z/\gamma^* \to ll$ events is significantly suppressed by requiring that only one muon is present in the event. In a few cases one of the two muons is not detected due to the limited detector acceptance and then the final states will result to be same as the ones of the $t\bar{t}$ processes.

The contribution of the $Z/\gamma^* \to ll$ production to the background of the $t\bar{t}$ production is small, due to the high efficiency of the CMS detector in identifying the muons.

2.5.4 Single top quark production

The process of producing the single top quark is discussed in Section 2.2.2. Single top events can have the same detector signature as the muon+jets final state of the $t\bar{t}$ decay. due to ISR and FSR which can produce additional jets. The overall contribution of the single top production to the $t\bar{t}$ background is rather small.

2.5.5 Diboson production

Another background for the $t\bar{t}$ process is the production of a pair of vector bosons. Since two vector bosons are produced, their further decays will result in a final state

similar to the $t\bar{t}$ events. Additional jets can be produced by ISR and FSR. The diboson background to the $t\bar{t}$ production is small.

CHAPTER 3

LHC AND THE CMS DETECTOR

3.1 The Large Hadron Collider (LHC)

The LHC (16) is a two-ring superconducting hadron accelerator and collider. It is located on the border between Switzerland and France where it was installed in the already existing 26.7 km circumference tunnel constructed for the CERN LEP machine. The LHC was designed to collide opposing particle beams of protons with an energy up to 7 TeV per beam, but it also offers the possibility of colliding lead nuclei (Pb^{82}_{208}) with an energy of up to 574 TeV per nucleus. The LHC was built by CERN in collaboration with over 10,000 scientists and engineers from hundreds of universities and laboratories residing in over 100 countries.

The proton beams were circulated for the first time in September 2008. The physics program started in 2010 at half of the nominal energy, 3.5 TeV per particle, and it will continue until the end of 2012, followed by an upgrade to reach the nominal energy of 7 TeV per proton. Short periods of heavy ions collisions are scheduled between pp runs.

3.1.1 LHC Layout

The basic layout of the LHC is presented in Figure 8. The LHC tunnel has eight sections and lies between 45 m and 170 m below the surface. Each straight section is approximately 528 m long and can serve as an experiment or utility insertion.

In Figure 8, the sections are denoted **Octants**, and a **Point** corresponds to each octant representing the locations of different experiments. Point 1 and Point 5 represent the locations for two high luminosity (peak luminosity L = 10^{34} cm^2s) experiments at the LHC: ATLAS (17) and CMS (18), respectively, which are both general purpose detectors designed for proton-proton collisions. The injection system for Beam 1 and the ALICE (19) experiment, designed to study high energy nucleus-nucleus collisions, are located at Point 2. Point 8 hosts the LHCb (20) experiment designed to study the physics of b-quarks, and the injection system for Beam 2. Both ALICE and LHCb are lower luminosity experiments: peak luminosity L = 10^{32} cm^2s. The straight sections located at Point 3 and Point 7 are equipped with collimation systems, while Point 4 holds two radio frequency systems needed to capture, accelerate and store the two beams. The beam dumping system is located at Point 6. In addition, two small experiments were built at LHC: TOTEM (21), which is located next to CMS at Point 5 and LHCf (22), which is located next to ATLAS at Point 1, and their goal is to detect "forward" particles that do not participate in collisions.

The LHC relies on superconducting magnets that are at the edge of present technology. There are separate magnetic fields and vacuum chambers in the main arcs, while the common sections are located only at the insertion regions. The magnets are cooled to a temperature below 2 K using superfluid helium and operate at fields above 8 T. The insulation vacuum for cryomagnets is 10^{-6} mbar, while the hydrogen density must be

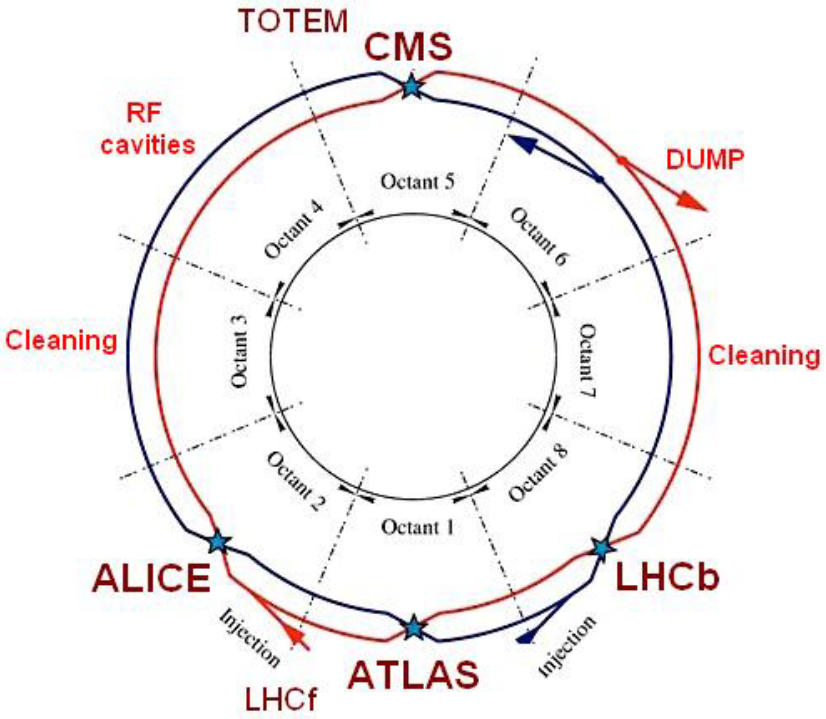

Figure 8. The structure of LHC.

10^{-15} H_2m^{-3} inside the beam pipes and 10^{-13} H_2m^{-3} at the interaction points in order to reduce the beam-gas interactions.

3.1.2 Protons Travel Path

The LHC is linked to the CERN accelerator complex (shown in Figure 9) through two transfer tunnels, each about 2.5 km in length.

The protons are obtained from injecting hydrogen gas into a metal cylinder (Duoplasmatron) surrounded by an electric field to break down the gas into its constituents (protons and electrons): $H_2 \rightarrow 2H^+ + 2e^-$. This process yields about 70% protons.

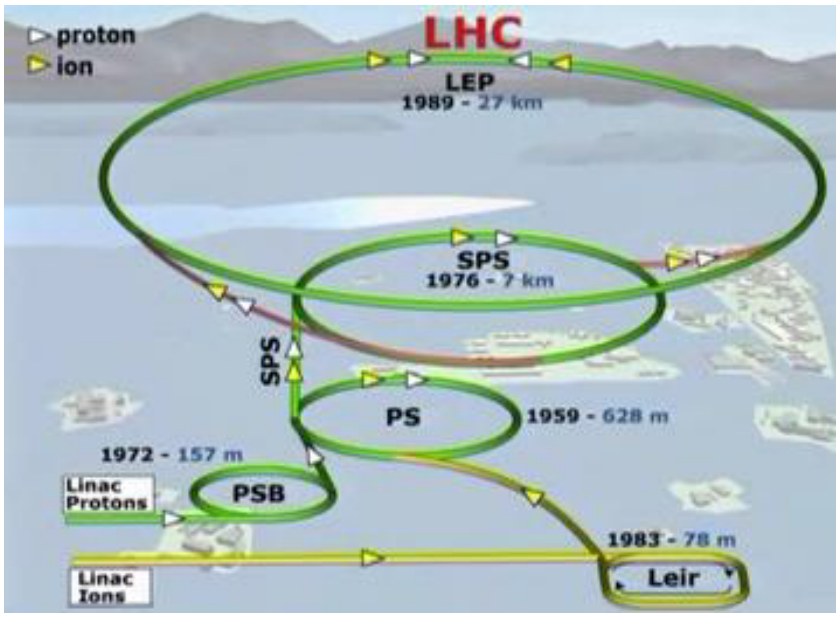

Figure 9. The LHC accelerator complex at CERN.

The protons are accelerated by a 90 kV supply and then sent to a Radio Frequency Quadrupole (QRF), which both speeds up and focuses the particle beam. The bunches of protons accelerated up to 750 keV are sent to the linear accelerator LINAC2. Within LINAC2, the particles are accelerated up to 50 MeV and then sent to the Proton Synchrotron Booster (PSB). The PSB accelerates the beam to 1.4 GeV in 530 m and injects it into the 628 m circumference circular accelerator, denoted the Proton Synchrotron (PS). The PS is responsible for accelerating the beam to 25 GeV and for providing 81 bunch packets with 25 ns spacing for LHC. The bunches formed in the PS are injected into the 7 km circumference circular accelerator, denoted Super Proton Synchrotron (SPS). The

SPS accelerates the protons to 450 GeV in 4.3 seconds and transfers them to the LHC in both clockwise and anticlockwise direction. The transit time for the bunches injected into the LHC is about 45 minutes and includes the 450 GeV injection plateau and the ramping to the high energy. When the beam lifetime (about 10 hours) is reached, the beam itself is exhausted and dumped. The LHC machine is designed to withstand about 20,000 such cycles in 20 years lifetime.

3.1.3 LHC Luminosity

The luminosity is a crucial parameter when performing precise physics measurements and during the commissioning process. The luminosity represents the number of particles that cross a given area in a given time interval. The luminosity at the LHC is not constant over a physics run, but decays due to the degradation of intensities and emittances of the circulating beams. The instantaneous luminosity is given by Equation 3.1:

$$\mathcal{L}_0 = \frac{N_p^2 N_b f \gamma F}{4\pi \epsilon_T \beta}, \tag{3.1}$$

where the parameters represent:

- N_p - is the number of particles per bunch
- N_b - is the number of bunches per beam
- f - is the revolution frequency
- γ - is the relativistic gamma factor

- F - is the geometric luminosity reduction factor, which in case of a circular beam is calculated as $F = (1 + (\frac{\theta\sigma_z}{2\sigma_T})^2)^{-1/2}$, where θ is the crossing angle at the interaction point, σ_z is the RMS bunch length, and σ_T is the transverse RMS beam size at the interaction point
- ϵ_T - is the normalized transverse beam emittance
- β - is the betatron function at the collision point.

The profile of the beam in the transverse plane can be described by an ellipse of constant area $\pi\epsilon_T$ and variable shape. The two semi-axes are defined as $\sqrt{\epsilon_T/\beta}$ and $\sqrt{\epsilon_T\beta}$. The nominal beam parameters for proton-proton collisions are listed in Table VII.

Beam Parameters	Protons
Injection energy per nucleon [GeV]	450
Collisions energy per nucleon [TeV]	7
Number of particles per bunch	1.1×10^{11}
Number of bunches	2808
Bunch separation [ns]	25
β	0.55
ϵ_T [μm]	3.75
Instantaneous luminosity [$\mathrm{cm}^{-2}\mathrm{s}^{-1}$]	10^{34}

TABLE VII

NOMINAL PARAMETERS FOR THE LHC BEAM FOR PROTON-PROTON COLLISIONS.

The integrated luminosity over the run is given by Equation 3.2:

$$L = \mathcal{L}_0 \tau [1 - \exp^{-T/\tau}], \tag{3.2}$$

where $\mathcal{L}_0$ is the initial instantaneous luminosity, τ is the luminosity lifetime and T is the duration of the run.

3.2 The Compact Muon Solenoid (CMS)

The CMS detector is a multi-purpose apparatus built at the LHC to investigate a wide range of physics. It is installed 100 m underground close to the French village of Cessy. The detector requirements for CMS to meet the goals of the LHC physics program can be summarized as follows:

- good muon identification and momentum resolution
- good charged particle momentum resolution and reconstruction efficiency in the inner tracker
- good electromagnetic energy resolution
- good missing transverse energy and dijet mass resolution

The CMS layout is shown in Figure 10. It is 21 m long, 15 m wide and 15 m high and weights $12,500$ tonnes. The main distinguishing features of the CMS detector are a high field solenoid, a silicon based inner tracking system, and an homogeneous scintillating crystals-based electromagnetic calorimeter.

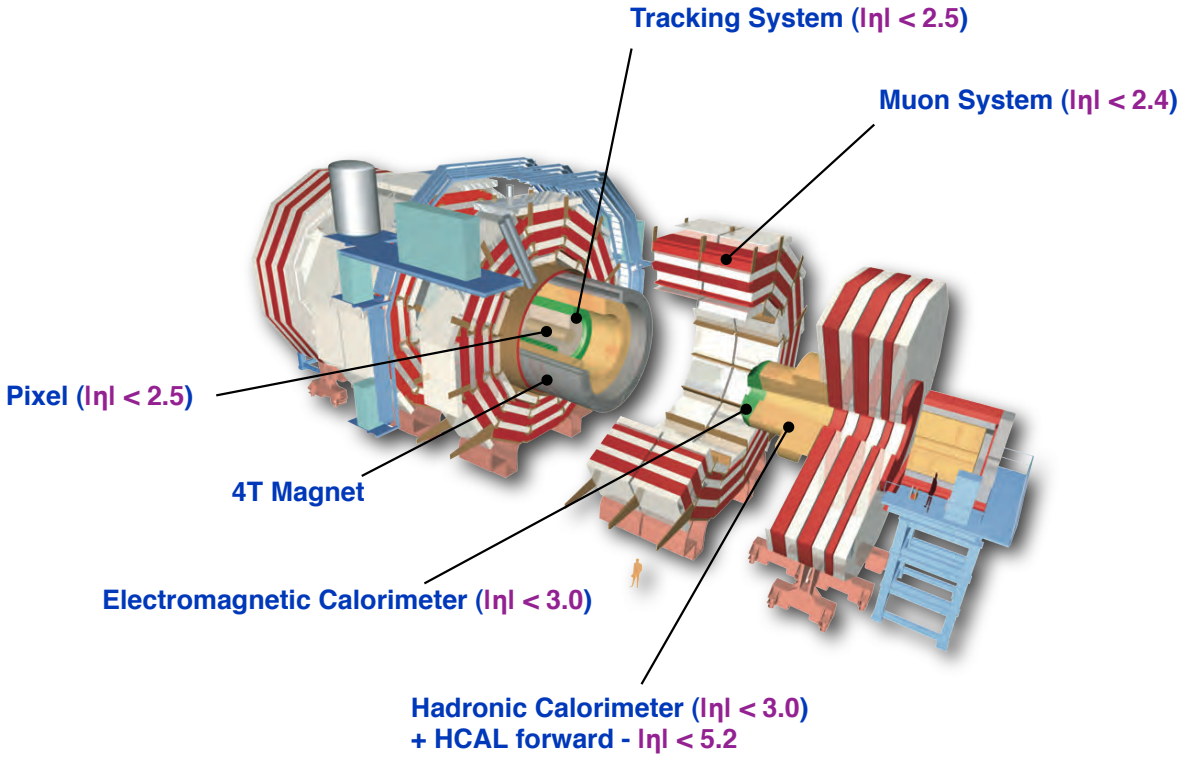

Figure 10. Schematic view of the CMS detector.

The system of coordinates adopted by CMS has the origin centered at the nominal collision point inside the experiment: the y-axis pointing upward, the x-axis pointing radially inward towards the center of the LHC, and the z axis points along the beam direction counter-clockwise. The azimuthal angle ϕ is measured from the x-axis in the

x-y plane and the radial coordinate in this plane is denoted by r. The polar angle θ is measured from the z-axis. The pseudorapidity is defined as

$$\eta = -ln(tan(\frac{\theta}{2})). \tag{3.3}$$

As shown in Figure 12, the particles emerging from collisions first meet the tracker. The tracks of the charged particles are bent due to the magnetic field and this allows the measurement of their momentum. Outside the tracker are the calorimeters that measure the energy of the particles: the Electromagnetic Calorimeter (ECAL) measures the energy for photons and electrons, while the Hadronic Calorimeter (HCAL) is designed to detect the hadrons. The muon system is designed to identify muons.

The CMS superconducting magnet (23) was designed to reach a 4 T magnetic field inside a free bore with the purpose of bending the paths of the charged particles. The bore has 6 m diameter and 12.5 m length, being large enough to accommodate the tracker and the calorimeter. The coil is composed of 4 layers of superconducting NbTi windings. The critical temperature of the superconductor in a 4 T magnetic field is 4.5 K. Currently, the magnet is operating at 3.8 T. The magnetic flux is returned through a yoke which interleaves with the muon detector. The yoke is composed of 5 wheels in the central region and 2 end-cap disks. Both the magnet coil and the yoke are shown in Figure 11.

At the design luminosity ($\sqrt{s} = 14$ TeV), there will be 10^9 inelastic events detected by CMS, and 20 inelastic collisions superimposed on the event of interest. This implies

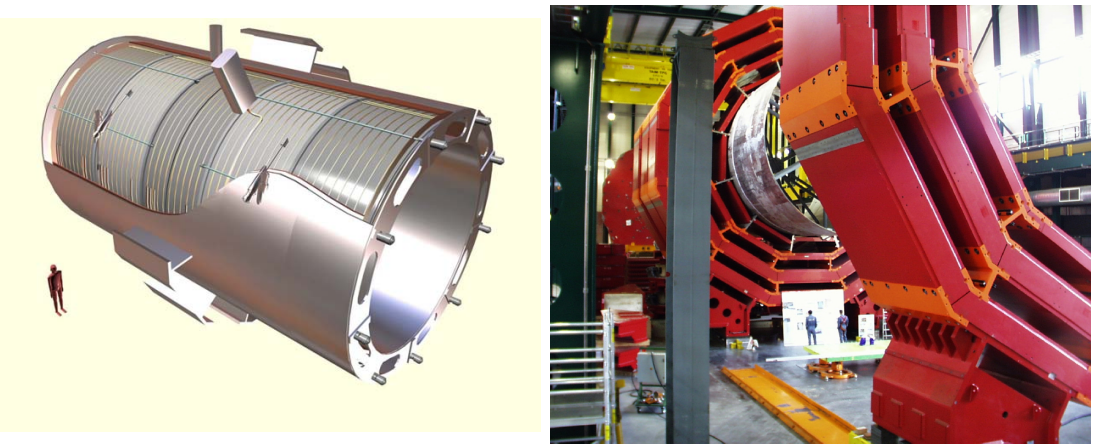

Figure 11. Coil (left) and yoke (right) of the CMS magnet.

that about 1000 charged particles will emerge from the interaction point for each proton bunch crossing.

Each of the CMS components will be described briefly in the following subsections. In addition to the described detectors, CMS also includes two very forward calorimeter systems (CASTOR and ZDC), situated outside the muon system. They are mainly used in the heavy ions and diffractive physics programs, a detailed description of them can be found in (18).

3.2.1 Pixel Detector

The Pixel system is the detector closest to the interaction region. It consists of three barrel layers and two end-cap disks shown in Figure 13.

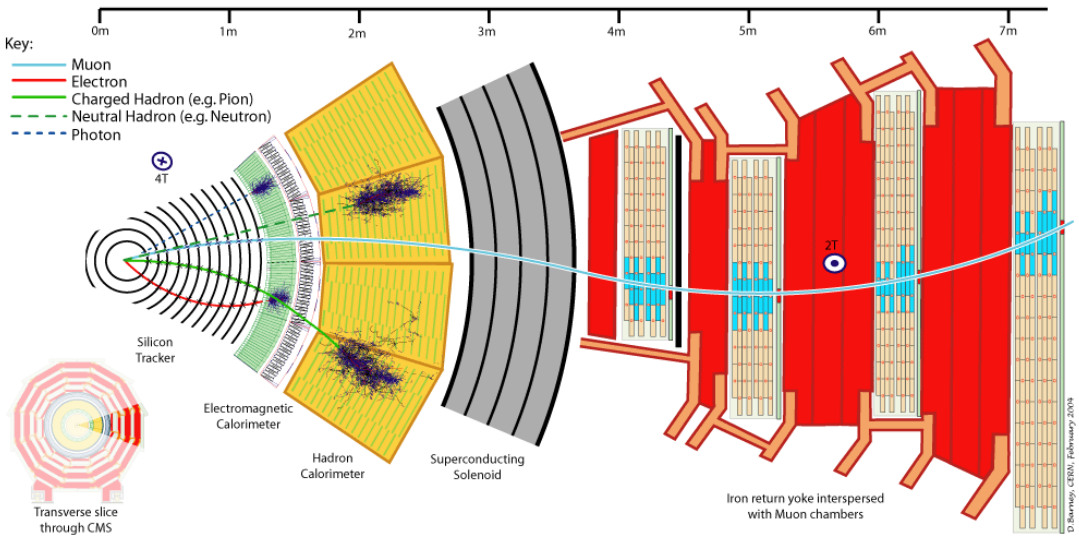

Figure 12. Slice through the CMS detector showing different particles incident on the CMS subdetectors.

The barrel layers have a length of 53 cm and are located at mean radii of 4.4 cm, 7.3 cm and 10.2 cm. The end-cap disks extend from 6 cm to 15 cm in radius and are placed on each side of the barrel at $z = \pm 34.5$ cm and $z = \pm 46.5$ cm. There are 66 million pixels included in 1440 pixel modules covering a total area of about 1.06 m^2. Figure 14 shows the geometrical layout of the Pixel Detector including the covered pseudorapidity range $(-2.5 < \eta < 2.5)$.

The Pixel Detector is the CMS subdetector most exposed to radiation. The hit density rate for each bunch crossing is 1 MHz/mm^2 for the first layer of the Pixel Detector. At full LHC luminosity, it is expected that the innermost layer of the Pixel Detector will

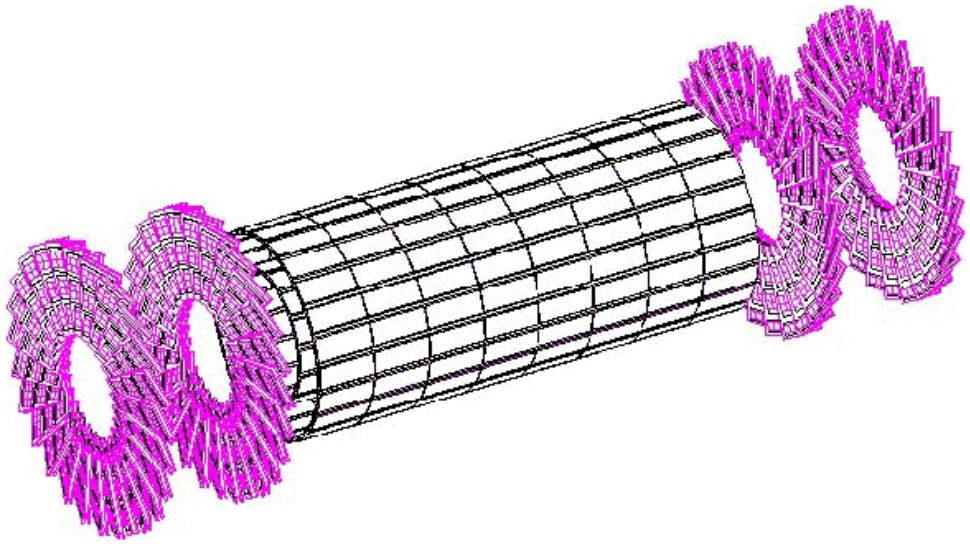

Figure 13. A sketch of the CMS Pixel detector showing both barrel and end-cap disks.

stay operational for at least 2 years while the outermost layer should be functional for more than 10 years.

The CMS Pixel Detector uses the n pixel on n substrate sensor design, that allows partial depleted operation even at very high particle flux. The pixel cell size is 100×150 μm^2 achieving a similar spacial resolution of 15-20 μm in both r $-$ ϕ and z directions. This characteristic allows a 3D vertex reconstruction, which is very important for identifying secondary vertices with low track multiplicity. The sensor signals are read by a readout chip bonded to the sensors. The operating mode is the zero suppression, meaning that only signals above a certain threshold are read out. In order to maintain a proper oper-

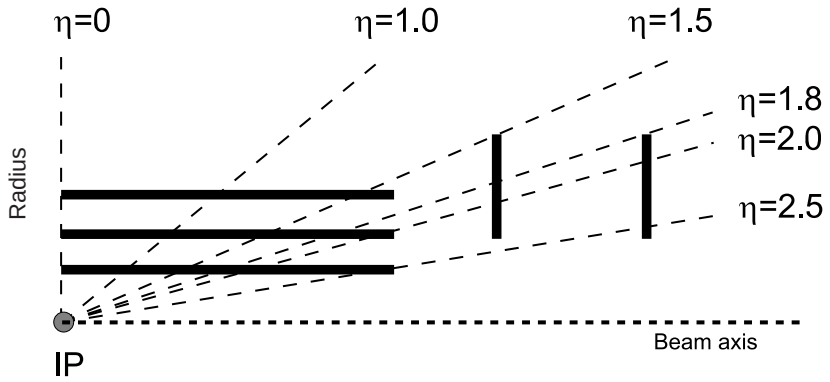

Figure 14. Geometrical layout of the pixel detector.

ation of the silicon sensors and reduce the radiation damage, the entire pixel volume is currently kept at a temperature of -20^o C.

3.2.2 Silicon Strip Tracker Detector

The Silicon Strip Tracker (SST) (shown in Figure 15) occupies the radial region between 20 cm and 116 cm within the CMS detector. It has a total of 9.3 million strips representing 198 m^2 of active silicon area. The acceptance of the SST is $\mid \eta \mid \leq 2.5$.

The SST is composed of the following parts: Tracker Inner Barrel (TIB), Tracker Inner Disks (TID), Tracker Outer Barrel (TOB) and Tracker Endcaps (TEC). A schematic drawing of the SST layout is shown in Figure 16 (24).

The sensor elements in the SST are single sided p-on-n silicon micro-strip sensors. They have been manufactured on 6 inch wafers using a standard planar process (25). In order to equip all regions in the CMS tracker, 15 different sensor geometries are needed,

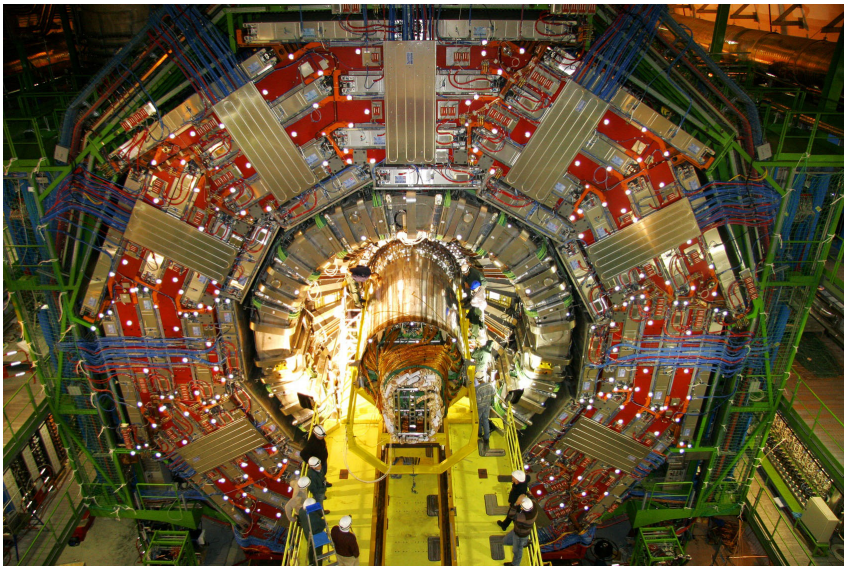

Figure 15. Insertion of the Silicon Strip Tracker into the CMS detector inside the cavern at Point 5.

a few of them being represented in Figure 17. The sensors have either 512 or 768 strips, reflecting the read-out modularity of 256 channels: two 128-channels front-end chips are connected to the same optical fiber.

The signals from the silicon sensors are amplified, shaped and stored by a custom integrated circuit, the APV25. Upon a positive first level trigger decision, the analog signals of all channels are multiplexed and transmitted via optical fibers to Front End Driver (FED) boards in the service cavern where the analog to digital conversion takes place. In this way an accurate pedestal and common mode subtraction is achieved.

The characteristics of the SST subsystems are the following:

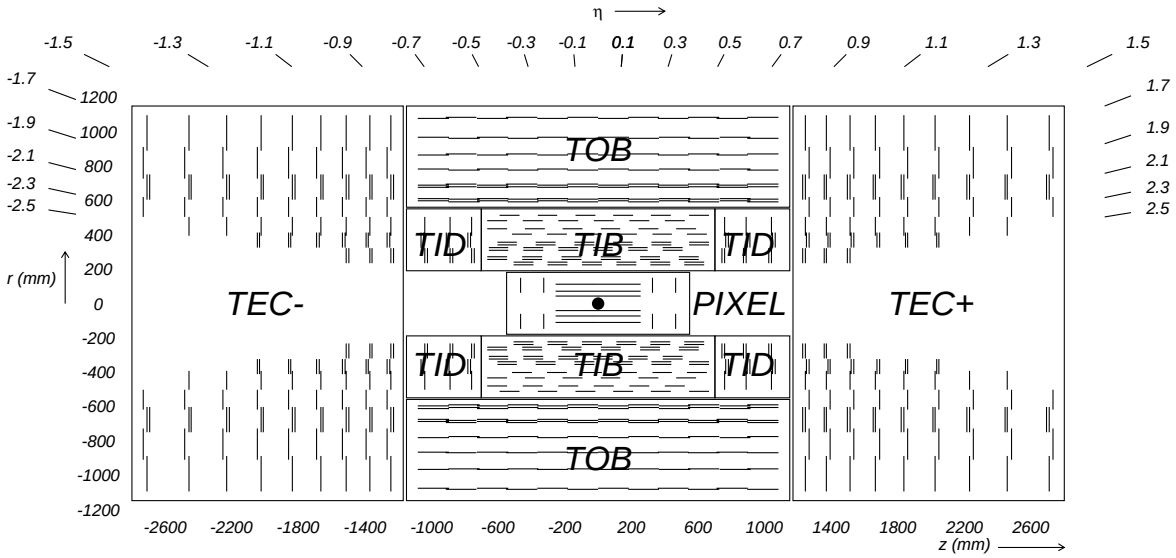

Figure 16. Geometrical layout of the pixel detector.

- **TIB** - is composed of 4 barrel layers and together with TID extends to the radius of 55 cm. The silicon sensors used have a thickness of 320 μm, with their strips parallel to the beam axis. The strip pitch is 80 μm on layers 1 and 2 and 120 μm on layers 3 and 4, leading to a single point resolution of 20-35 μm for the r-ϕ measurements.

- **TID** - is composed of 3 disks at each end of the TIB. The silicon sensors have the same thickness as the ones used for the TIB, 320 μm. The strips are arranged in the radial direction, and the pitch varies between 100 μm and 141 μm, providing a single point resolution of 30-40 μm for the z-ϕ measurements.

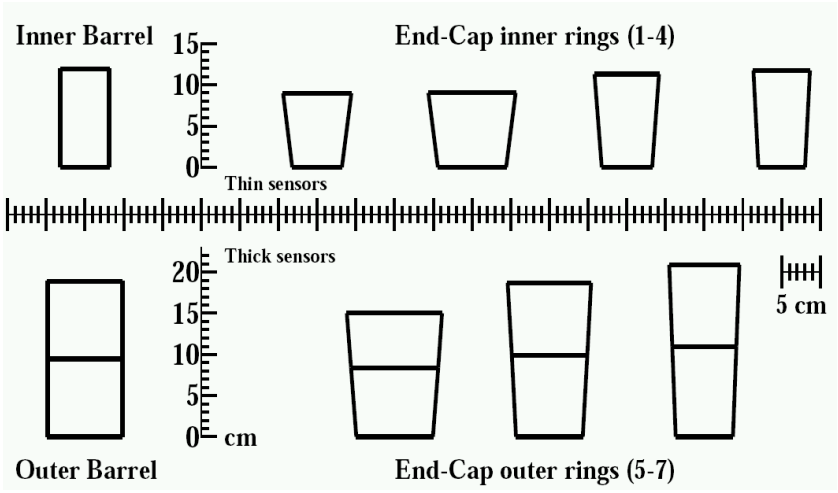

Figure 17. Silicon sensor geometries utilized in the CMS tracker. In the outer layers, the sensors are paired to form a single module.

- **TOB** - is composed of 6 barrel layers. It surrounds the TIB and TID and it extends in z to ± 118 cm and in r to 116 cm. The sensors used for TOB have a thickness of 500 μm and a bigger length, in order to keep the signal-to-noise ratio above 10 for 10 years of operation. The pitch varies from 183 μm in the first four layers to 122 μm on the outermost two layers, and the single point resolution varies between 35-53 μm for the r-ϕ measurements.

- **TEC** - is composed of two parts, denoted TEC+ and TEC-, depending on which side of the interaction point it is located, each TEC being composed of 9 disks. The disks carry up to 7 rings of silicon micro-strip detectors with different thicknesses: 320 μm on the inner 4 rings, and 500 μm on rings 5-7. The strips are arranged in

the radial direction, and the pitch varies between 97 μm and 184 μm, leading to a point resolution of 50-60 μm for the z-ϕ measurements.

As shown in Figure 16, the modules in the first two layers of TIB and TOB, the first two rings of the TID and the rings 1,2 and 5 of the TECs carry a second micro-strip detector module which is mounted back-to-back with a stereo angle of 100 mrad in order to provide a measurement of the third coordinate: z in the barrel and r on the disks. The point resolution is not as good as for the other two coordinates (about 500 μm in the endcaps).

After the Pixel detector, the SST receives the largest amount of radiation. The hit density rate for each bunch crossing is 60 kHz/mm^2 for the first TIB layer and 3 kHz/mm^2 for the outermost TOB layer. In order to maintain a proper operation, the entire tracker detector is cooled to a temperature of -20^o C, using a C_6F_{14} cooling fluid. It is expected that the SST will remain fully operational for 10 years, at a temperature of about -27^o C.

Studies of the performance of the SST involved the percent of noisy/dead strips and the signal-to-noise ratio. The percent of strips flagged as problematic (noisy or dead) is below 1% for all the SST subdetectors. The signal-to-noise is a benchmark for the performance of the SST. The signal-to-noise is about 25-30 for thin modules and 31-36 for the thick ones. The thick sensors collect a factor of 5/3 more charge than the thin sensors, but this does not simply scale up the signal-to-noise ratio, as the noise increases

with the increase of the sensors lengths. The signal-to-noise proved to be stable within 5% over time.

3.2.3 Electromagnetic Calorimeter

The radiation loss of a charged particle with energy E can be described as: $-\frac{dE}{dx} = \frac{E}{X_0}$, where X_0 represents the radiation length and its value is material dependent (26). The electrons emit photons through bremsstrahlung, losing energy until a critical value E_c for the energy is reached. Below E_c, the dominant interactions are ionization and excitation. An important characteristic of the material is also the Moliere radius, which determines the spread of the particles shower resulting from the bremsstrahlung. For a good electromagnetic calorimeter, it is necessary that it is build of material with short X_0 and small Moliere radius, so that the electrons and photons will deposit their entire energies before reaching the edge of the detector.

The electromagnetic calorimeter is an hermetic homogeneous calorimeter composed of a central Barrel and two Endcaps. A Preshower detector is placed in front of the Endcaps. A layout of the ECAL detector is shown in Figure 18.

The ECAL is made of lead tungstate ($PbWO_4$) crystals, having 61200 crystals in the ECAL Barrel and 7324 crystals in each of the Endcaps. The $PbWO_4$ crystals have a short radiation length (0.89 cm) and a small Moliere radius (2.2 cm), providing a compact calorimeter with fine granularity. The ECAL granularity in $\eta - \phi$ space is 0.0174×0.0174,

while the momentum resolution depends on the momentum of the particle (≈10% for $p_T > 100$ GeV and ≈20% for $p_T < 50$ GeV (27)).

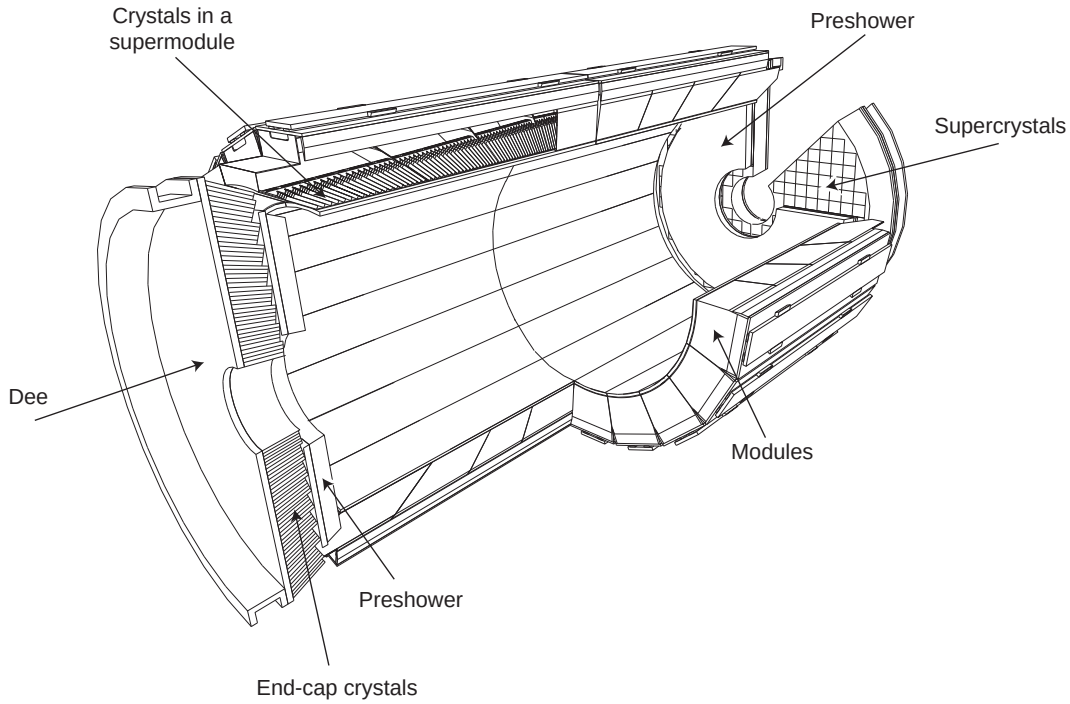

Figure 18. Layout of the CMS ECAL showing the arrangement of crystal modules in the barrel and the supercrystals in the endcaps, with the preshower in front.

The ECAL subdetectors are characterized by the following:

- **ECAL Barrel (EB)** - extends into the CMS detector to a radius of 1.29 m and covers the pseudorapidity range $|\eta| < 1.479$. The crystals have a trapezoidal shape,

with 22×22 mm^2 at the front of the crystal and 26×26 mm^2 at the rear face. The crystal length is 230 mm corresponding to 25.8 X_0. The crystals are assembled into modules of different shapes, each module containing 400 or 500 crystals. Four modules are assembled into a supermodule, containing 1700 crystals. The scintillation light is collected and amplified in the EB by the avalanche photodiodes (APDs) specially developed for CMS ECAL. A pair of APDs is mounted on each crystal.

- **ECAL Endcaps (EE)** - extend to 315.4 cm on the z axis from the interaction point, and cover the pseudorapidity range $1.479 <| \eta |< 3.0$. The EE consist of identically shaped crystals grouped in mechanical units of 5×5 crystals, named supercrystals. Each endcap is divided into two halves, named Dees. The crystals have a front cross section of 28.62×28.62 mm^2, a rear cross section of 30×30 mm^2 and a length of 220 mm corresponding to 24.7 X_0. The scintillation light in the EE is amplified by the vacuum phototriodes (VPTs) equipped with a single gain stage.
- **ECAL Preshower** - is placed in front of the EE. It contains one lead radiator layer and one silicon strip sensor layer. Its purpose is to improve the position resolution for electrons and photons. It has a thickness of 20 cm and a coverage in pseudorapidity of $1.6 <| \eta |< 2.6$. Both layers together provide about 3 X_0.

The number of photons emitted by crystals and the amplification of the APDs and VTPs are temperature dependent. The nominal operating temperature for ECAL is 18^o C.

3.2.4 Hadron Calorimeter

The HCAL (28) is important for the measurement of the hadron jets and the missing transverse energy. The absorption of hadrons by matter as a function of the distance x is described by $N = N_0 e^{-x/\lambda_I}$, where N_0 represents the initial number of hadrons and λ_I is the average interaction length. The interaction length depends on the material and can be calculated as $\lambda_I = \frac{A}{N_A \times \rho \times \sigma_{inel}}$, where A represents the atomic weight, N_A is Avogadro's number, ρ is the density of the material, and σ_{inel} is the inelastic cross section. The interaction length for the CMS HCAL is 16.4 cm.

The hadron calorimeter is composed of different parts, located either inside the solenoid (barrel and endcaps) or outside the magnet coil (outer and forward). The layout of the HCAL is presented in Figure 19.

The HCAL subdetectors are characterized by the following:

- **Hadron Barrel Calorimeter (HB)** - is restricted to the space between the ECAL Barrel and the solenoid. It has a coverage in pseudorapidity of $| \eta | < 1.3$. The HB is made of 16 layers of absorber plates with a variable thickness between 40-75 mm. The absorber plates are made of brass, except the first and the last absorber layers which are made of stainless steel to provide additional structural strength. The absorber plates are interleaved with tiles of plastic scintillation fibers. The HB granularity in $\eta - \phi$ space is 0.087×0.087. The absorbers provide between 5.82 λ_I and 10.6 λ_I, from $\eta = 0$ (where the HB is thinner) to $\eta = 1.3$, respectively. Another

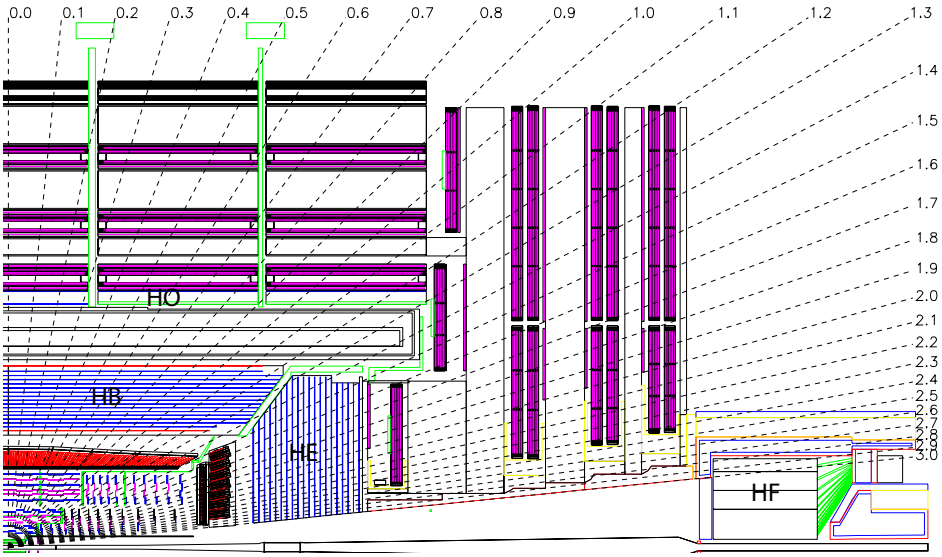

Figure 19. Layout of the CMS HCAL showing the locations of the hadron barrel (HB), endcap (HE), outer (HO), and forward (HF) calorimeter.

λ_I is added to the absorbers due to the ECAL presence in front of the HB. The scintillation light is collected by the wavelength shifting fibers, which are arranged into read-out towers and transport the light to the hybrid photodiode (HPD).

- **Hadron Endcaps Calorimeter (HE)** - covers a substantial portion of rapidity range: $1.3 <| \eta |< 3.0$. It is made of cartridge brass absorber plates of 79 mm thickness interleaved with trays of plastic scintillators. The HE has the same granularity and the same read-out as the HE. The absorber provides 10.6 λ_I.

- **Hadron Outer Calorimeter (HO)** - is placed outside the solenoid in the region of $| \eta |< 1.3$. Its presence is necessary because the EB and HB do not provide sufficient

containment for the hadron showers. It is used to identify the late starting showers. An HO layer is placed in front of each of the 5 rings of iron yoke. At $\eta = 0$ the HO has two layers of plastic scintillators, while all the other rings have only one layer at a radius of 4.1m. The total depth of the calorimeter is thus extended to a minimum of 12 λ_I.

- **Hadron Forward Calorimeter (HF)** - is placed outside the muon endcaps, near the beam pipe at a distance of 11.2 m from the interaction point, being exposed to a very large flux of particles. Its presence extends the pseudorapidity coverage for the HCAL to $\mid \eta \mid = 5$. The active material for the HF are quartz fibers, which emit Cherenkov light when they are traversed by charged particles. The fibers have different lengths, making possible the distinction between the showers generated by hadrons from the showers generated by electrons and photons. The HF granularity in $\eta - \phi$ space is 0.175×0.175 and its depth is 10 λ_I.

The CMS detector uses signals from the HF to determine the instantaneous luminosity in real time using two methods. The first method is based on "zero counting", in which the average fraction of empty towers is used to infer the mean number of interactions per bunch crossing. The second method exploits the linear relationship between the average transverse energy per tower and the luminosity.

Although all HF towers are outfitted with luminosity firmware, the best linearity is obtained by limiting the coverage to four azimuthal (2π) rings in the range $3.5 < \mid \eta \mid < 4.2$.

The main reason for this limitation is to avoid taking into account the η rings which have different probabilities for occupied towers in a single interaction event.

The "zero counting" method is the default online method because it offers a better performance at the relatively low luminosities delivered by the LHC. Additional offline methods were developed at CMS for computing the luminosity, for which the absolute scale for the offline luminosity is done by Van der Meer scans. A detailed description can be found in (29).

3.2.5 Muon System

The Muon System (30) has three functions: identification, momentum measurement and triggering. Due to the shape of the solenoid magnet, the muon system was designed to have a cylindrical shape with a barrel section and two planar endcap regions.

The muon system subdetectors have the following characteristics:

- **Muon Barrel** - consists of barrel drift tube (DT) chambers and offers a coverage of $| \eta | < 1.2$. The DT has a transverse dimension of 21 mm, an anode wire 2.4 m long and is filled with a gas mixture of 15% Ar + 85% CO_2. When a muon traverses the DT, the gas mixture is ionized and the ions (electrons) are accelerated towards the cathode (anode) by an electric field. The accelerated particles will continue ionizing, forming an electromagnetic shower. Each charged particle is collected with a delay corresponding to the distance travelled from its relative position to the collecting

electrode. The distribution of the collected charge density is used to determine the trajectory of the muon inside the DT.

There are 4 muon stations composed of DT chambers intermixed with the iron yoke. The DT chambers are made of 3 or 2 super layers (SL). The 3 inner stations consist of 8 chambers each, 4 chambers provide an r $-$ ϕ measurement and the other 4 provide a z-direction measurement, according to the orientation of the SLs in the chambers. The fourth outer station does not provide a z-direction measurement. The Muon Barrel is shown in Figure 20.

Since the magnetic field in the Muon Barrel is uniform and has a relatively low strength, the DTs are successfully used as tracking detectors. The position resolution in r$-\phi$ for the Muon Barrel is about 100 μm.

- **Muon Endcaps** - are made of trapezoidal cathode strip chambers (CSC) and offer a coverage of $0.9 <| \eta |< 2.4$. The CSC contain 6 planes of wires (gold plated tungsten) arranged on the ϕ direction with spacing of 3.2 mm. The planes of wires are interleaved with 7 plates of cathode strips, arranged along the radial direction with a distance of 0.5 mm between them. The CSC setup allows a precise position measurement of the muons in all three coordinates. The gas used in the CSC is a mixture of 40% Ar + 50% CO_2 + 10% CF_4.

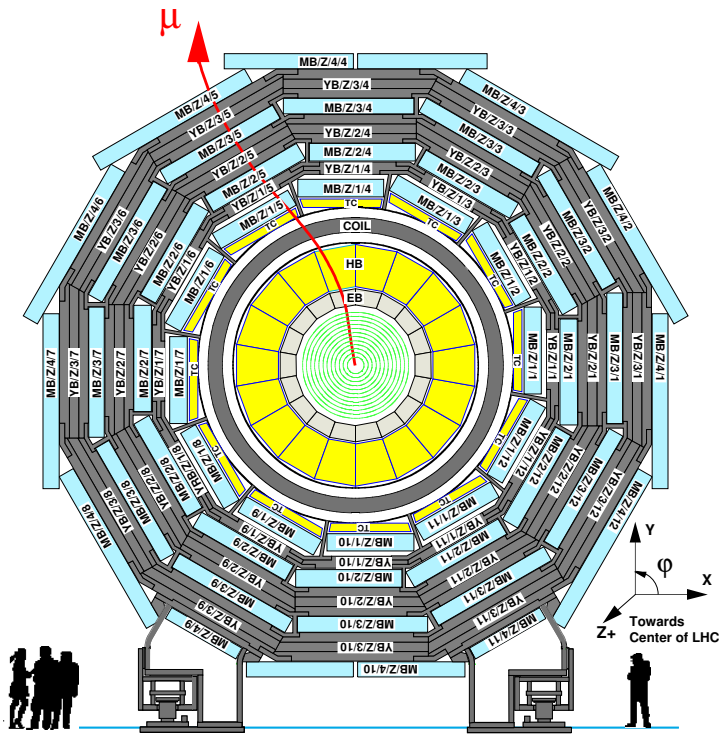

Figure 20. Layout of the CMS barrel muon DT chambers.

The Muon Endcaps contain in total 468 CSCs. One Muon Endcap is shown in Figure 21. The spatial resolution in r$-\phi$ for the Muon Endcap is about 75 μm for the inner chambers and 150 μm for all others.

- **Resistive Plate Chamber (RPC) system** - is a trigger system added in both barrel and endcap regions in order to improve the time response. The RPCs are gaseous parallel-plate detectors that combine adequate spacial resolution (1.5 cm) with a time resolution (1.5 ns) comparable to that of scintillators. The gas used in

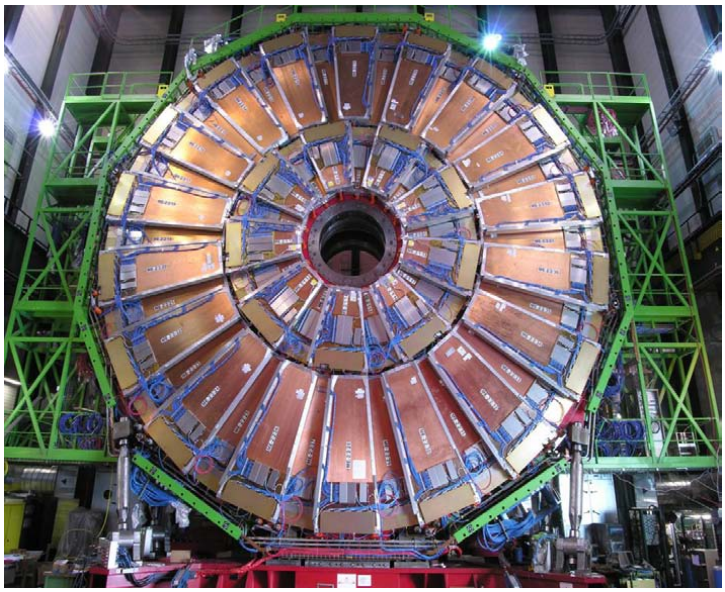

Figure 21. Layout of the CMS Endcap muon CSC chambers.

the RPCs is a mixture of 96.2% $C_2H_2F_4$ + 3.5% C_4H_{10} + 0.3% SF_6. The plates operates in avalanche mode with pick-up read-out strips between them.

The RPCs are embedded into the iron yoke layers. For the barrel, the RPCs are arranged on each side of the first two layers, while for the outer two layers, the RPCs are located only on the inner side of the layers. For the endcaps, the RPCs are arranged on both sides of the three layers. The pseudorapidity coverage of the RPC system is $| \eta | < 2.1$. The arrangement of the RPCs in the Muon System allows the trigger algorithm to perform a good reconstruction of the muons, even for the low p_T particles.

3.3 CMS Trigger System

The proton-proton collisions at the LHC occur at high interaction rates, yielding about 20 simultaneous pp collisions. For the design luminosity of 10^{34} $cm^{-2}s^{-1}$ and a bunch crossing interval of 25 ns, a crossing frequency of 40 MHz is expected. Due to the difficulty of storing and processing such a large amount of data, a drastic rate reduction needs to be achieved. This task is performed by the trigger system, which is the start of the physics event selection process. The rate is reduced in two steps: Level-1 Trigger (L1T) and High-Level trigger (HLT). The L1T consists of custom-designed, largely programmable electronics, and it has the capacity of reducing the amount of data to 100 KHz. The HLT is a software system implemented in a filter farm of computers, reducing the event rate to about 200 Hz.

3.3.1 Level-1 Trigger

The L1T uses coarsely segmented data from the calorimeters and the muon system. It has local, regional and global components as shown in Figure 22. The Local Triggers, also called Trigger Primitive Generators, are based on energy deposits in calorimeter trigger towers and hit patterns in muon chambers. The Regional Triggers combine the information from the Local Triggers and build objects such as electron and muon candidates based on momentum, position and quality flags. The Global Calorimeter and Global Muon Triggers determine the highest rank calorimeter and muon objects across the entire experiment and transfer them to the Global Trigger. The decision to reject or

to accept an event is taken by the Global Trigger and then passed to the HLT for further evaluation.

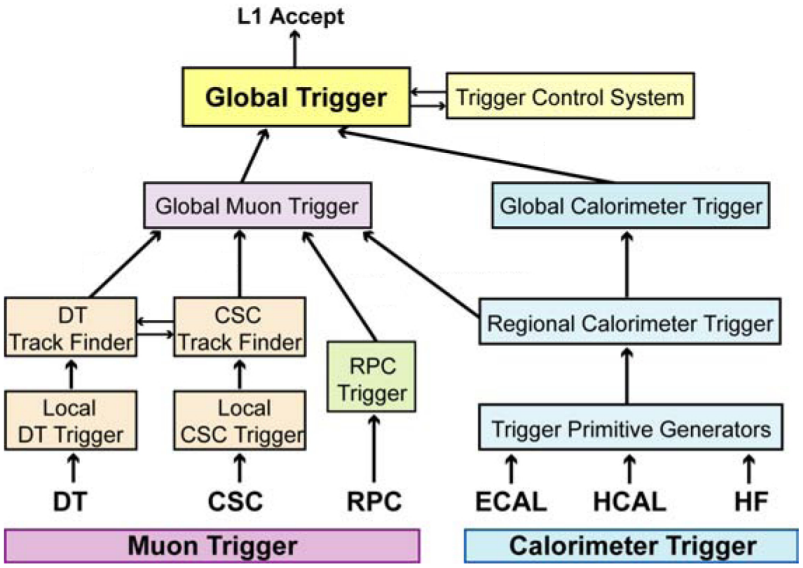

Figure 22. Local, regional and global components of the Level-1 Trigger.

3.3.1.1 Muon Triggers

All three muon systems take part in the trigger. The barrel DT chambers provide local trigger information in form of track segments in the $\phi-$projection and hit patterns in the $\eta-$projection in the central region. The coded information (position, transverse momentum and track quality) is transmitted from the detector to the DT regional trigger, named Drift Tube Trigger Track Finder(DTTF), through high speed optical links. The

endcap CSCs deliver 3D track segments to the CSC Track Finder (CSCTF). Both DTTF and CSCTF identify muon candidates, determine their momenta, locations and quality. The candidates are sorted by rank, which is a function of p_T and quality. For the overlapping region between the barrel and the endcaps, the DTTF and CSCTF exchange information about the reconstructed muon tracks. The RPC trigger chambers have excellent timing resolution and deliver their own track candidates based on regional hit patterns.

The DTTF and the CSCTF deliver each up to 4 muon candidates to the L1 Muon Global Trigger, while the RPC trigger delivers up to 8 candidates. The Global Muon Trigger combines the information from all three subdetectors, achieving an improved momentum resolution and efficiency compared to the stand alone systems. The rapidity coverage of the muon trigger for the first data was $\mid \eta \mid \leq 2.1$. The four highest p_T muon candidates are forwarded to the global trigger.

3.3.1.2 Calorimeter Triggers

The Trigger Primitive Generators (TPG) represents the first step of the Calorimeter Trigger pipeline. For triggering purposes, the calorimeters are subdivided in trigger towers. The TPGs sum the transverse energies measured in ECAL crystals or HCAL read-out towers to obtain the trigger tower E_T and attach the correct bunch crossing number. In the region $\mid \eta \mid < 1.74$ each trigger tower has a $\eta - \phi$ coverage of 0.087×0.087. Beyond that boundary the towers are larger.

The Regional Calorimeter Trigger determines the electron/photon candidates and transverse energy sums per calorimeter region. A region consists of 4×4 non-overlapping towers from ECAL and HCAL. Information relevant for muons about isolation and compatibility with minimum ionizing particles is also calculated. The added information is passed to the Global Calorimeter Trigger.

The Global Calorimeter Trigger determines jets through clustering techniques . The jets together with the total transverse energy, the missing transverse energy, the number of jets and the scalar transverse energy sum of all jets (H_T) are provided to the Global Trigger. Electrons and photons are identified for $\mid \eta \mid < 2.5$.

3.3.1.3 Global Trigger

The main task of the Global Trigger is to accept or reject events based on the objects delivered by the Muon and Calorimeter Triggers, for further processing by the HLT. The Global Trigger can accommodate up to 128 algorithms. In addition, up to 64 algorithms can be simple on/off signals, called technical triggers, which are pre-defined hardware signals.

3.3.2 High-Level Trigger

The HLT (31) uses a filter farm of processors to run an offline reconstruction software to select events for further storage. The reconstruction involves different trigger paths, specific to different types of events. A trigger path contains reconstruction modules and selection filters which are reconstructed in a specific order. All the filters existent in a

trigger path must be passed in order to select an event. The HLT paths are specific to different reconstructed objects: jet triggers, muon triggers, etc. The events selected by the HLT are saved in different datasets. An event can be saved in more than one dataset, if it is selected by multiple HLT triggers.

The **Data Acquisition System (DAQ)** together with the Trigger System form the Trigger and Data Acquisition System (TriDAS) (32). The DAQ system is shown in Figure 23. Once a L1T signal is sent to the HLT, the data from the front-end buffers is read and passed to the DAQ. The event builder represents the first block of the DAQ which delivers the event fragments from the CMS cavern to the surface and then stores them in the readout units. A readout builder consists of multiple readout units, builder units and a single event manager connected to a common switching network. The event manager handles the data flow between different components and allocates events to builder units upon request. The complete events are transferred to different filter units for data consistency checks, reconstruction and filtering. The accepted events are forwarded to a storage manager.

3.4 CMS Computing Infrastructure

Physically, an event is the result of a single readout of the detector electronics which register signals generated by the interaction of particles produced in pp collisions with the detectors. From out of millions of events recorded in the detector, the most interesting 100 events per bunch crossing are selected by the online TriDAS system. The recorded

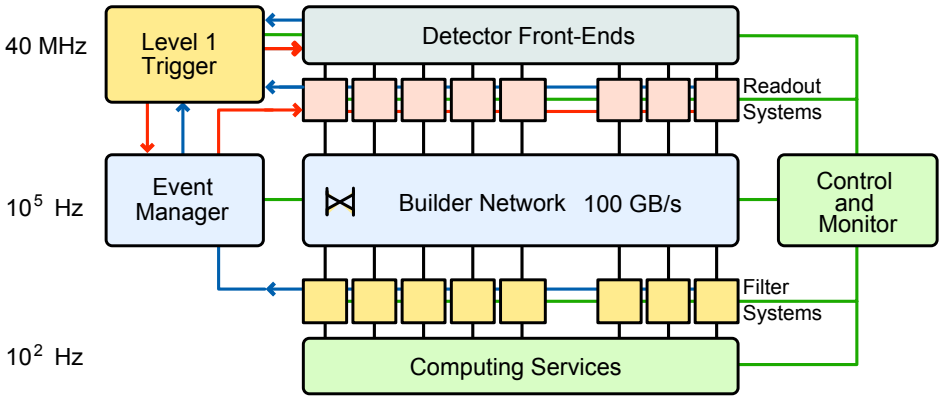

Figure 23. Architecture of the CMS DAQ system.

data is further stored, transferred and manipulated by the CMS computing infrastructure, which consists of multiple computing centers located all over the world. The centers used to store the data are structured in the following way:

- **Tier-0** - a single center located at CERN. It accepts the events from DAQ and copies them to a permanent mass storage. It also performs the first reconstruction of the raw data.
- **Tier-1** - a few large computing facilities located in different countries. Both raw and first reconstructed data are transferred to Tier-1 from Tier-0, where additional data reconstruction is performed. The Tier-1 centers also store the simulated data.

- **Tier-2** - several centers placed at different institutions. The Tier-2 centers support physics analysis activities by providing local storage for various data samples. They are also used in the production of the Monte Carlo (MC) simulated data.

All CMS computing centers are integrated into a single coherent system by the use of the grid middleware (33). It provides a standardized interface to storage and computing facilities and allows for remote job submission and data access. The events are stored as Root (34) files and can be visualized using the Fireworks (35) event display, which provides a simple interface with several graphical and textual views.

The overall collection of software used at CMS, referred to as CMSSW, is built around a Framework, Event Data Model and Services needed by the simulation, calibration, alignment and reconstruction modules that process event data used in the analysis. The Event Data Model is split in four categories: event data producers that add new objects to the event, event data filters that select events based on specific conditions, event data analyzers that produce summary information from events and event data input/output modules for disk storage and DAQ. Different data formats are used at CMS, containing different level of details:

- **RAW** - used to store the full recorded information from the detector and trigger decision. A RAW event occupies 1.5 MB of disk space.

- **RECO** - stores reconstructed objects obtained by applying multiple pattern recognition and compression algorithms to the RAW data. A RECO event occupies 0.5 MB of disk space.
- **AOD** - is a more compact format, produced by filtering the RECO data. It contains selected parameters of the high-level physics objects. An AOD event occupies 100 kB of disk space.

CHAPTER 4

EVENT RECONSTRUCTION AND OBJECT IDENTIFICATION

4.1 Reconstruction Procedure

The reconstruction is the operation of constructing physics quantities from the RAW data collected in the experiment. The reconstruction process is seen as a collection of independent units, each one providing a set of corresponding reconstructed objects. The reconstruction process can be divided into 3 steps:

1. **Local reconstruction**: provides the reconstruction at the level of an individual detector module. The local reconstruction uses as a input the RAW data or simulated data, in the form of digitized electronic signals called **digis**. The output from the local reconstruction are the reconstructed hits, named **RecHits**, which are typically position measurements in tracking-type detectors (Muon and Tracker systems) and energy depositions in Calorimeter systems.
2. **Global reconstruction**: combines information from the different modules of a subdetector. It uses as a input the RecHits produced in the local reconstruction and produces candidates for the reconstructed objects.
3. **Final reconstruction**: combines the reconstructed objects from individual subdetectors to produce higher-level reconstructed objects suitable for high-level triggering or for physics analysis.

In general, all the objects can be reconstructed both locally and globally using specific algorithms for each case. In addition, CMS has implemented the **Particle Flow** algorithm (PF) which combines the information from all sub-detectors to identify and reconstruct all particles produced in the collision, namely charged hadrons, photons, neutral hadrons, muons and electrons, with an optimal determination of their energy and type even for the particles with low momenta and energies. The resulting list of particles can be used to build jets, to determine the missing transverse energy, to reconstruct and identify taus from their decay products, to quantify charged lepton isolation with respect to other particles, to tag b jets, etc. A detailed description of identifying different objects using the PF algorithm can be found in (36).

4.2 Charged Tracks

As stated above, the first step in charged tracks reconstruction is the local reconstruction. The strip and pixel signals produced by the charged particles on the silicon detectors of the tracking system are clustered into "hits". The position of these hits are known with a precision of 15-20 μm due to the excellent spacial resolution of the Pixel and Silicon Tracker detectors (Section 3.2.1). Starting from the reconstructed hits, the track reconstruction proceeds in several steps (37): seed generation, track finding and track fitting and smoothing, described in the following sections.

4.2.1 Seed generation

The trajectory seeds define the starting trajectory parameters and uncertainties used to initiate the iterative track finding procedure to identify the track candidates. The following parameters are needed to properly define a seed: a trajectory curvature, at least 2 hits and a beam constraint. Track candidates based on three hits (triplets) allow reconstruction of the primary vertex and can be used in simple algorithms for online event selection, but are not fully efficient.

The trajectory seeds are constructed on the pixel layers of the tracking system, due to the following reasons:

- the occupancy on the inner pixel subsystem is lower than the outer layers of the strip subdetector, due to the high density of read-out channels
- the pixel sensor hits provide two-dimensional measurements, which allow better defined parameters for constructing the seeds than using the hits from the strip subdetector
- many particles (e.g. pions, electrons) suffer destructive interactions before they reach the outermost layers of the tracking system.

Due to the characteristics of the CMS detector, a highly efficient track finding algorithm can be obtained with a reconstruction of charged particle trajectories starting on the innermost layer of the tracker and moving inside-out.

4.2.2 Track finding

The track finding is based on the combinatorial Kalman filter method (38),(37). The filter proceeds within an iterative process, starting from the seeds and a rough estimate of the trajectory parameters. Information regarding the successive detection layers is added one by one, up to the last point, where the full tracker information is included. The Kalman method includes the following steps:

- **navigation**: determines which layers are compatible with the initial seed trajectory. The trajectory is extrapolated to each of these layers using the equations of motion of a charged particle in a constant magnetic field. The multiple scattering and energy loss in the traversed material is also taken into account.
- **search for compatible hits**: is performed using the Lorentz drift of the ionization charge carriers inside the silicon bulk. Since several hits on the new layer may be compatible with the predicted trajectory, several new trajectory candidates are created, one per hit. In case the track did not leave any hit on that particular layer, a fake hit is created, called invalid hit.
- **trajectory parameters update**: once a hit is found to be compatible with the trajectory, the track candidate is extended to the next compatible layer. If several hits from the same layer are found to be compatible with the trajectory, a track-candidate for each hit is built. Additional track-candidates are created using also the fake hits.

To avoid an exponential increase of the number of track-candidates, only a limited number of these are retained at each step. Hence, the algorithm is configurable through several parameters and it can provide either a high track finding efficiency (usually for offline reconstruction) or a very fast CPU performance suitable for its use in the HLT. The main parameters are (37):

- number of track candidates propagated at each step ≤ 5
- χ^2 for the hits considered compatible with the predicted track state ≤ 30
- number of invalid hits ≤ 1
- transverse momentum ≥ 0.9
- number of hits per track ≥ 5

In addition, a "stopping condition" can be used, and the pattern recognition is interrupted before reaching the outermost layers of the tracker. This is the case for HLT, when the accuracy of the track parameters is often reached after 5 of 6 hits and a continuation to 12 hits would be an unnecessary CPU consumption. Due to the specific geometry of the CMS Tracker, it is easy to navigate between barrel layers during the track finding. However, the navigation is more complex for the transition between the barrel and the end-cap regions, where a single layer can be connected to several others. As the propagation distance to some of these layers can be quite large, the uncertainty of the compatibility between the trajectory and the new hits is large, and also the probability of finding fake hits increases. Once the trajectory reconstruction reaches the end-cap

disks, the navigation is again simplified, since many of the trajectory fake hits are quickly discarded, usually because they fail the condition on the minimum trajectory's χ^2 or the maximum number of invalid hits.

There are a few cases which can produce ambiguities in track finding: the same track is reconstructed more than once starting from different seeds, or more than one track-candidate is built from the same seed. These ambiguities must be resolved in order to avoid that the same charged particle is counted multiple times. The identification of the track duplicates within the track finding algorithm is performed by using the fraction of shared hits between two track candidates (39):

$$f_{shared} = \frac{N_{shared}^{hits}}{min(N_1^{hits}, N_2^{hits})}, \tag{4.1}$$

where N_i^{hits} represents the number of hits in the i track candidate. If this value exceeds the cut value $c = 50\%$, the track with the least number of hits is removed from the track candidates collection. If both tracks have the same number of hits, the one with the highest χ^2 is discarded. The ambiguity resolution method prefers the long trajectories, since the tracks with the small number of hits are always discarded. This can introduce a bias, because even though this criteria is successful in most of the cases, it fails in some special conditions, such as the reconstruction of low energy hadrons and electrons. This effect is diminished by using a configurable value for the cut value c, depending on the reconstructed object.

4.2.3 Track fitting and smoothing

The track finding results in a collection of trajectories, each associated to a number of hits and an estimate of the track parameters. The estimate of the trajectories can be biased by the constraints applied during the seed generation. Therefore, the trajectory is refitted using a combination of the standard Kalman filter and smoother. The Kalman filter is initialized at the location of the innermost hit with an estimate obtained during seeding. The corresponding covariance matrix and the track parameters are updated with each valid hit added to the trajectory. This first filter is complemented with a smoothing stage: a second filter is initialized with the result of the first one and runs backwards towards the beam line. The filter and smoothing procedure yields optimal estimates of the parameters at the surface associated with each hit. The estimates on other surfaces, e.g. at the impact point, are then derived by extrapolation from the closest hit.

For an average LHC event, the combinatorial track finder described above yields a significant fraction of fake tracks that can be removed by a series of cuts on the normalized χ^2, the longitudinal and transverse impact parameters with respect to the beam spot, and their significances. Tracks that fail the loosest selection are rejected, while those that pass the tightest selection are labeled **highPurity**.

4.2.4 Track reconstruction performance

The performance of the track reconstruction algorithm is evaluated by the efficiency and resolution on different categories of simulated samples. There are two types of

efficiencies that can be measured: **the algorithmic efficiency** which represents the efficiency of reconstructing correctly the simulated particles which have at least three hits on three different tracker layers, and **the global efficiency** which represents the fraction of charged particles, among all the simulated ones, that have a partner within the collection of reconstructed tracks. The algorithmic efficiency is exclusively related to the ability of the tracking algorithm to perform the track-finding task, all other stages being fully efficient. The global efficiency includes besides the efficiency of the algorithm, the acceptance, the hit efficiency and any other factor influencing the reconstruction. For this reason, the algorithmic efficiency is slightly higher than the global efficiency.

The global efficiencies for single particles like muons, pions and electrons measured in MC simulated data are shown respectively in Figure 24, Figure 25 (top) and Figure 26 (top) (39) as a function of pseudorapidity, for different transverse momenta. At high $|\eta|$, the drop in the efficiency is mainly due to the lack of coverage in the forward pixel disks. The fake rate for reconstructing tracks for pions and electrons measured in MC data are also shown in Figure 25 (bottom) and respectively Figure 26 (bottom). Since both pions and electrons are subject to multiple scattering and energy loss by ionization, their track reconstruction efficiency is lower than the corresponding one for muons. The decrease in efficiency in the transition and end-cap regions is due to the fact that the merging of separate trajectories is more common in these regions. This effect is correlated with the increase of the fake rate.

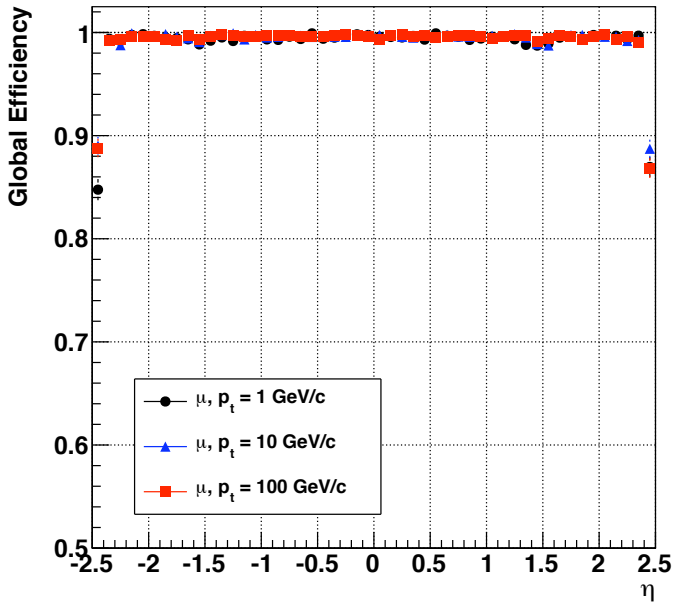

Figure 24. Global track reconstruction efficiency as a function of pseudorapidity for muons with different transverse momenta measured in MC.

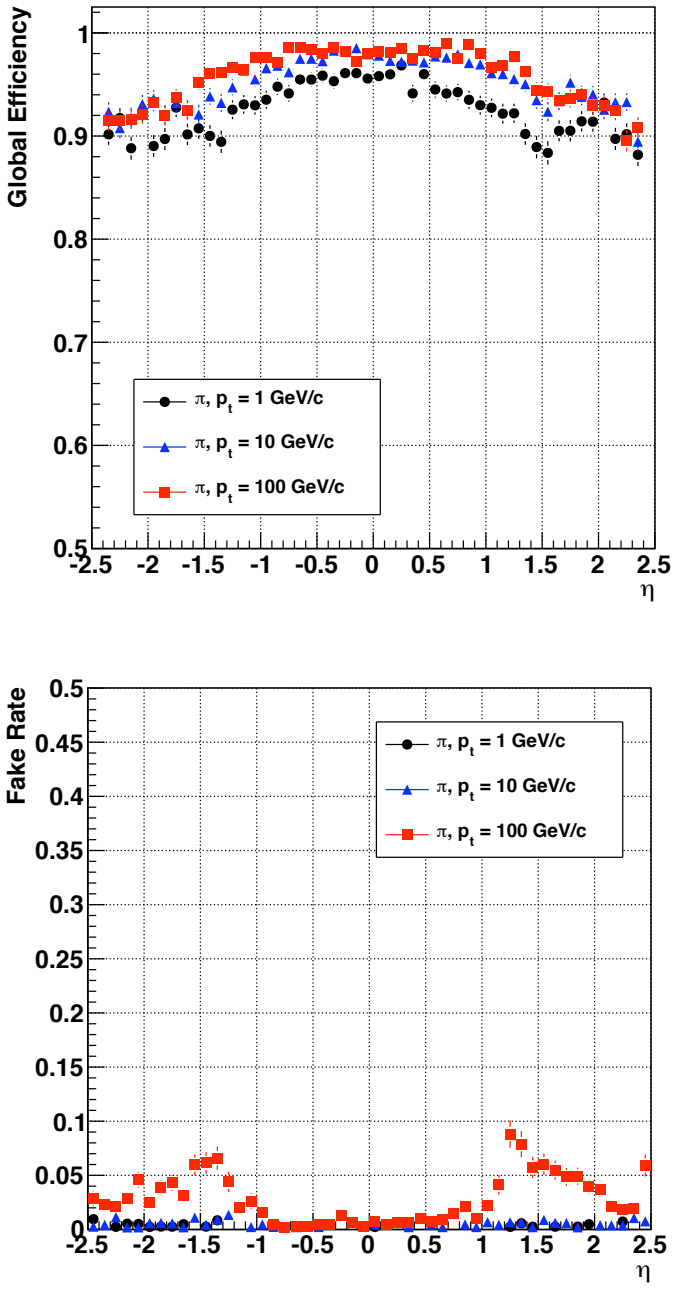

Figure 25. Global track reconstruction efficiency as a function of pseudorapidity for pions with different transverse momenta (top) and fake rate (bottom) for tracks reconstruction for pions measured in MC.

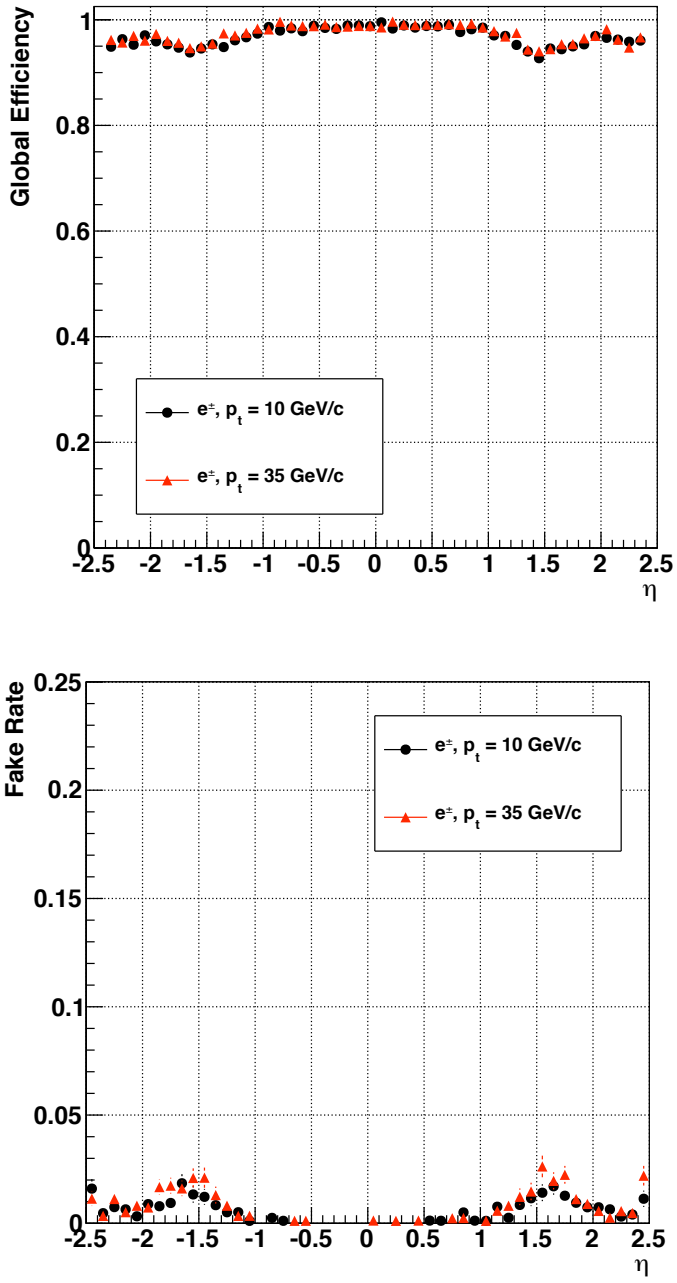

Figure 26. Global track reconstruction efficiency as a function of pseudorapidity for electrons with different transverse momenta (left) and fake rate (bottom) for tracks reconstruction for electrons measured in MC.

The track resolution is measured as a function of the pseudorapidity of the simulated charged particle by the standard deviation of the Gaussian fit of the residuals distribution, i.e. the differences between the reconstructed and simulated parameters. For single particles at high momentum, the impact parameter resolution is fairly constant ≈ 10 μm, while at low momenta, the impact parameter resolution is progressively degraded by multiple scattering. In the case of transverse momentum, the track reconstruction resolution is $1-2\%$ up to a pseudo$-$rapidity of $|\eta| = 1.6$. There is a degradation observed for $|\eta| > 1.6$, which is due to the gap between the barrel and the end-cap disks and to the lower hit resolution of the hits from the outermost layers of the tracker.

4.3 Primary Vertex

The Primary Vertex (PV) represents the location where the proton-proton collision takes place. To reconstruct a collision event, it is very important to precisely determine the PV position.

The PV reconstruction involves two steps: the vertex finding and the vertex fitting. The **vertex finding** determines the position and the uncertainty of an interaction vertex by grouping tracks into vertex candidates. The tracks originating from the interaction region are selected based on:

- location: $-$ given by the transverse impact parameter significance with respect to the beam line (d_0^{BS}). The impact parameter significance is defined as the mea-

sured track impact parameter with respect to the beam line divided by its uncertainty (40).

- quality: − given by the number of hits associated to each track and the normalized track χ^2.

Tracks are clustered based on the z component of the measured impact parameter with respect to the beam line $d_0^{BS}(z)$. The vertex candidates are formed by grouping tracks that have $d_{0_i}^{BS}(z) - d_{0_j}^{BS}(z) < 1$ cm, where the i and j tracks are the closest neighbors.

Next, the vertex candidates containing at least two tracks are fit with an adaptive **vertex fit** (41), which computes the best estimate of the vertex parameters. To each track in the vertex is assigned a track weight between 0 and 1 based on its compatibility with the common vertex. A track consistent with the common vertex has a weight close to 1 which can be used to select the real proton-proton interactions.

The vertex reconstruction efficiency is measured in simulated data by asking how often its position is consistent with the true value. The primary vertex efficiency depends strongly on the number of tracks in the cluster and hence can be influenced by the fake rate of the reconstructed tracks. This bias can be avoided if only tracks with transverse momentum above 0.5 GeV are used. The vertex reconstruction efficiency can be measured in both simulated samples and real data, using the tag and probe method described in (40). The efficiency is measured by counting how often the probe vertex is matched to the original vertex, given that the tag vertex is reconstructed and matched to the original

vertex. Figure 27 shows the primary vertex efficiency as a function of the number of tracks associated to the vertex. The efficiency is estimated to be close to 100% if there are more than 2 tracks with transverse momenta greater than 0.5 GeV in the vertex.

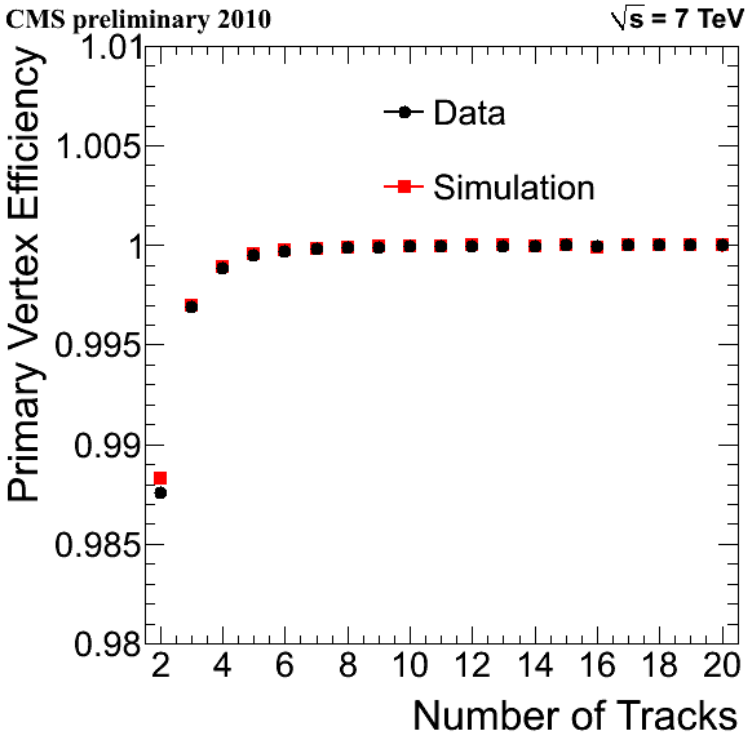

Figure 27. Primary vertex efficiency as a function of the number of tracks in a cluster measured in MC.

The resolution of the reconstructed vertex is measured (40) to be between 250 and 100 μm for a low number of tracks in the cluster (2 to 6 tracks) and below 50 μm for 10 or more tracks associated to the same vertex. For minimum bias events at 7 TeV the resolution is between 20 and 25 μm, for primary vertices using more than 30 tracks.

4.4 Muons

4.4.1 Muon reconstruction

The ability to reconstruct and identify muons with high efficiencies over a wide range of energies and the entire geometric acceptance of the detector is essential for the proper identification of physics signatures at the LHC. The reconstruction algorithms use the information from the entire detector or only a specific subdetector, resulting in three categories of reconstructed muons:

- **Stand-alone muon** - is reconstructed using only the information from the muon system, but the resolution can be affected, especially for muons with low energy. The seeds are built based on the DT and CSC segments. The seed trajectory state parameters are propagated to the innermost compatible muon detector layer and a pre-filter is applied in the inside-out direction in order to refine the seed state before the true filter. The final filter, applied from the outside-in direction, builds the muon trajectory. In order to finally accept the trajectory as a muon track, at least two measurements, one of which must be a DT or CSC type, must be present in the fit, this allowing the rejection of fake DT/CSC segments. Moreover,

the inclusion of the resistive plate chambers (RPC) measurements can improve the reconstruction of the low momentum muons and those muons which escaped through the interspace between the wheels. After the fake track suppression, the parameters are extrapolated to the point of the closest approach to the beam line.

- **Global muon** - is reconstructed using information from both muon and tracker detector. The momentum resolution of the muon tracks with $p_T < 200$ GeV/c reconstructed in the muon system alone is dominated by multiple scattering. At low momentum, the best momentum resolution for muons is obtained from the silicon tracker. At higher momentum, however, the characteristics of the muon system allow the improvement of the muon momentum resolution by combining the muon track from the silicon detector with the muon track from the muon system into a global muon track.

  The first step in reconstructing a global muon track is to find a matched silicon tracker track for each stand-alone muon track. The matching procedure involves the comparison of trajectory parameters for the two tracks, after propagating them onto a common surface. The next step is to fit a track using the hits from the tracker track and the stand-alone muon track. If there are more tracker tracks that fit to one stand-alone muon track, the constructed global muon track with the best χ^2 is chosen.

- **Tracker muon** - is reconstructed using the tracks from the tracker detector matched with the segments from the muon detector. Since in some cases the hit and segment information in the muon system is minimal, the stand-alone muon reconstruction fails. The tracker muon provides a complementary approach to reconstruct and identify muons starting from a silicon tracker track and then searching for compatible segments in the muon detectors. The energy deposition in the calorimeter can also be used for muon identification. The RecHits (Section 4.2) are used by default to compute the energy deposits, but in the calorimeter, the CaloTowers can also be used, allowing for a more precise muon-missing transverse energy correction.

For the stand-alone and global muon reconstruction, the efficiencies are factorized as following:

$$\begin{aligned} \epsilon_{seed} &= \epsilon_{seed-algo} \times \epsilon_{\mu-acceptance} \\ \epsilon_{sta} &= \epsilon_{seed} \times \epsilon_{sta-algo} \\ \epsilon_{glb} &= \epsilon_{sta} \times \epsilon_{tk} \times \epsilon_{matching} \end{aligned} \tag{4.2}$$

where:

- $\epsilon_{seed}-$ is the efficiency of finding a seed which is matched to the simulated hits, including the detector acceptance and the muon energy loss in the material
- $\epsilon_{sta}-$ represents the stand-alone muon reconstruction efficiency

- $\epsilon_{glb}-$ represents the global muon reconstruction efficiency
- $\epsilon_{tk}-$ represents the muon reconstruction efficiency in the inner tracker system

Figure 28 shows the efficiency of different muon reconstructions as a function of muon η and transverse momentum. The loss in ϵ_{sta} efficiency at $\mid \eta \mid = 0.3$ is due to a geometrical effect, since in that region there is a discontinuity between the central wheel and its neighbors. The region $0.8 < |\eta| < 1.2$ is also known to be problematic as DT and CSC segments are used together to estimate the seed state. The loss in the tracker muon efficiency at $\mid \eta \mid \approx 0$ is due to the tracker geometry, but overall the muon efficiency reconstruction in the tracker is quite stable and above 99%. The global muon reconstruction efficiency is dominated by the stand-alone muon reconstruction efficiency, all other efficiencies, including the matching efficiency being very close to 1 over the entire η and p_T range (42). The reconstruction efficiency of the tracker tracks (General Tracks) for the muons is also shown, for comparison.

4.4.2 Muon identification

There are several well identified processes through which muons are produced in proton-proton collisions: **prompt muons** - come from decays of W and Z bosons, Drell-Yan or top quark production, **muons from heavy flavors** - come from a b or c quark or a tau lepton, **muons from light flavors** - come from decays in flight of π or K or the decay of particles produced in nuclear reactions in the detector material, **hadron punch-through** - the hits from the muon chambers were produced by particles other

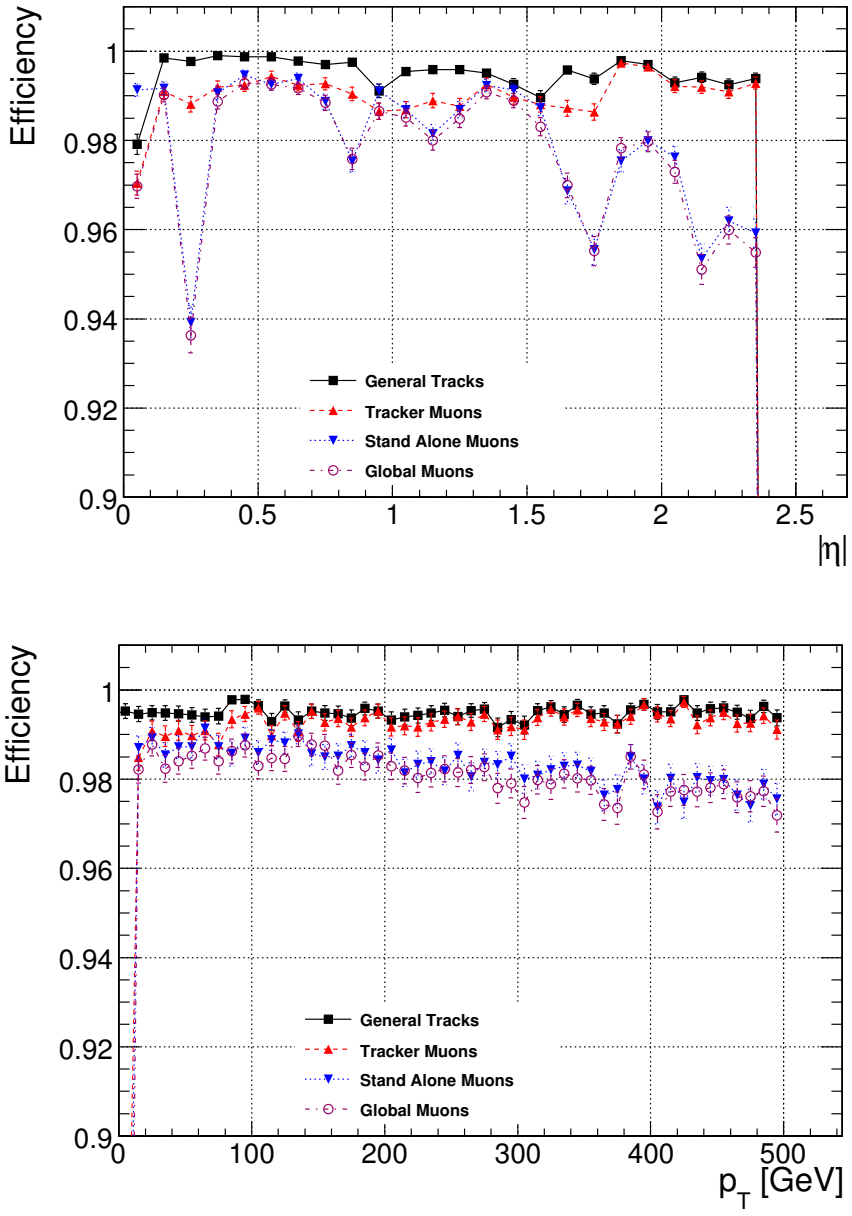

Figure 28. The efficiencies of different muon reconstructions as a function of muon pseudorapidity (top) and transverse momentum (bottom).The reconstruction efficiency of the tracker tracks (General Tracks) for the muons is also shown, for comparison.

than muons that escaped the calorimeters, **duplicate muons** - are due to instrumental effects or imperfections which affect the trajectory building step in the reconstruction process.

During the reconstruction process, information about the muon quality is also stored:

- number of muon stations matched with segments identified from the inner tracked extrapolation - is a powerful tool to reject muons from light flavors
- transverse impact parameter of the muon with respect to the primary vertex - is useful to separate prompt muons from muons from heavy flavor decays
- track fit χ^2 in the silicon tracker - is a good discriminant for the decays in flight
- muon isolation - is a good parameter to select muons among other particles using tracker and calorimeter information.

In general, the muon selection is analysis dependent, this might require a balance between the muon identification efficiency and purity. In the analysis described in this thesis, we used the muons reconstructed with the PF algorithm. In order to identify the PF muons, the selection is applied on all the muon candidates reconstructed using the standard algorithms. The selection has been optimized to identify muons in jets with high efficiency, keeping fake rate from misidentified charged hadrons low.

A detailed description of the muon identification parameters and their efficiencies in selecting muons in both data and MC is presented in (43). In particular for the

measurement described in this thesis, the results of the muon selection efficiency are shown in Section 5.5.

4.5 Electrons

The electron reconstruction relies on the energy deposited in the ECAL crystals. The presence of the tracker results in bremsstrahlung and photon conversions. The photon also deposits energy in the calorimeter and hence, the reconstruction of the electrons depends on how well photons are identified.

The electron reconstruction also proceeds with the seeding and trajectory building steps. At the track seeding stage, there are two complementary algorithms used: the **tracker driven** and the **ECAL driven** seeding. The tracker driven seeding is more suitable for electrons with a low transverse momentum and electrons inside jets. The ECAL driven seeding starts with the reconstruction of the clusters seeded by ECAL cells with the highest amount of energy deposited. Because of the strong magnetic field, the energy reaching the calorimeter is spread in ϕ. The energy is therefore clustered in the ECAL barrel using a supercluster, obtained by grouping clusters within a ϕ window around the starting crystal up to a maximum extension of 0.3 rad. In the ECAL endcaps, the clusters are formed within a window of 5×5 matrices crystals and the superclusters are formed by grouping such clusters whose position lie within $\phi < 0.3$ rad.

For building the trajectory, the regular Kalman filter used in the CMS track reconstruction is not optimal due to the energy loss of the electrons though radiative

interactions in the tracker material. Hence a new algorithm, the Gaussian Sum Filter (GSF), has been developed and implemented in CMS to model the energy losses due to Bremsstrahlung radiation in the tracker layers (44). This method takes into account the sudden changes in the curvature radius caused by radiative photon emission in the layer to layer propagation. The GSF track is extrapolated from its outermost measurement to the ECAL, and it is associated to the seed cluster - the closest ECAL cluster with energy deposited by the electron. If no such seed cluster is found, the GSF track is no longer considered for electron reconstruction.

For each tracker layer, a search for a Bremsstrahlung photon emission is performed by extrapolating a straight line tangent to the direction of the GSF track up to the ECAL entrance. If an ECAL cluster can be linked to this line, its energy is assigned to the total electron energy.

A set of variables is used to identify the electron. These variable can be calorimeter-based (e.g. the shower width) or tracker-based, like the fraction of the initial electron energy carried by Bremsstrahlung photons as measured by the GSF track (f_{brem}), the transverse momentum uncertainty from the GSF fit, the χ^2 of the GSF fit, etc. (45). An efficiency of $\approx 90\%$ was achieved to select the electrons based on loose cuts on the identification variables. A good agreement between data and MC simulation was observed for all the electron identification variables (46), (45).

4.6 Jets

Jets are localized streams of particles generated through quark and gluon radiation emerging from the p-p collisions. At CMS, jet production has a large cross section and hence understanding their energy calibration and resolution is crucial for many analyses with jets in the final state.

4.6.1 Jets reconstruction

Jets were reconstructed using the $anti{-}k_T$ clustering algorithm (47) with the size parameter $R = 0.5$. At CMS there are three types of reconstructed jets, depending on the way in which the information from different detectors is combined: calo-jets, jet-plus-track jets and particle flow jets. The calo-jets are reconstructed using the calorimeter towers. The calorimeter tower contains the combined energy of an HCAL cell and 25 underlying ECAL crystals. A calo-jet is build if the tower energy exceeds 1 GeV. The jet-plus-track jets are reconstructed by adding the information from the calorimeter towers with the information from the tracking detectors. The PF jets are reconstructed using the PF algorithm. The PF jets are reconstructed after the list of identified particles is obtained with the PF algorithm. The jet momentum and spacial resolution are improved with respect to the calo-jets and jet-plus-track jets because the tracking detectors and the excellent granularity of the ECAL allow a precise measurement of the charged hadrons and photons inside jets, which constitute 90% of the jet energy. In the measurement presented in this thesis, we used the PF jets.

In order to identify and select jets, CMS developed a jet quality criteria (**jet ID**) proven to retain the majority of real jets in MC while rejecting most fake jets arising from calorimeter and/or readout electronics noise. The jet ID cuts for PF jets consist of the following cuts (27):

- charged hadron fraction $CHF > 0.0$ within the tracking region of $\eta < 2.4$
- neutral hadron fraction $NHF < 1.0$
- charged electromagnetic (electron) fraction $CEF < 1.0$
- neutral electromagnetic (photon) fraction $NEF < 1.0$

In the simulated samples, all the generated stable particles, except for neutrinos, give rise to **gen-jets**. The gen-jets are obtained by clustering the stable particles produced during the hadronization process that follows the hard interaction. The reconstructed jets are matched to the closest gen-jet in the (η, ϕ) plane. The fraction of the reconstructed jets that are matched to generated jets represents the matching jet efficiency. The matching jet efficiency for both PFjets and calo-jets are shown in Figure 29. The matching jet efficiency for calo-jets is found to be greater than 80% for jets with $p_T > 20$ GeV/c, reaching the 100% plateau above 40 GeV/c. The jet efficiency for PFjets is 100% for jets with $p_T > 20$ GeV/c.

4.6.2 Jet energy calibration

The energy of a jet that is reconstructed and measured in the detector is usually different from the energy of the corresponding particle jet. The causes for this difference

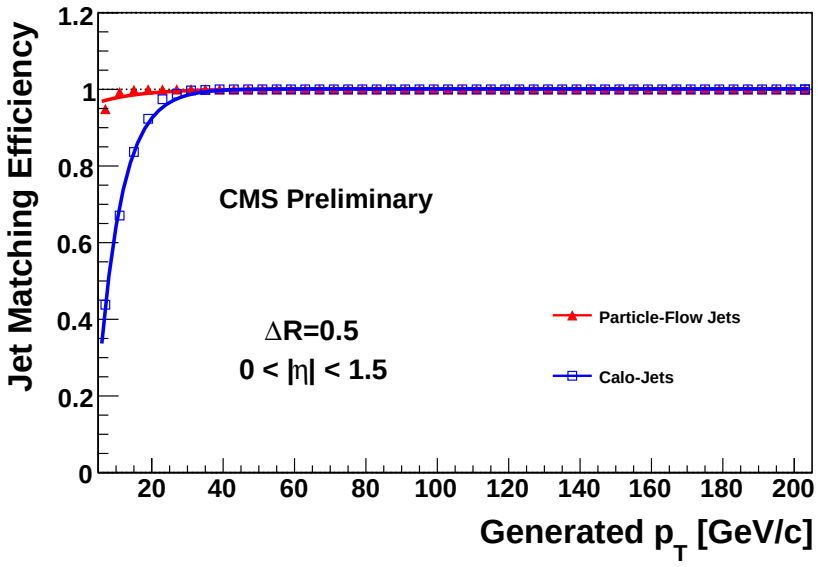

Figure 29. Jet matching efficiency as a function of generated jet p_T for calo-jets and PF jets in the barrel region.

are the non-uniform and non-linear response of the CMS calorimeters, electronics noise and pile-up. The purpose of the jet energy calibration is to relate, on average, the energy measured in the detector to the energy of the corresponding particle jet. The jet energy calibration at CMS involves the following corrections:

1. **Offset correction** - is used to correct the jet energy for electronics noise and pile-up.

2. **Relative correction** - removes variations in jet energy response versus jet η. For this purpose, a control region, usually the barrel ($\eta < 1.3$), is taken as a reference. A dijet p_T balance technique (48) is used to search for back-to-back (in ϕ) dijet

events, where for a given jet in the reference region there will be a corresponding jet in the probe region, at arbitrary η, with the same transverse momentum. The difference between the transverse momentum of the reference and probe jets is used to compute the relative jet energy response. The relative correction is parametrized versus jet η separately for different jet p_T regions.

3. **Absolute correction** - removes variations in jet energy response versus jet p_T. The absolute jet energy response is measured in the reference region (barrel) using γ+jets events, with two different methods: the missing transverse energy projection fraction (MPF) and the p_T balance. Both methods exploit the balance between the photon and the opposite jet in the transverse plane.

Jet energy corrections are applied as shown in Equation 4.3:

$$E_{corrected} = (E_{uncorrected} - E_{offset}) \times C_{rel}(\eta, p_T^{''}) \times C_{abs}(p_T^{'}) \quad (4.3)$$

where $p_T^{''}$ represents the transverse momentum of the jet on which the offset correction was applied, and $p_T^{'}$ represents the transverse momentum of the jet corrected by offset and relative corrections. The corrections are initially obtained using simulated samples and used as the first step to correct the reconstructed jet energy (27). Next, small residual corrections are applied for relative and absolute jet energy scale (JES) measured in data. The residual corrections used to correct the jet energy scale are very small, indicating

that the Monte Carlo truth jet energy calibration is a good starting point. Detailed descriptions of the methods used to compute the corrections are found in (27) and (48).

The overall jet energy correction factor and its uncertainty is shown separately as a function of jet p_T and jet η in Figure 30.

4.7 Jet tagging

The identification of the b-quarks is crucial for distinguishing events like top quarks, Higgs bosons or Supersymmetric particles from background processes containing only light flavor (u, d, c, s, g) jets. The b jet identification relies mainly on the distinct properties of the b hadrons, such as large lifetime ($\tau_b \approx 1.5$ ps, $c\tau_b \approx 450$ μm), large mass, decays to a large number of particles (on average 5 charged tracks), relatively large semileptonic BRs (b hadrons decay into muons or electrons in about 20% of the cases), and a hard fragmentation function (the b hadron in a b jet carries a large fraction of the jet energy).

The b-tagging algorithms developed at CMS use a combination of two or more properties of the b-quark in order to discriminate between the b-quarks and light flavor quarks. The key ingredient for all lifetime-based b-tagging algorithms at CMS is the efficient track reconstruction, in particular the precise reconstruction close to the interaction point. The following track selection was used in the b-tagging algorithms with the first 7 TeV data at CMS (49):

- number of tracker hits ≥ 8 (among them, at least 2 hits are in the pixel detector)

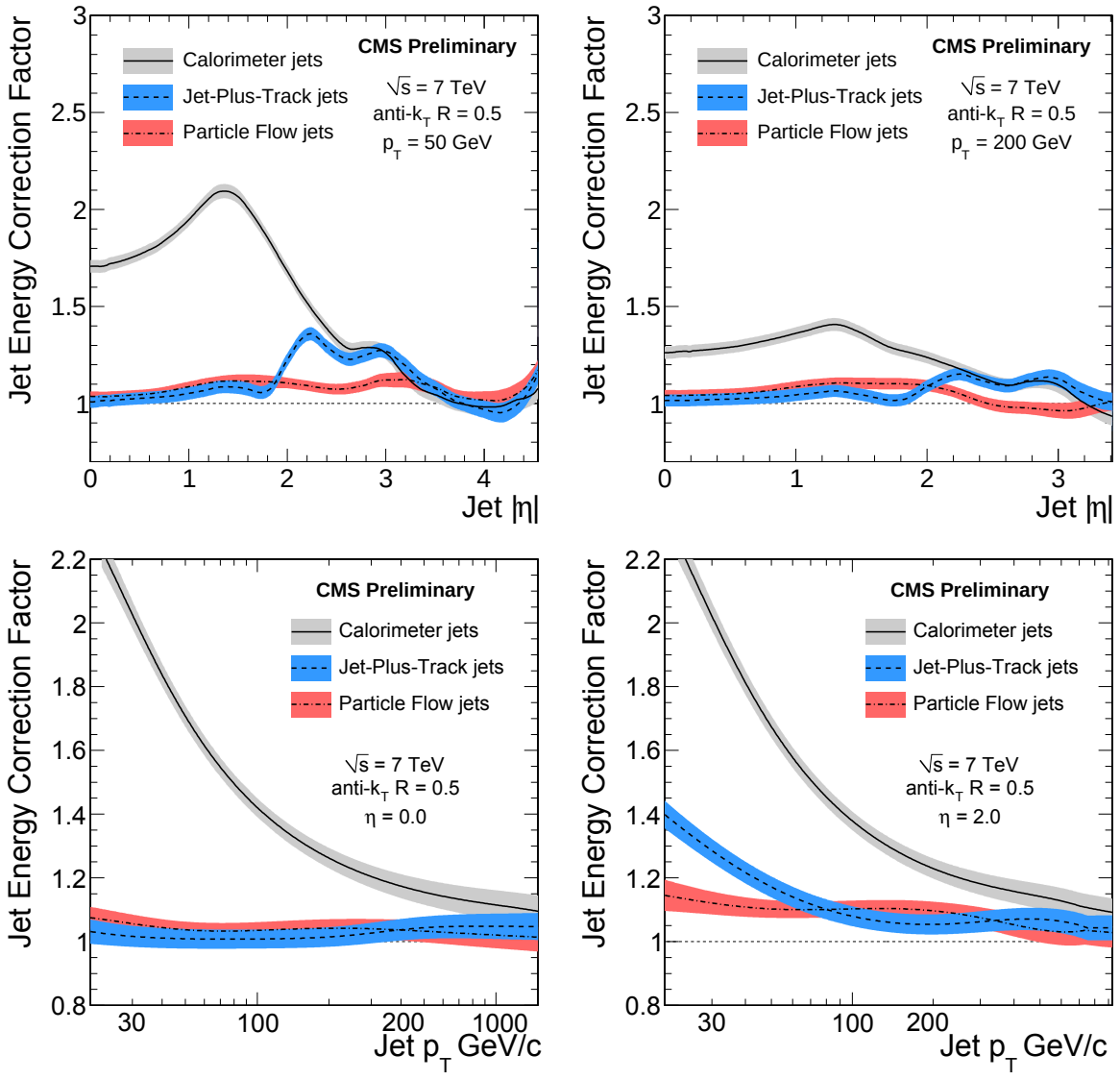

Figure 30. Total jet energy correction factor and its uncertainty band as a function of jet η for fixed values of jet p_T: 50 GeV/c (top, left) and 200 GeV/c (top, right) and respectively as a function of jet p_T for fixed values of jet η: 0.0 (bottom, left) and 2.0 (bottom, right).

- χ^2/ndof of the track fit < 2
- $p_T > 1$ GeV/c
- unsigned transverse impact parameter $d_{0xy} < 0.2$ cm
- unsigned longitudinal impact parameter $d_{0_z} < 17$ cm
- distance of closest approach to the jet axis < 0.07 cm
- decay length < 5 cm.

The good resolution of the CMS Pixel detector allows for a precise reconstruction of the b-decay vertex, named the secondary vertex. The tracks associated to the b-decay products originate from the secondary vertex, which is found to be displaced with respect to the primary vertex, as shown in Figure 31. In comparison, the tracks associated to the light flavor jets originate mainly from the primary vertex. The displacement of the secondary vertex with respect to the primary vertex is named decay (flight) length and its sign is considered positive if the tracks associated to the secondary vertex are produced in the direction of the b-jet and negative, otherwise. The negative values of the decay length are mostly corresponding to the light flavor jets, since they are more affected by the resolution compared to the heavy flavor jets. Hence, the measurement of the mistag rates in data is based on the negative values of the discriminator.

4.7.1 Simple Secondary Vertex b−Tagging algorithm

The Simple Secondary Vertex (SSV) is one of the b−tagging algorithms (**taggers**) (50) implemented at CMS. SSV is based on the reconstruction of at least one secondary ver-

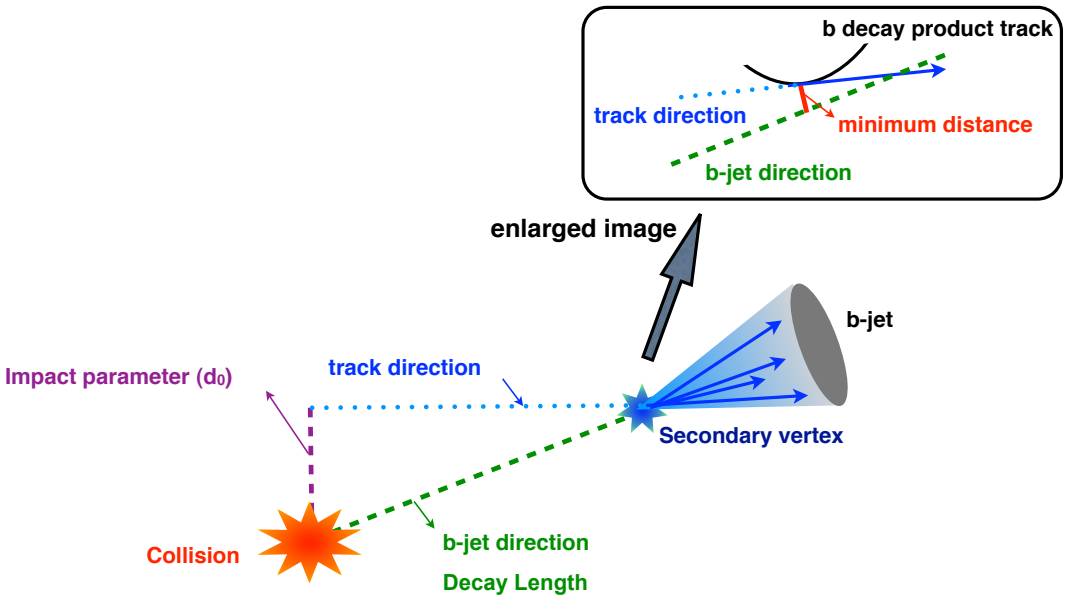

Figure 31. Representation of the secondary vertex, from which the b decay tracks originate. The track impact parameter measured with respect to the primary vertex is also shown.

tex. The b-tagging algorithm produces as output a numerical **discriminator**, denoted D which is computed from the three dimensional decay length L_{3D}:

$$D = sign(L_{3D})log(1 + \frac{\mid L_{3D} \mid}{\sigma_{L_{3D}}}). \tag{4.4}$$

There are also 2 taggers associated to this algorithm, based on the number of tracks N_{trk} associated to the decay vertex: **simple secondary vertex high efficiency (SSVHE)**, for $N_{trk} \geq 2$ and **simple secondary vertex high purity (SSVHP)**, for $N_{trk} \geq 3$. The operating points are chosen based on the amount of light partons passing the SSV

requirements in simulated samples: **loose (L)** – corresponding to 10% mistag for light flavor partons, **medium (M)** – 1%, or *tight (T)* – 0.1%.

4.7.2 b-Tagging performance

The efficiency to tag a certain flavor "q" of the jet as a b jet is defined as

$$\epsilon_q = \frac{\text{Number of jets of flavor q tagged as b}}{\text{Number of jets of flavor q}}. \quad (4.5)$$

For b-jets, the measured efficiency is named b-tagging efficiency. For jets with light flavors, the ϵ_q represents the mistag rate. This efficiency can be directly measured in Monte Carlo simulated samples, since the true flavor of a reconstructed jet is determined by analyzing the parton content in a cone around the jet direction. A reconstructed jet is matched to the initial parton from the primary process if it is within a cone of radius $\Delta R = \sqrt{(\Delta\phi^2 + \Delta\eta^2)} < 0.3$.

In data, the measurement of the b-tagging efficiency is done with data-driven methods, like the **System8** method. The System8 method is based on two samples containing muons within jets, one sample being a subset of the other, and two separate taggers. The correlations between the two taggers and the two samples form a system of eight equations. By solving the eight equations, the b-tagging efficiencies are obtained as a function of jet transverse momentum and jet pseudorapidity. A detailed description of the System8 method and its performance is found in (51).

For the measurement presented in this thesis, the Simple Secondary Vertex b-tagging algorithm with High Efficiency and Medium operating point was used (SSVHEM). The cut for the discriminator D in SSVHEM is 1.79. The b-tagging efficiency for SSVHEM measured in data with the System8 method is shown in Figure 32 (top) as a function of jet p_T.

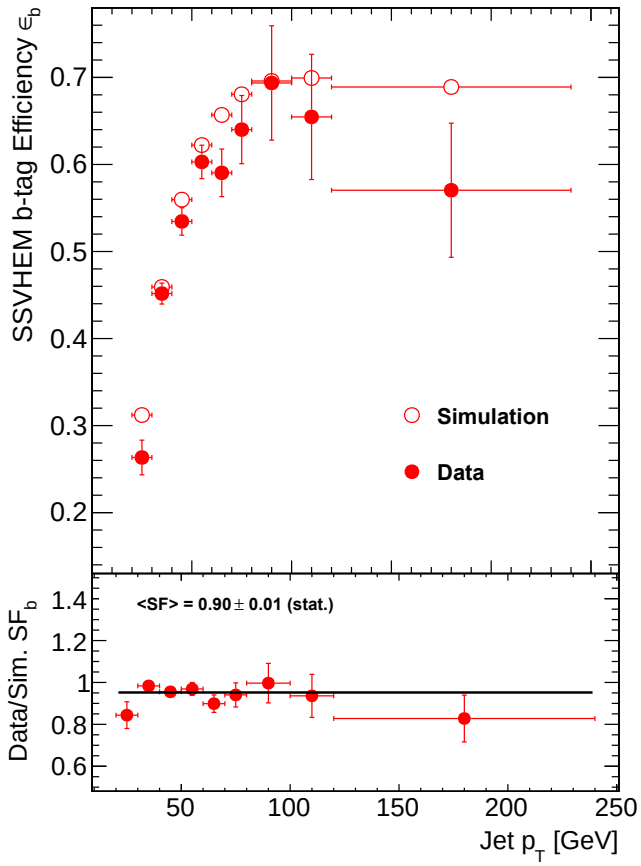

Figure 32. The b−tagging efficiency and the scale factor for the SSVHEM measured with System8 method as a function of jet p_T.

The b-tagging efficiency is also measured in Monte Carlo simulated samples with the same method. The difference between the efficiencies measured in data and Monte Carlo samples is expressed as a ratio between the two efficiencies, denoted b-tagging scale factor (SF). The scale factor is also parametrized as a function of jet p_T and shown in Figure 32 (bottom). The efficiency increases with higher jet p_T ranging from 30% to 65%. The scale factors are ≈ 0.9 for all jet p_T and jet η ranges.

4.8 Missing Transverse Energy

Neutral particles like neutrinos escape detection, but their presence can be inferred by an imbalance in the transverse momentum of the event. The vector momentum imbalance in the transverse plane to the beam direction is known as missing transverse momentum, denoted as $\overrightarrow{\not{E}}_{\mathrm{T}}$. Its magnitude represents the missing transverse energy and is denoted $\not{E}_{\mathrm{T}}$.

In general, $\overrightarrow{\not{E}}_{\mathrm{T}}$ is calculated as the negative of the vector sum of the components of transverse momentum of all final state particles reconstructed in the detector. There are three types of algorithms developed at CMS to reconstruct $\overrightarrow{\not{E}}_{\mathrm{T}}$: calorimeter $\not{E}_{\mathrm{T}}$, track-corrected $\not{E}_{\mathrm{T}}$ and particle flow $\not{E}_{\mathrm{T}}$. In this thesis, we used the particle flow $\not{E}_{\mathrm{T}}$, which is based on the PF algorithm. The PF $\not{E}_{\mathrm{T}}$ is the negative vector sum of all reconstructed particles in the event: muons, electrons, photons and charged and neutral hadrons.

The performance of the particle flow $\not{E}_{\mathrm{T}}$ reconstruction algorithm was studied in both data and Monte Carlo simulated samples and more details can be found in (52) and (53). A good agreement between data and MC is observed.

CHAPTER 5

DATA AND EVENT SELECTION

5.1 Data Sample

For the measurement presented in this thesis we used pp collisions data at $\sqrt{s} = 7$ TeV center of mass energy, collected between April and November 2010 with the CMS detector. Data is comprised of three periods shown in Table VIII and corresponds to a total integrated luminosity of 36.15 pb^{-1}.

Runs	HLT Trigger	Data Set	Integrated Luminosity
135821−144114	HLTMu9	/Mu/Run2010A-Nov4ReReco v1/AOD	3.18 pb^{-1}
146240−147195	HLTMu9	/Mu/Run2010B-Nov4ReReco v1/AOD	5.06 pb^{-1}
147196−149711	HLTMu15	/Mu/Run2010B-Nov4ReReco v1/AOD	27.91 pb^{-1}

TABLE VIII

DATA SAMPLES CORRESPONDING TO THE INTEGRATED LUMINOSITY OF 36 pb^{-1}.

The trigger used in this analysis is based on the presence of at least one charged muon. Because the instantaneous luminosity increased with time, the minimum transverse momentum p_T of the muon required in the trigger increased from 9 GeV (corresponding

to HLTMu9) to 15 GeV (corresponding to HLTMu15). The data is reconstructed with the 3.8.X release of CMSSW. All the samples used in this thesis, both data and simulated samples, were processed with the Physics Analysis Toolkit (PAT), which is used to create PAT-tuples. A very loose selection, including loose cuts like $p_T^{\mu} > 15$ GeV/c and $p_T^{jet} > 15$ GeV/c, was included in the PAT-tuples production, in order to reduce the size of the samples. The MC samples used in this analysis were generated with 3.8.X release of CMSSW.

5.2 Signal and Background Modeling

The simulation of $t\bar{t}$ events is performed using MADGRAPH (54). The top quark pairs are generated together with 0 to 3 hard jets. PYTHIA (55) was also used to provide the hadronization of the generated particles. The shower matching is done following the Kt-MLM prescription (54). The generated events are then passed through the GEANT4-based (56) CMS detector simulation. The top quark pair production cross section has been calculated to next-to-leading order (NLO) as $\sigma_{t\bar{t}} = 157^{+23}_{-24}$ pb (12), using MCFM (57). The uncertainty in the cross section includes the scale uncertainties, estimated by varying the factorization and renormalization scales by a factor of 2 and 0.5 around the central scale choice of 173 GeV, and the uncertainties from the PDFs and the values of α_s, following the results obtained by using the MSTW2008 (58), CTEQ6.6 (59), and NNPDF2.0 (60) sets. The uncertainties are then combined according to the PDF4LHC prescriptions (61).

The electroweak production of single top quarks is considered a background process, and is simulated using MADGRAPH. The NLO cross section of the single top t-channel is set to $\sigma_t = 64.6^{+3.4}_{-3.2}$ pb (13) using MCFM (57). The uncertainty is defined in a similar way as for the top quark pair production. For both $t\bar{t}$ and single top quark production, renormalization and factorization scales were set to $m_t = 173$ GeV. Similarly, the cross sections for tW-channel and s-channel are predicted to be $\sigma_{tW} = 10.6 \pm 0.8$ pb (14) and $\sigma_s = 4.2 \pm 0.2$ pb (15). A measurement of the t-channel cross section has been performed by CMS and is found to be consistent with the value predicted by the standard model.

W(Z) + jets processes are also simulated using MADGRAPH and QCD multi-jet samples were produced using PYTHIA. The inclusive NNLO cross section for the production of W bosons decaying into leptons is set to $\sigma_{W\to l\mu} = 31314 \pm 1558$ pb using FEWZ (62), setting renormalization and factorization scales to $m_W = 80.398$ GeV. The uncertainty was determined in a similar way as for top quark pair production. The Drell-Yan production cross section at NNLO has been calculated using FEWZ as $\sigma_{Z/\gamma^*\to ll}(m_{ll} > 50\text{GeV}) = 3048 \pm 132$ pb, where the scales were set to $m_Z = 91.1876$ GeV. In addition, a diboson sample WW was used, which was generated in a similar way as the W(Z) + jets samples, for which the production cross section has been calculated using FEWZ as $\sigma_{WW} = 43 \pm 2.6$ pb. Also, a complementary sample for W(Z) + jets, denoted VQQ, was used in order to increase the statistics for the heavy flavor processes. The leading order (LO) theoretical cross section used for the VQQ sample is $\sigma_{VQQ} = 35.8$ pb.

Table IX shows the list with all the samples used in this analysis with the specific names as they appear in the CMS database. The corresponding cross sections calculated for each sample are also shown. QCD and W + jets constitute the main background for the $t\bar{t}$ signal before tagging the b-jets. We refer to all the samples other than QCD, including the $t\bar{t}$ signal, as **W-like**.

Sample	Cross Section (pb)
/TTJets TuneD6T 7TeV-madgraph-tauola/Fall10-START38 V12-v3	157.5^{+23}_{-24}
/WJetsToLNu TuneD6T 7TeV-madgraph-tauola/Fall10-START38 V12-v1	31314 ± 1558
/DYJetsToLL TuneD6T M-50 7TeV-madgraph-tauola/Fall10-START38 V12-v2	3048 ± 132
/VQQJetsToLL TuneD6T 7TeV-madgraph-tauola/Fall10-START38 V12-v2	35.8 (LO)
/TToBLNu TuneZ2 s-channel 7TeV-madgraph/Fall10-START38 V12-v1	4.2 ± 0.2
/TToBLNu TuneZ2 t-channel 7TeV-madgraph/Fall10-START38 V12-v2	$64.6^{+3.4}_{-3.2}$
/TToBLNu TuneZ2 tW-channel 7TeV-madgraph/Fall10-START38 V12-v2	10.6 ± 0.8
/QCD Pt-20 MuEnrichedPt-15 TuneZ2 7TeV-pythia6/Fall10-START38 V12-v1	2.966×10^8
/WWtoAnything TuneZ2 7TeV-pythia6-tauola/Fall10-START38 V12-v1	43 ± 2.6

TABLE IX

MONTE CARLO SIMULATED DATA SAMPLES FOR DIFFERENT PROCESSES. THE SAMPLE NAMES ARE SPECIFIC TO CMS DATABASE. THE CORRESPONDING CROSS SECTIONS ARE SHOWN ON THE RIGHT COLUMN.

5.3 Composition of the W+jets sample

W bosons produced in association with jets are the largest irreducible background to $t\bar{t}$ production. In the W+jets MC samples the jets can have different flavors depending on the originating partons. The flavor assignment of each individual jet is done by matching the closest generator-level b or c quark to the reconstructed jet, as following:

- the jet is considered a **b-jet** , if ΔR(jet, b parton) < 0.5
- the jet is considered a **c-jet** , if ΔR(jet, c parton) < 0.5 and no b parton is found within ΔR(jet, parton) < 0.5
- the jet is considered light flavor jet if no b or c partons are found within ΔR(jet, parton) < 0.5.

The simulated W sample is constructed for this analysis by combining the W+jets and VQQ Monte Carlo simulated samples (presented in section 5.2). The number of W+heavy flavor (b and c flavor) events in the simulated W+jets sample is augmented by adding the VQQ sample. In order to remove the overlaps, a dedicated tool from CMSSW is used, named **Flavor History Filter** (63). The filter creates 11 **paths** which correspond to different processes. Paths 1-4 correspond to processes where the heavy flavor jets originate from the Matrix Element calculation, while for the processes associated to paths 5-6, they originate from the parton shower. In particular, $b\bar{b}$ events are taken from paths 1+5, b events are taken from path 2, $c\bar{c}$ events are taken from paths 3+6 and c events are taken from path 4. Light flavor events are taken from path 11.

Paths 7-10 correspond to the overlapping phase space between parton shower and matrix element calculations and their result is ambiguous. Hence we do not use these paths, meaning that we discard events with two heavy flavor partons generated close together from the Matrix Element calculation, and events with two heavy flavor partons generated far apart from the Parton Shower calculation.

The simulated samples W+jets and VQQ are combined using the prescription described in (63). The normalization of these events is taken into account, as shown in equation Equation 5.1.

$$\begin{aligned} N_{W+jets} &= \sigma_{VQQ} \times \frac{N_{VQQ}^{path1}}{N_{VQQ}} + \sigma_{Wjets} \times \frac{N_{Wjets}^{path2}}{N_{Wjets}} + \sigma_{Wjets} \times \frac{N_{Wjets}^{path5}}{N_{Wjets}} \\ &+ \sigma_{VQQ} \times \frac{N_{VQQ}^{path3}}{N_{VQQ}} + \sigma_{Wjets} \times \frac{N_{Wjets}^{path6}}{N_{Wjets}} \\ &+ \sigma_{Wjets} \times \frac{N_{Wjets}^{path4}}{N_{Wjets}} \\ &+ \sigma_{Wjets} \times \frac{N_{Wjets}^{path11}}{N_{Wjets}}, \end{aligned} \tag{5.1}$$

where σ_{VQQ} and σ_{Wjets} correspond to the theoretical cross section for the MC VQQ and W+jets sample (Table IX). $N_{VQQ}^{\text{path i}}$ and $N_{Wjets}^{\text{path i}}$ represent the number of events selected with path i for MC VQQ and W+jets, respectively. Table X shows the fraction of heavy flavor and light events in the combined W+jets simulated sample. The fractions were calculated with respect to the total number of events in the separated W+jets sample, normalized to the NNLO theoretical cross section. From the table it is observed that

only a small fraction ($\approx 2\%$) of the total number of events were discarded by not using the results from the paths 7-10.

N jets	Wb(b)	Wc(c)	W+light
≥ 1	0.008	0.160	0.806
≥ 2	0.016	0.166	0.776
≥ 3	0.025	0.162	0.746
≥ 4	0.031	0.146	0.726
$= 1$	0.007	0.159	0.813
$= 2$	0.014	0.167	0.783
$= 3$	0.024	0.166	0.751

TABLE X

FRACTION OF Wb(b), Wc(c) AND W+LIGHT JETS EVENTS IN THE W+JETS COMBINED SAMPLE.

5.4 Event Selection

The good quality of the data is ensured by a series of conditions, described below. A scraping event filter is applied to the data in order to reject interactions from the beam halo. The filter removes events with less than 10 good tracks with $p_T > 250$ MeV/c. In addition, a cleaning procedure to reject anomalous HCAL noise is applied to the data through the HBHENoiseFilter. The filter removes the HCAL noise which is not due to

the electronics (pedestal) noise, but rather due to instrumentation issues associated with HPDs and readout boxes.

The signal events are selected by requiring that they pass the muon trigger HLTMu9 or HLTMu15, as described in 5.1. In addition, the presence of a good primary vertex within the event is required. For our purposes, a good primary vertex must satisfy:

- the primary vertex is not fake
- the number of degrees of freedom $ndof > 4$
- the projection on the z axis of the primary vertex is within 24 cm from the center of the detector: $|z| < 24$ cm
- $\rho = \sqrt{(x^2 + y^2)} < 2.0$ cm, where (x,y) represents the primary vertex coordinates on the (x,y) plane.

The signature of a top quark pair event in the μ+jets channel includes one isolated muon, at least 3 jets and a neutrino. The selected muon must satisfy the requirements described below:

1. The muon must be both Global Muon and Tracker Muon, as described in section 4.4.1
2. $p_T^{\mu} > 20$ GeV/c. As can be seen in Figure 33 a large fraction of the QCD sample is removed by this cut.
3. The muon is required to be within the Tracker and Muon Detector acceptance: $|\eta^{\mu}| < 2.1$

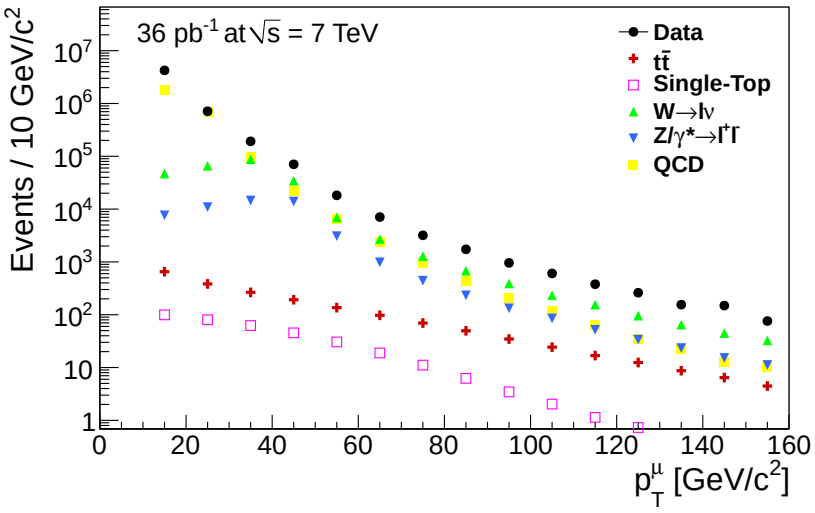

Figure 33. Transverse momentum of the muon for different samples. A preselection cut $p_T^{\mu} > 10$ GeV/c was applied in the PAT-tuples production. The events passed the selection cuts regarding the primary vertex and the Global and Tracker Muon.

4. The global track associated with the muons satisfies normalized $\chi^2 < 10$ and the number of valid muon hits > 0

5. The inner track associated with the muon must have the number of valid hits associated to it > 10

6. The inner track of the muon must contain at least one pixel hit

7. The muon must originate from the primary vertex, which means that the distance between the z vertex of the inner track of the muon (z) and the z coordinate of the primary vertex (PVz) satisfies: $|z - PVz| < 1$ cm

8. The segments matched to a global muon must be part of at least two muon stations

9. The 2D impact parameter calculated with respect to the beam spot must satisfy $|d_0| < 0.02$ cm. In Figure 34 it can be seen that a large number of muons from the QCD sample have large impact parameters, while most of the muons from the signal sample have impact parameters below 0.02 cm.

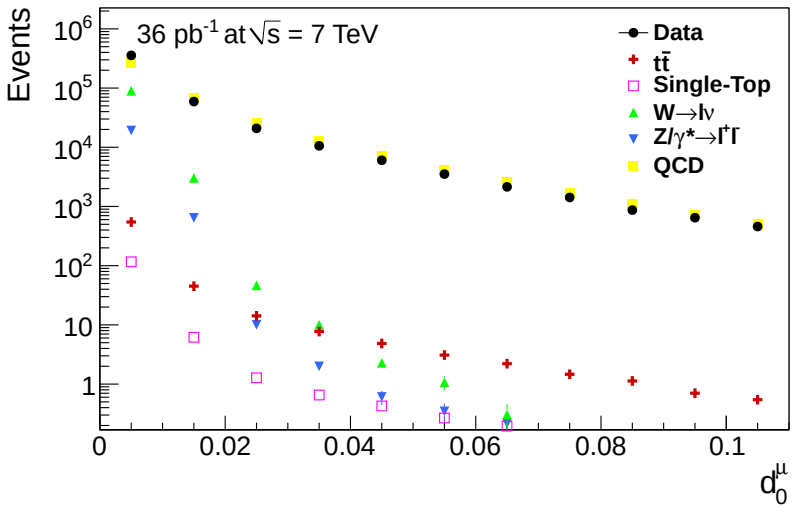

Figure 34. Impact parameter of the muon calculated with respect to the beam spot for different samples. The events passed the selection cuts regarding the primary vertex, the Global and Tracker Muon, the muon acceptance, the muon p_T, the muon χ^2, the number of hits for the muon tracks, the number of segments for the muon and the distance from the muon z and PVz.

The main difference between muons originating from the W boson decay and those from the semi-leptonic decay of heavy quarks is that the former tend to be more isolated, while the latter appear within the cone of the heavy quark jets. Thus, asking for a muon to be isolated will reduce considerably the QCD background, where most of the muons are found within the jet cone. The muon isolation requirements used in this thesis were based on both geometrical properties and the energy deposited in the detector. The **combined relative isolation** defined as:

$$I^{\mu}_{rel} = \frac{I_{Trk} + I_{Ecal} + I_{Hcal}}{p^{\mu}_{T}}, \tag{5.2}$$

must satisfy $I^{\mu}_{rel} < 0.1$. I_{Trk}, I_{Ecal} and I_{Hcal} are isolation variables defined in the tracker, electromagnetic calorimeter and hadronic calorimeter, respectively. The tracker isolation variable is defined as the sum of all the tracks within a sold angle cone of $\Delta R(\eta, \phi) > 0.3$ around the muon

$$I_{Trk} = \sum_{\substack{i \neq j \\ \Delta R < 0.3}} (p^{track}_{T,i}), \tag{5.3}$$

where j is the track associated to the muon. The calorimeter isolation variables are defined as sums of the energy deposited within a solid angle cone of $\Delta R(\eta, \phi) > 0.3$ around the muon

$$I_{Ecal} = \sum_{\substack{i \\ \Delta R < 0.3}} (E^{Ecal}_{T,i}),\ I_{Hcal} = \sum_{\substack{i \\ \Delta R < 0.3}} (E^{Hcal}_{T,i}). \tag{5.4}$$

Figure 35 shows the combined relative isolation for data and different simulated samples. It can be seen that muons from most of the QCD events have $I^{\mu}_{rel} > 0.1$, while muons from the signal events have lower I^{μ}_{rel}.

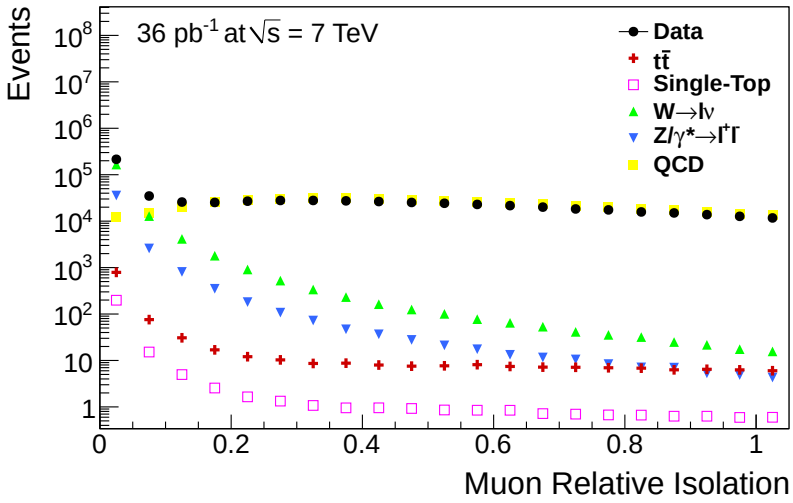

Figure 35. Combined relative isolation of the muon for different samples. The events passed the selection cuts regarding the primary vertex, the Global and Tracker Muon and all the muon selection cuts, except the isolation and $\Delta R(\mu, jet)$.

We also required that the **distance between the muon and the closest jet** should be $\Delta R(\mu, jet) > 0.3$, where $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. From the distribution of $\Delta R(\mu, jet)$ presented in Figure 36 it can be seen that most of the muons from QCD processes have at least one jet within $\Delta R = 0.3$, while most of the muons from W-like processes appear

further from the jets. The figure also shows that most of the data events in the first bin are QCD events.

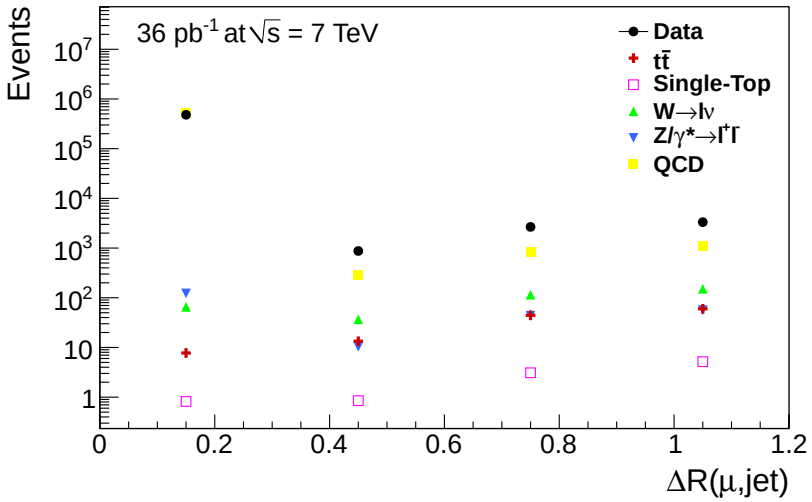

Figure 36. The distribution of $\Delta R(\mu, jet)$ for different samples. The events passed the selection cuts regarding the primary vertex, the Global and Tracker Muon and all the muon selection cuts, except the $\Delta R(\mu, jet)$. Most of the data events with low $\Delta R(\mu, jet)$ (corresponding to the first bin) are QCD events.

$t\bar{t}$ signal events in the μ+jets channel include exactly one isolated muon. The presence of a second isolated muon would most likely be the result of a different process. As a consequence, we remove events with a second muon satisfying the following conditions:

- the muon is a Global Muon

- $p_T^{\mu} > 10$ GeV
- $|\eta^{\mu}| < 2.5$
- the combined relative isolation $I_{rel}^{\mu} < 0.2$.

Likewise, we also remove events containing an electron satisfying:

- the transverse energy of the electron $E_T^e > 15$ GeV
- the electron is found within the detector acceptance $|\eta^e| < 2.5$
- the electron is isolated:

$$\frac{\sum_{i,\Delta R<0.3}(p_{T,i}^{track}) + \sum_{i,\Delta R<0.3}(E_T^{ECal,RecHit_i}) + \sum_{i,\Delta R<0.3}(E_T^{HCal,Tower_i})}{E_T^e} < 0.2. \tag{5.5}$$

The neutrino is reconstructed as the missing transverse energy using the particle flow algorithm as described in section 4.8. Together with the muon, the neutrino is a decay product of one of the two W bosons present in a top quark pair event. Hence the missing transverse energy is expected to be quite large for the W-like events and small in general for the QCD events. From Figure 37 it can be observed that many QCD events have missing transverse energy below 20 GeV, while the $t\bar{t}$ events have missing transverse energy above 20 GeV. Hence, a cut on missing transverse energy $\not{E}_{\mathrm{T}} > 20$ GeV also reduces the QCD background.

Jets used in this analysis were reconstructed using the particle flow algorithm as described in section 4.6 and must satisfy the following requirements:

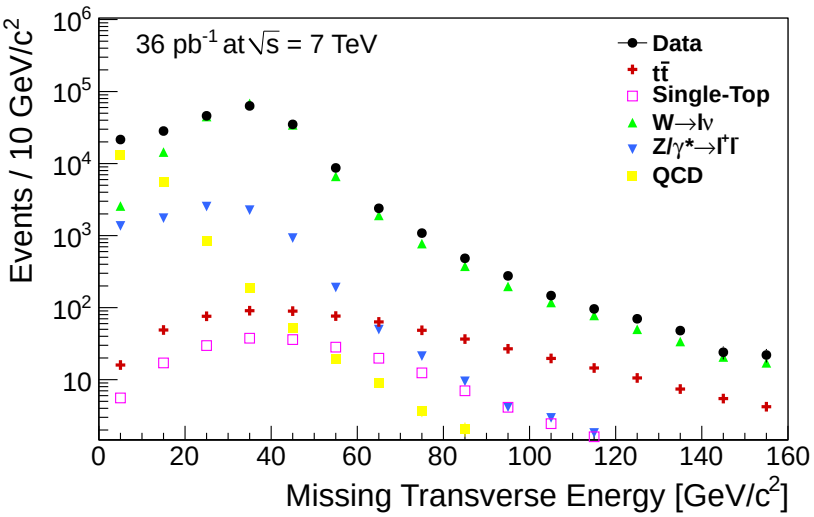

Figure 37. Missing transverse energy for different samples. The events passed the selection cuts regarding the primary vertex, the Global and Tracker Muon, all the muon selection cuts, the cuts for veto events with electrons and events with more than one muon.

- jet transverse momentum must satisfy: $p_T^{jet} > 25$ GeV
- jet must be within the detector acceptance $|\eta^{jet}| < 2.4$
- electromagnetic fraction > 0.01
- number of RecHits containing 90% of the jet energy > 1
- fraction of energy in the hottest HPD readout < 0.98

The number of events passing the selection criteria in data and MC are shown in Table XI and Table XII, respectively.

Cut	Number of events
Clean Filters	49034698
HLT	7727624
Good PV	7718212
One Isolated Muon	214368
Veto Loose Muon	208313
Veto Electron	207536
$\not{E}_{\mathrm{T}} > 20$ GeV	157654
=1 jet	20012
=2 jet	4506
=3 jet	1111
$\geq$4 jets	459

TABLE XI

NUMBER OF EVENTS PASSING THE SELECTION FOR DATA WITH AN INTEGRATED LUMINOSITY OF 36 pb^{-1}.

Cut	$t\bar{t}$	$W \to l\nu$	$Z \to l^+l^-$	QCD	t channel	tW channel	s channel	WW	Total
Processed	5670.0 ± 5.0	1127304.0 ± 293.0	109728.0 ± 68.8	3048454.8 ± 561.2	753.5 ± 1.1	381.6 ± 0.5	54.0 ± 0.1	1548.0 ± 1.1	4293893.9 ± 636.9
HLT	1784.8 ± 2.8	253269.6 ± 138.9	34376.0 ± 38.5	2586537.3 ± 517.0	261.7 ± 0.6	104.3 ± 0.3	19.7 ± 0.0	272.2 ± 0.5	2876625.7 ± 536.7
Good PV	1784.4 ± 2.8	252315.1 ± 138.6	34301.6 ± 38.5	2585475.4 ± 516.8	261.5 ± 0.6	104.3 ± 0.3	19.7 ± 0.0	271.7 ± 0.5	2874533.6 ± 536.5
One Iso mu	765.1 ± 1.8	174836.8 ± 115.4	14210.1 ± 24.8	19928.9 ± 45.4	157.3 ± 0.5	54.0 ± 0.2	10.7 ± 0.0	184.3 ± 0.4	210147.2 ± 126.5
Loose mu veto	741.7 ± 1.8	174829.4 ± 115.4	9283.8 ± 20.0	19906.5 ± 45.4	156.9 ± 0.5	52.4 ± 0.2	10.7 ± 0.0	179.4 ± 0.4	205160.8 ± 125.6
Electron veto	643.0 ± 1.7	174440.2 ± 115.2	9063.8 ± 19.8	19735.6 ± 45.2	153.8 ± 0.5	44.9 ± 0.2	10.5 ± 0.0	158.8 ± 0.3	204250.6 ± 125.4
MET> 20 GeV	578.4 ± 1.6	157490.0 ± 109.5	5968.4 ± 16.0	1121.5 ± 10.8	136.8 ± 0.5	39.4 ± 0.2	9.3 ± 0.0	137.1 ± 0.3	165480.8 ± 111.3
1 jet	32.6 ± 0.4	19209.8 ± 38.2	665.6 ± 5.4	647.8 ± 8.2	51.8 ± 0.3	6.3 ± 0.1	2.9 ± 0.0	51.7 ± 0.2	20668.5 ± 39.5
2 jets	116.2 ± 0.7	3900.6 ± 17.2	148.5 ± 2.5	185.9 ± 4.4	47.8 ± 0.3	13.1 ± 0.1	4.2 ± 0.0	36.9 ± 0.2	4453.1 ± 18.1
3 jets	184.6 ± 0.9	680.8 ± 7.2	31.8 ± 1.2	41.5 ± 2.1	18.6 ± 0.2	12.2 ± 0.1	1.3 ± 0.0	9.9 ± 0.1	980.7 ± 7.7
≥4 jets	241.9 ± 1.0	180.2 ± 3.7	9.8 ± 0.7	8.3 ± 0.9	7.2 ± 0.1	6.8 ± 0.1	0.3 ± 0.0	2.4 ± 0.0	456.9 ± 4.1

TABLE XII

NUMBER OF EVENTS PASSING THE SELECTION FOR MC SIMULATED SAMPLES NORMALIZED USING THEORETICAL CROSS SECTIONS TO 36 PB^{-1}

5.5 Measurement of the muon selection efficiencies

The muon efficiency is factorized into a combined efficiency of muon identification (ID) and isolation (ISO), and trigger efficiency. The former is the probability for a reconstructed muon to pass the muon quality selection, the latter is the probability for a muon which has already passed ID and ISO cuts to fire a certain single muon trigger.

The muon efficiency was measured on the $Z \rightarrow \mu\mu$ data sample using the Tag and Probe data driven method. The tag muon is a high quality muon matched to a trigger object. The probe muon is the one under investigation. We also require that the tag muon and the probe muon are oppositely charged, and invariant mass of the two is within 15 GeV/c^2 of the Z boson mass, which ensures the high purity of the $Z \rightarrow \mu\mu$ sample. The Z mass peak is fitted with a Gaussian from which we subtract the background fitted with an exponential function. The fitting is performed before and after the cuts, so the ratio of the number of $Z \rightarrow \mu\mu$ events obtained from fitting will represent the muon efficiency. The same procedure was repeated for both collision data and Z+jets MC. The ID and ISO efficiency is shown in Figure 38 and Figure 39 as the function of several kinematic variables.

The ratio between the results for Data and MC represents the scale factors (SF). Since the SF are found to be almost constant as a function of the kinematic variables, a single correction factor is used for the measurement presented in this thesis. The muon selection efficiencies for both data and simulated sample and the scale factors together

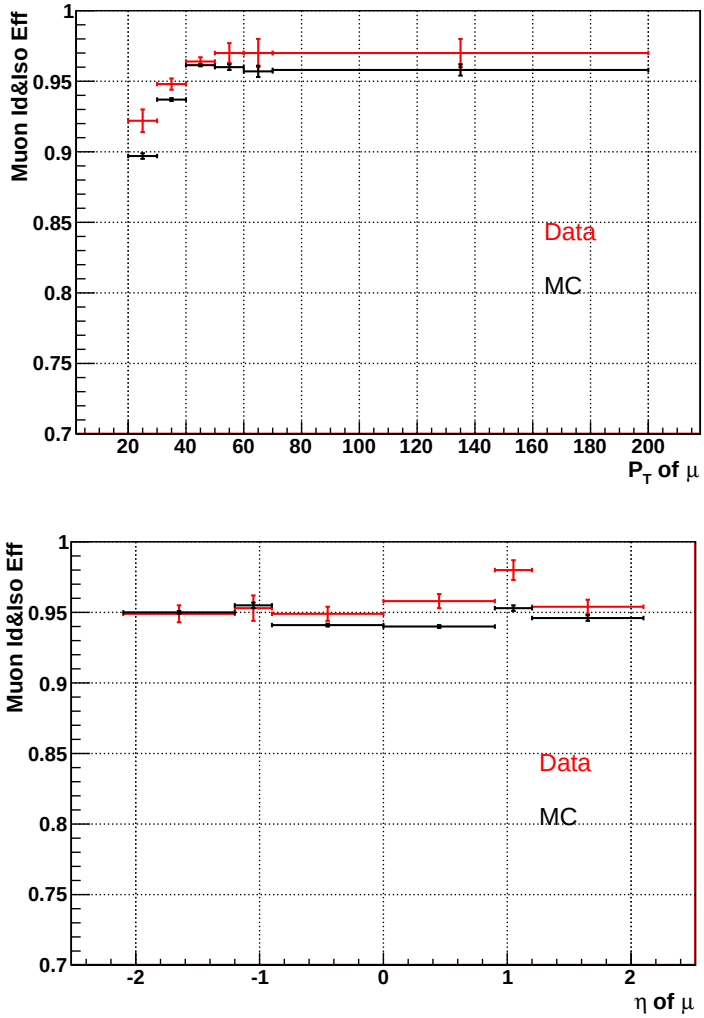

Figure 38. Muon ID and ISO efficiencies as functions of the transverse momentum of the muon p_T^μ (top) and the muon position η^μ (bottom).

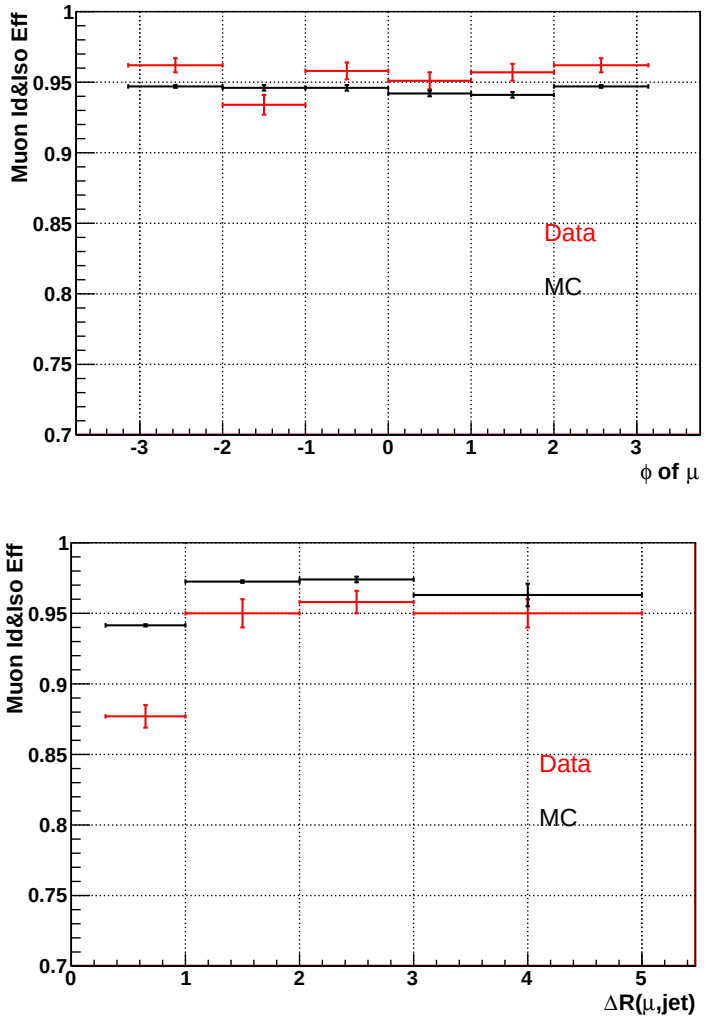

Figure 39. Muon ID and ISO efficiencies as functions of ϕ^{μ} (top) and the distance between the muon and the closest jet $\Delta R(\mu, jet)$ (bottom).

with the statistical uncertainties are summarized in Table XIII. The average values for the muon selection efficiencies and the SF are shown for both the inclusive sample and the sample with at least two jets.

	Trigger Efficiency	ID and ISO Efficiency (≥ 0 jets)	ID and ISO Efficiency (≥ 2 jets)
Data	$0.92 \pm 0.3\%$	$0.96 \pm 0.2\%$	$0.97 \pm 1.0\%$
MC	$0.93 \pm 0.1\%$	$0.94 \pm 0.1\%$	$0.96 \pm 0.2\%$
Scale Factor	$0.99 \pm 0.2\%$	$1.01 \pm 0.1\%$	$1.01 \pm 1.0\%$

TABLE XIII

MUON EFFICIENCIES IN DATA AND MONTE CARLO SIMULATION, AND THE SCALE FACTORS. THE VALUES FOR ID AND ISO EFFICIENCIES ARE SHOWN FOR BOTH THE INCLUSIVE SAMPLE AND ≥ 2 JETS. THE UNCERTAINTIES SHOWN ARE STATISTICAL ONLY.

The efficiency and the scale factor versus $\Delta R(\mu, jet)$ for the sample with at least two jets are shown in Figure 40. The statistics is reduced in this case, but overall we can observe that the scale factor is constant as a function of $\Delta R(\mu, jet)$ within the statistical uncertainties, therefore the fixed values from Table XIII were used in this analysis.

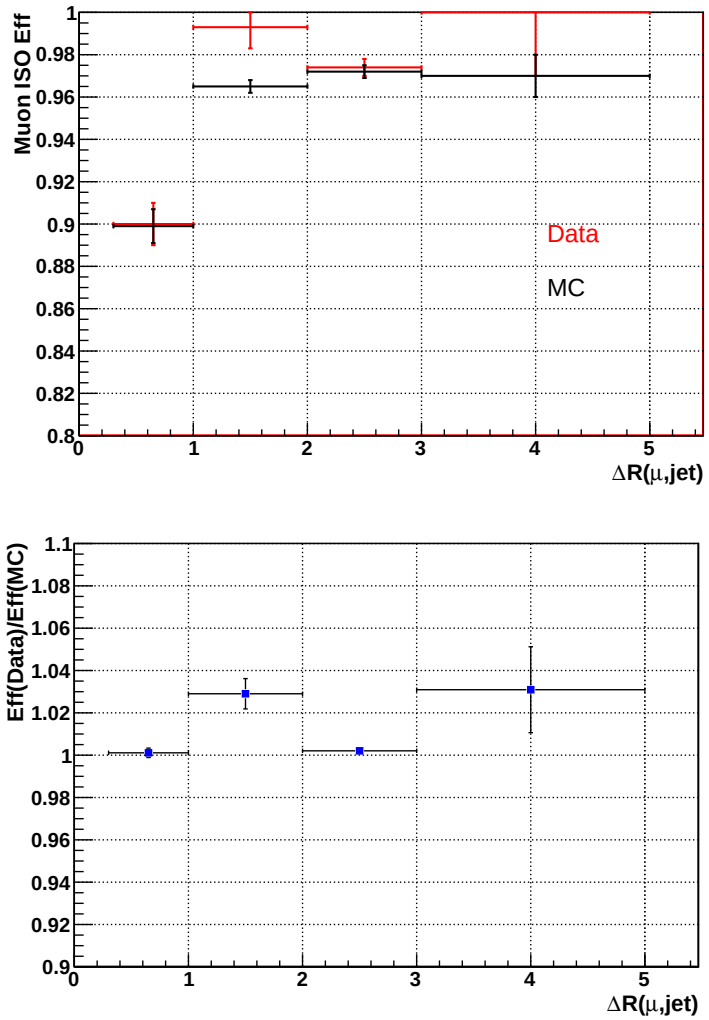

Figure 40. Muon ID and ISO efficiencies (top) and the corresponding scale factors (bottom) as a function of the distance between the muon and the closest jet $\Delta R(\mu, jet)$ for the sample with at least two jets.

CHAPTER 6

BACKGROUND ESTIMATION IN THE PRE-TAGGED DATA SAMPLE

The backgrounds to the $t\bar{t}$ muon+jets sample are QCD, W/Z+jets, single top and diboson processes. Among all the background processes, QCD and W+jets are the largest. Since the understanding of the perturbative strong interaction is limited, the QCD background is taken directly from data. The W+jets contribution is normalized using a data driven method. The electroweak processes are modeled using MC.

6.1 Estimation of the QCD background in the pre-tagged data sample

The background characterized as QCD represents the largest background before any selection is applied, but is greatly reduced due to very low efficiency for passing the selection criteria listed in 5.4. In particular, the muons present in the QCD events are failing the isolation requirements applied in the selection.

The QCD multijet background is estimated with the Matrix Method, which is a data driven method that relies minimally on MC simulation.

6.1.1 Matrix Method

The Matrix Method relies on two data sets: the **tight** sample that corresponds to the preselected sample and contains N_T events, and the **loose** sample that consists of N_L events that pass the selection described in 5.4 from which the combined relative isolation requirement on the muon was removed. Hence, the tight selected sample is a

subset of the loose selected sample. The events with leptons originating from a W decay are considered to be the W-like events and their number is denoted by the superscript **W-like**. The number of events originating from QCD multijet production is denoted by the superscript **QCD**. N_L and N_T can be written as:

$$\begin{aligned} N_L &= N^{W-like} + N^{QCD}, \\ N_T &= \varepsilon_{W-like} N^{W-like} + \varepsilon_{QCD} N^{QCD}. \end{aligned} \tag{6.1}$$

In Equation 6.1, ε_{W-like} is the efficiency for a loose muon from a W decay to pass the tight criteria; it is measured in the W-like sample obtained by combining the W, Z, $t\bar{t}$ and single top samples. The measured value for ε_{W-like} is corrected by the SF derived from $Z \to \mu\mu$ events as described in 5.5. We assigned an uncertainty of 2% on ε_{W-like}, which is a more conservative estimate than the one obtained from the Tag and Probe method. ε_{QCD} is the (fake) rate at which a loose muon in the QCD multijet sample passes the tight criteria; it is measured in a low missing transverse energy ($\not\!\!E_T$) data sample, that is dominated by QCD multijet events.

The linear system in Equation 6.1 can be solved for N^{QCD} and N^{W-like}. The resulting equations are :

$$\begin{aligned} N^{W-like} &= \frac{N_T - \varepsilon_{QCD} N_L}{\varepsilon_{W-like} - \varepsilon_{QCD}}, \\ N^{QCD} &= \frac{\varepsilon_{W-like} N_L - N_T}{\varepsilon_{W-like} - \varepsilon_{QCD}}. \end{aligned} \tag{6.2}$$

The number of W-like events in the preselected sample is obtained as $N_T^{W-like} = \varepsilon_{W-like} N^{W-like}$, and the number of QCD multijet events as $N_T^{QCD} = \varepsilon_{QCD} N^{QCD}$. A detailed description of the error calculation is described in (64).

The method relies on the fact that the fake rate measured in the QCD sample is independent of the $\not\!\!E_T$ in the event. Figure 41 shows the fake rate distribution for the MC simulated QCD sample for events with two or more jets. The distribution is quite flat with respect to $\not\!\!E_T$, the fluctuations observed at higher values of $\not\!\!E_T$ are due to the lack of statistics. The uncertainty lines represent the total uncertainty assigned to ε_{QCD}, as described in section 6.1.2.

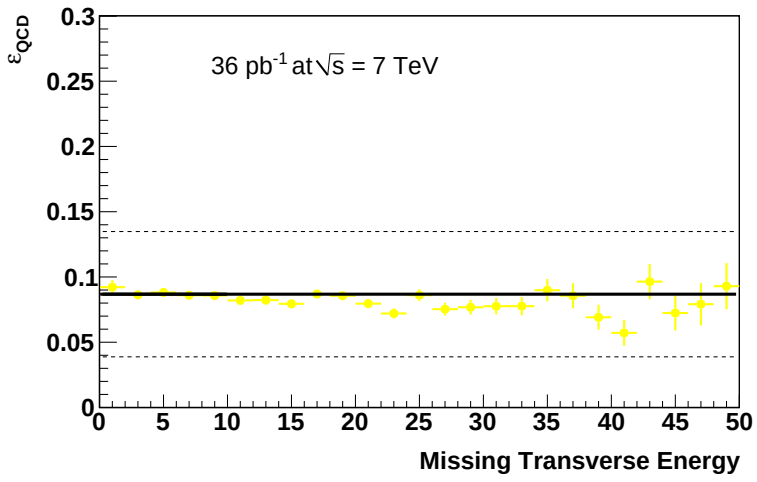

Figure 41. ε_{QCD} for QCD sample as a function of $\not\!\!E_T$. The uncertainty lines represent the total uncertainty assigned to ε_{QCD}.

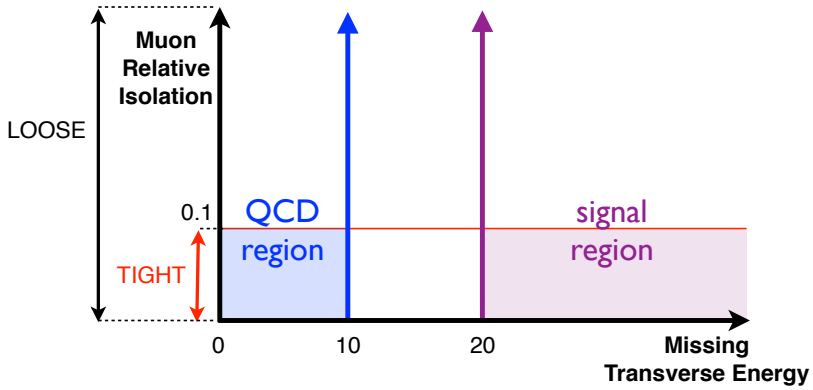

Figure 42. ε_{QCD} is measured in the QCD region ($\not\!\!E_T < 10$ GeV), while the estimation of the QCD multijet background is done in the signal region ($\not\!\!E_T > 20$ GeV). The tight sample is a subset of the loose sample, consisting of those events with muon combined Relative Isolation < 0.1.

In order to minimize the signal contamination, the region corresponding to $\not\!\!E_T < 10$ GeV was chosen to measure ε_{QCD}. Figure 42 shows the region where we measure ε_{QCD}, labeled **QCD region**, as having $\not\!\!E_T < 10$ GeV, and the region where we estimate the QCD multijet background in the final sample, labeled **signal region** having $\not\!\!E_T > 20$ GeV. This cartoon also indicates the fact that the muon combined Relative Isolation < 0.1 is the requirement which represents the difference between the loose and the tight samples (as used in Equation 6.1).

Figure 43 show the two dimensional distribution of the events with two jets or more, selected as a function of the muon combined Relative Isolation and $\not\!\!E_T$ for $t\bar{t}$, QCD and

W+jets MC, respectively. Figure 43 shows that the main contribution in the QCD region comes from QCD processes.

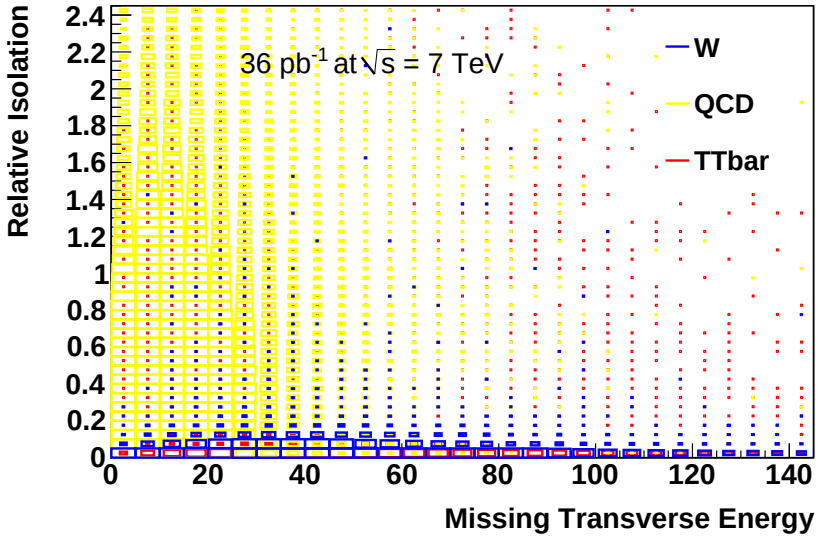

Figure 43. Two dimensional distribution for $t\bar{t}$, QCD and W+jets MC events with two jets or more, shown as a function of the muon combined Relative Isolation and $\not\!\!E_T$. The main contribution in the QCD region comes from QCD processes.

The measured ε_{QCD} proved to be sensitive to the threshold used to define the QCD dominated region. We attribute this behavior to the fact that some W and Z events are present in the $\not\!\!E_T < 10$ GeV region, which we refer to as **signal contamination**. The choice of the threshold is limited by the amount of statistics. In order to obtain a cleaner

QCD sample, we subtract the signal contamination present in the region $\not\!\!E_T < 10$ GeV using the predictions given by the MC.

The values of ε_{QCD} measured for each jet multiplicity bin, before and after subtracting the W(Z) amount from the $\not\!\!E_T < 10$ GeV region are summarized in the Table XIV. It can be observed that the values of ε_{QCD} corresponding to the 2 and 3 exclusive jet bins and the 4 or more inclusive jet bin are comparable within statistical uncertainties. We therefore use ε_{QCD} measured for the 2 or more jets bin to estimate the background in the 2, 3 and 4 or more jet bins. The topology for events with 1 jet is very different and we use the ε_{QCD} obtained in the 1 jet bin sample to estimate the background for those events.

Njets	default	W(Z) subtracted
$= 1$	0.23 ± 0.00	0.19 ± 0.00
≥ 2	0.21 ± 0.01	0.151 ± 0.01
$= 2$	0.21 ± 0.01	0.15 ± 0.01
$= 3$	0.21 ± 0.02	0.14 ± 0.02
≥ 4	0.24 ± 0.04	0.17 ± 0.04

TABLE XIV

ε_{QCD} VALUES MEASURED BEFORE SUBTRACTING THE W(Z) CONTRIBUTION (DENOTED AS DEFAULT) AND AFTER SUBTRACTING THE W(Z) CONTRIBUTION (DENOTED AS W(Z) SUBTRACTED) FROM THE $\not\!\!E_T < 10$ GeV REGION.

6.1.2 Matrix Method Results

The estimated number of W-like events and QCD multijet events obtained with this method are summarized in Table XV for the samples with 1, 2, 3 and 4 or more jets. The results are for a data sample with a total luminosity of 36 pb^{-1} . The signal contamination, calculated using the MC samples for W and Z, is subtracted from the data in the $\not\!\!E_T < 10$ GeV region. By measuring the ε_{QCD} in events with 2 or more jets, we are confident that the contribution from $t\bar{t}$ and single top in this region is negligible. An uncertainty of 2% is assigned to ε_{Sig}, as explained in 6.1.1. As shown in Table XIV, ε_{QCD} values are affected by W-like contamination in the QCD region. We therefore assign a systematic uncertainty as a difference between the values of ε_{QCD} measured with and without subtracting the W(Z) contribution to the QCD region. This uncertainty also covers the differences between the calculated ε_{QCD} for each jet bin and the value used in the Matrix Method for the events with 2, 3 and 4 or more jets. Standard error propagation, based on Equation 6.2, is used to derive the uncertainties on the number of W-like and QCD multijet events.

Table XVI shows the scale factors obtained for each sample depending on the number of jets. The scale factors were calculated for both QCD and W-like events, as being the ratio between the number of events estimated in data with the Matrix Method and the number of events predicted by MC.

Njets	N^{W-like}	N^{QCD}	ε_{sig}	ε_{QCD}
1	18358 ± 483	1654 ± 461	0.98 ± 0.02	$0.19 \pm 0.00 \pm 0.04$
2	4113 ± 192	393 ± 180	0.97 ± 0.02	$0.15 \pm 0.01 \pm 0.06$
3	1003 ± 60	108 ± 49	0.97 ± 0.02	$0.15 \pm 0.01 \pm 0.06$
≥ 4	426 ± 26	33 ± 15	0.96 ± 0.02	$0.15 \pm 0.01 \pm 0.06$

TABLE XV

NUMBER OF W-LIKE AND QCD MULTIJET EVENTS IN THE CANDIDATE SAMPLE AS PREDICTED BY THE MATRIX METHOD. THE ε_{QCD} VALUE MEASURED FOR THE EVENTS WITH 2 OR MORE JETS IS USED FOR THE EVENTS WITH 2, 3 AND 4 OR MORE JETS. BOTH STATISTICAL AND SYSTEMATIC UNCERTAINTIES ARE SHOWN FOR ε_{QCD} .

Njets	SF_{W-like}	SF_{QCD}
1	0.92 ± 0.01	2.54 ± 0.12
2	0.97 ± 0.02	2.10 ± 0.19
3	1.08 ± 0.05	2.57 ± 0.47
≥ 4	0.96 ± 0.07	3.98 ± 1.55

TABLE XVI

SCALE FACTORS FOR THE ESTIMATED NUMBER OF QCD MULTIJET EVENTS, DENOTED SF_{QCD} AND W-LIKE EVENTS, DENOTED SF_{W-LIKE} CALCULATED WITH RESPECT TO MC PREDICTIONS.

For the measurement presented in this thesis, the estimation of the QCD multijet background is taken from the results of the Matrix Method presented in Table XV.

6.2 Estimation of the W+jets background in the pre-tagged data sample

The Berends-Giele Scaling Method relies on estimating the amount of W+jets in the events with ≥ 1 jets and events with ≥ 2 jets and use this information to estimate the amount of W+jets in events with ≥ 3 jets. In each jet bin, the amount of W+jets is computed as the difference between the data and the MC predictions for the Z+jets, single top and diboson samples and the Matrix Method estimations for the QCD background, using Equation 6.3:

$$\begin{aligned} N_{W+jets}^{pretagged} &= N_{data}^{pretagged} \\ &- N_{QCD,data-driven}^{pretagged} \\ &- N_{Z+jets,MC}^{pretagged} - N_{singleTop,MC}^{pretagged} - N_{WW,MC}^{pretagged} \\ &- N_{t\bar{t},MC}^{pretagged}. \end{aligned} \tag{6.3}$$

Table XVII shows the predictions for both Data and MC samples in different jet multiplicity bins. Figure 44 shows the number of W events as a function of number of jets. As can be seen, data and MC show a similar behavior.

jetBin	N_{data}	$N_{t\bar{t}}^{MC}$	N_{Z+jets}^{MC}	$N_{singleTop}^{MC}$	N_{WW}^{MC}	$N_{QCD}^{data-driven}$
≥ 1	26086.00	577.6±1.6	859.3±6.1	205.3±0.8	101.3±0.3	2188.0±497.5
≥ 2	6076.00	544.9±1.5	190.9±2.9	130.1±0.6	49.4±0.2	534.0±187.2
≥ 3	1570.00	428.3±1.4	41.8±1.4	53.3±0.4	12.3±0.1	141.0±51.2
≥ 4	459.00	242.9±1.0	9.9±0.6	16.4±0.2	2.4±0.0	33.0±15.0

TABLE XVII

PREDICTED NUMBER OF EVENTS FOR DATA AND MC SAMPLES FOR DIFFERENT JET MULTIPLICITY BINS. THE AMOUNT OF QCD IS ESTIMATED WITH THE MATRIX METHOD, WHILE $t\bar{t}$ AND THE SMALL BACKGROUNDS ARE NORMALIZED TO THE CORRESPONDING THEORETICAL CROSS SECTIONS.

To verify that the number of events decreases at a constant rate when increasing the number of jets, the ratio between the number of events with W+n jets and the number of events with W+(n+1) jets is calculated as shown in Equation 6.4,

$$C(n) = \frac{N_W^{njets}}{N_W^{n+1jets}}. \tag{6.4}$$

Figure 45 shows the values for the ratios C(n) as a function of number of jets. As can be observed, C(n) increases with the number of jets for the W+jets estimated from data. We attribute this increase to the effect of the signal contamination in larger jet multiplicity bins. The C(n) for the simulated W+jets sample is constant versus number of jets. The values of C(1) agree between data and MC.

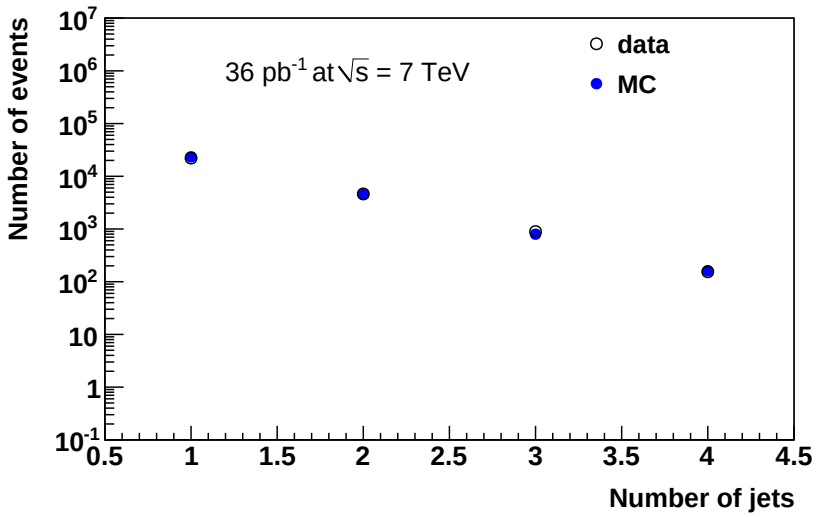

Figure 44. Number of events for both combined simulated W+jets sample and the estimated W+jets sample using equation Equation 6.4 as a function of the number of jets.

We use the value of C(1) measured in data together with the number of events in the first jet bin (where the signal is negligible) to estimate the amount of W+jets in the events with 3 or more jets and 4 or more jets, as shown in Equation 6.5:

$$\begin{aligned} N^{pretagData}_{W+jets,\geq 3} &= \frac{N^{pretagData}_{W+jets,\geq 1}}{C(1)C(2)} \\ N^{pretagData}_{W+jets,\geq 4} &= \frac{N^{pretagData}_{W+jets,\geq 1}}{C(1)C(2)C(3)}, \end{aligned} \tag{6.5}$$

where the values C(2) and C(3) were taken to be equal to C(1): C(2)=C(3)=C(1). In order to account for the differences observed between the values of C(n) calculated from

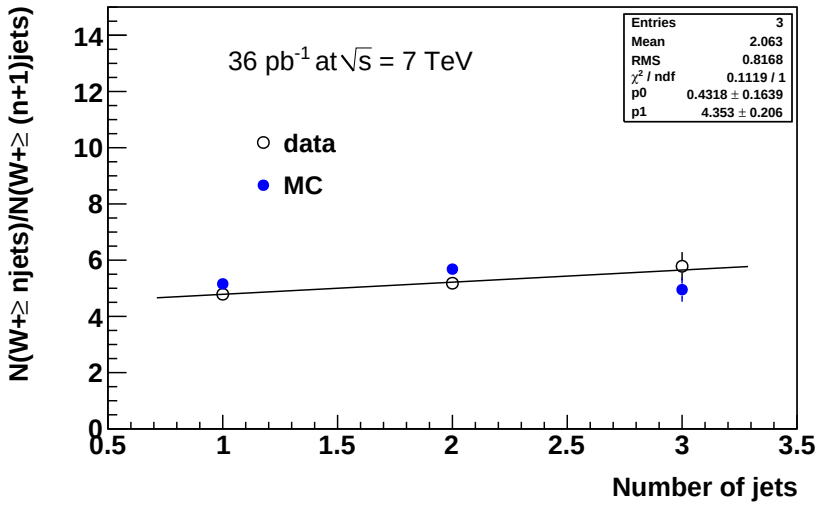

Figure 45. C(n) for data and MC as a function of the number of jets. The values measured from data are fitted with a linear function.

data as a function of the number of jets, we fit their distribution shown in Figure 45 with a linear function and extract new C(n) values from the fit. The differences between these values and C(1) were used to assign the systematic uncertainty on N_{W+jets} .

The effect of subtracting the signal contribution as given by the theoretical cross section is studied by scaling up and down the $t\bar{t}$ theoretical cross section by 30%. This will result in a systematic uncertainty on the measured $\sigma_{t\bar{t}}$ which is rather small, due to the limited amount of $t\bar{t}$ events in the first two jet bins, as shown in Section 8.1.

The estimated number of W+jets events with ≥ 3 and ≥ 4 jets are shown in Table XVIII. The uncertainty is dominated by the systematics which is one-sided $(-)$, since all the values for C(n) extracted from the fit are larger than C(1).

N jets	$N_{W+jets}^{\text{pretag Data}}$	$N_{W+jets}^{\text{pretag MC}}$
≥ 3	967.5−80.1	799.7 ± 7.8
≥ 4	202.2−45.1	161.7 ± 3.5

TABLE XVIII

ESTIMATED NUMBER OF EVENTS FOR THE W+JETS PRETAGGED SAMPLE WITH ≥ 3 AND ≥ 4 JETS. THE UNCERTAINTY IS DOMINATED BY THE SYSTEMATICS AND ONE-SIDED $(-)$, SINCE ALL THE VALUES FOR C(N) EXTRACTED FROM THE FIT ARE LARGER THAN C(1). THE COMPARISON WITH THE PREDICTION IS ALSO SHOWN.

6.3 Other backgrounds

The contributions from the low rate electroweak physics backgrounds to the preselected sample are estimated from MC. The following processes are considered:

- **diboson production (denoted as VV)**: $WW \rightarrow \mu + jets$, $WZ \rightarrow \mu + jets$
- **singe top production** in the s, t and tW channels
- **Z+jets**: $Z \rightarrow \mu + jets$.

For a given process i, the number of events before tagging is determined as

$$N_i^{presel} = \sigma_i \epsilon_i^{presel,tot} BR_i \mathcal{L}, \tag{6.6}$$

where σ_i, BR_i and $\mathcal{L}$ represents the cross section(Section 5.2), BR and the integrated luminosity, respectively. The total preselection efficiency $\epsilon_i^{presel,tot}$ for the process i is defined as

$$\epsilon_i^{presel,tot} = \epsilon_i^{trig} \epsilon_i^{presel}, \tag{6.7}$$

where ϵ_i^{trig} and ϵ_i^{presel} are the trigger and the preselection efficiencies, respectively. The trigger efficiency is the marginal efficiency only for events that pass the preselection, and is obtained by folding into the MC the per-muon and per-jet trigger efficiencies measured in data. The preselection efficiency is entirely determined from MC with the appropriate scale factors applied. The small background contributions in the pretagged sample are presented in Table XIX.

6.4 Composition of the Pre-tagged Data Sample

Table XIX shows the number of events for data and the predicted number of events for MC samples before tagging. QCD is scaled to the Matrix Method predictions. W+jets is normalized to data using the Berends Scaling Method. The small backgrounds (Z+jets, diboson and single top) are normalized to the corresponding theoretical cross sections,

as presented in section 5.2. The $t\bar{t}$ MC sample normalized to the NNLO theoretical cross section is also shown, for comparison.

Njets	N_{data}	$N_{t\bar{t}}^{MC}$	N_{W+jets}^{MC}	N_{Z+jets}^{MC}	$N_{singleTop}^{MC}$	N_{VV}^{MC}	$N_{QCD}^{data-driven}$
$= 1$	20010.00	32.6±0.4	17527.8±35.5	665.6±5.4	75.3±0.4	51.9±0.1	1654.0±461.0
$= 2$	4506.00	116.2±0.7	3733.4±17.2	148.5±2.5	76.7±0.5	37.1±0.1	393.0±180.0
$= 3$	1111.00	184.6±0.9	765.3−63.3	31.8±1.2	36.9±0.3	9.9±0.1	108.0±49.0
≥ 4	459.00	241.9±1.0	202.2−4.6	9.8±0.6	16.4±0.2	2.4±0.0	33.0±15.0

TABLE XIX

NUMBER OF EVENTS FOR DATA AND THE PREDICTED NUMBER OF EVENTS FOR MONTE CARLO SIMULATED SAMPLES BEFORE TAGGING. QCD IS SCALED TO THE MATRIX METHOD PREDICTIONS. W+JETS IS NORMALIZED TO DATA USING THE BERENDS SCALING METHOD.

CHAPTER 7

BACKGROUND ESTIMATION IN THE TAGGED DATA SAMPLE

7.1 b-Tagging

For the measurement presented in this thesis, we used the simple secondary vertex b-tagging algorithm high efficiency with medium operating point. The algorithm is described in section 4.7.1. The b-tagging efficiencies are measured as a function of jet transverse momentum and pseudorapidity. The tagging efficiencies for b-jets measured in MC are slightly different than the tagging efficiencies measured in data, as shown in Figure 32. The same discrepancy is observed for c-jets and light flavor jets. We therefore apply a correction factor as a function of jet flavor when b-tagging is performed on the simulated samples.

We used direct tagging for both data and MC. The simulated samples are corrected for the tagging efficiency measured in data by using the scale factors obtained as the ratio between the b-tagging efficiency in data and b-tagging efficiency in MC, as shown in Equation 7.1,

$$SF_x = \frac{\varepsilon_x^{Data}}{\varepsilon_x^{MC}}. \tag{7.1}$$

A similar scale factor is applied to the non-tagged jets as shown in Equation 7.2,

$$\bar{SF}_x = \frac{1-\varepsilon_x^{Data}}{1-\varepsilon_x^{MC}}. \quad (7.2)$$

For the data, the b-tagging efficiency is measured using data driven methods, as described in Section 4.7.1. For MC, the b-tagging efficiency is measured in a QCD sample containing jets with different flavors. Figure 46 (top) shows a comparison between the b-tagging efficiency measured in a $t\bar{t}$ MC sample and in a MC QCD sample for the SSVHEM operating point. Figure 46 (bottom) compares the mistagging rate for the same samples and operating point.

From Figure 46 it can be observed that the b-tagging efficiency measured in QCD is very similar to the b-tagging efficiency measured in the $t\bar{t}$ sample. The b-jets from the QCD events have in general low transverse momentum, hence the measurement of the b-tagging efficiency for the high p_T jets should be performed in a different sample. We chose to use the values measured in the $t\bar{t}$ simulated sample for the high p_T jets.

For light jets, the MC efficiencies and the SF_l are parametrized as a function of the energy of the jet $\mathrm{E}_t^{\mathrm{jet}}$ and the jet rapidity $|\eta^{\mathrm{jet}}|$. For b and c jets, the MC efficiencies are obtained from $t\bar{t}$ MC as a function of the jet transverse momentum $\mathrm{p}_T^{\mathrm{jet}}$ and jet pseudo-rapidity $|\eta^{\mathrm{jet}}|$. Figure 47 shows the tagging efficiency as a function of $\mathrm{p}_T^{\mathrm{jet}}$, separately for different $|\eta^{\mathrm{jet}}|$ bins.

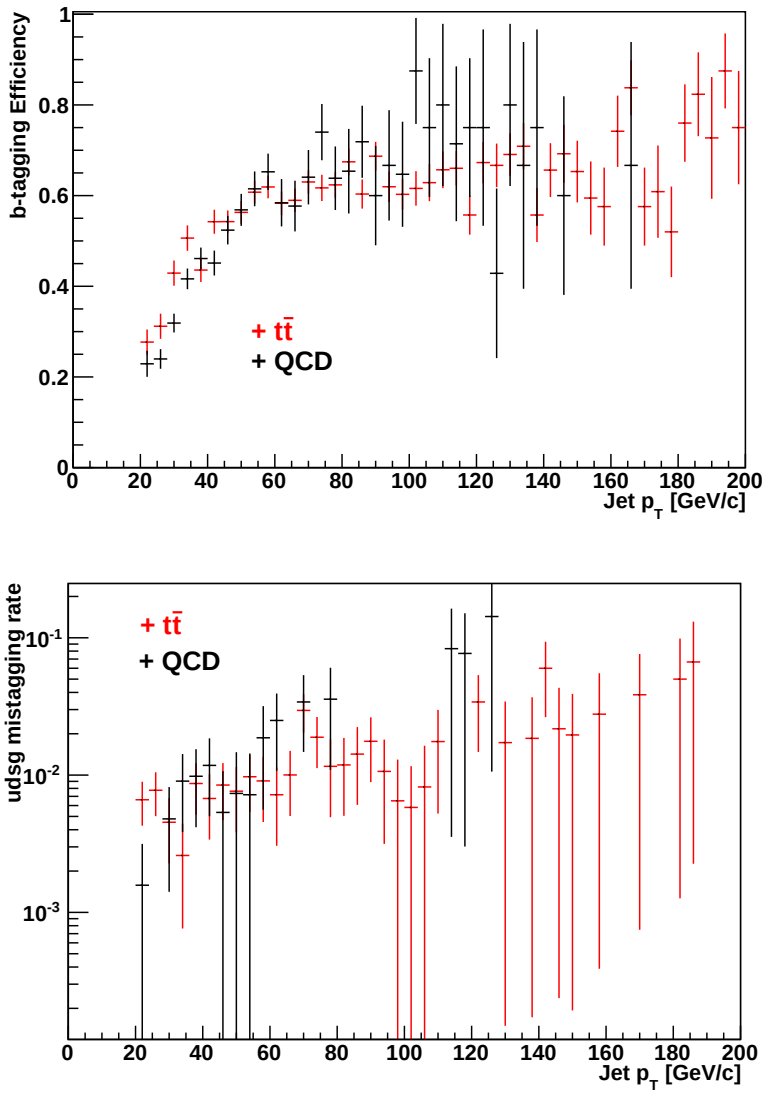

Figure 46. (top) Comparison of the b-tagging efficiency in a MC sample of $t\bar{t}$ and QCD events for the SSVHEM operating point. (bottom) Comparison of the mistagging rate in a MC sample of $t\bar{t}$ and QCD events for the SSVHEM operating point.

As shown in Figure 32, the scale factors for the heavy flavor jets are quite constant as a function of jet p_T. This is not the case, though, for the light jets. We use $\mathrm{SF}_b = 0.9$ for both b-jets and c-jets and a parametrized SF_l as a function of the energy of the jet $\mathrm{E}_t^{\mathrm{jet}}$ and the jet rapidity $|\eta^{\mathrm{jet}}|$ for the light jets (u,d,s,g).

The scale factors applied to the non-tagged jets $\bar{\mathrm{SF}}_x$ can be computed using the scale factors for the tagged jets SF_x and the Monte Carlo tagging efficiencies $\varepsilon_x^{\mathrm{MC}}$, as shown in Equation 7.3,

$$\bar{SF}_x = \frac{1 - SF_x \varepsilon_x^{MC}}{1 - \varepsilon_x^{MC}}. \tag{7.3}$$

Each MC simulated event is thus weighted using both the scale factors for the tagged jets, SF_x and the non-tagged jets, $\bar{\mathrm{SF}}_x$, as defined in the equation Equation 7.4,

$$weight_{event} = \prod_{Ntaggedjets} SF_{taggedjet} * \prod_{Nnon-taggedjets} \bar{SF}_{non-taggedjet}. \tag{7.4}$$

7.2 Background estimation after applying b-tagging

Both data and MC simulated samples were tagged directly, as explained in section 7.1. The pre-tagged selected simulated samples 6.3 for the low rate electroweak physics processes are tagged directly as described in 7.1. Table XX shows the contributions to the tagged sample of the low rate electroweak processes for different jet multiplicity bins with at least one tagged jet.

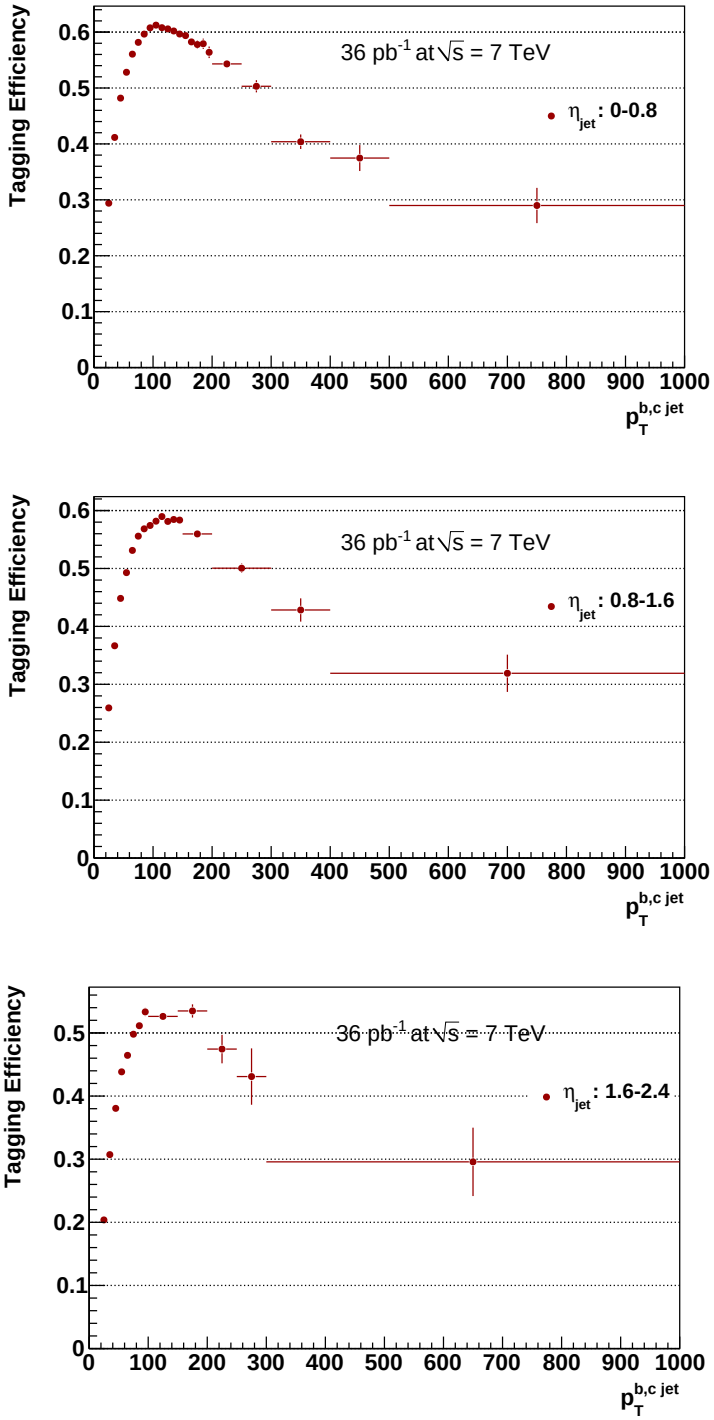

Figure 47. Tagging efficiency for b and c jets obtained from $t\bar{t}$ MC as a function of p_T^{jet}. The efficiency distribution is shown separately for different $|\eta^{jet}|$ bins: $0-0.8$ (top left), $0.8-1.6$ (top right), $1.6-2.4$ (bottom).

N jets	$N_{Z+jets}^{\geq 1tag}$	$N_{singleTop}^{\geq 1tag}$	$N_{VV}^{\geq 1tag}$
$= 1$	14.3	31.7	1.8
$= 2$	6.8	41.0	2.4
$= 3$	2.8	22.8	0.9
≥ 4	1.1	11.0	0.3

TABLE XX

NUMBER OF Z+JETS, SINGLE TOP AND DIBOSON EVENTS OBTAINED AFTER TAGGING.

Table XXI shows the expected number of QCD multijet events with at least one tagged jet.

N jets	$N_{QCD}^{\geq 1tag}$
$= 1$	123.0
$= 2$	57.1
$= 3$	22.5
≥ 4	10.6

TABLE XXI

EXPECTED NUMBER OF QCD MULTIJET EVENTS WITH AT LEAST ONE TAGGED JET. THE PREDICTION OF MC SIMULATION HAS BEEN SCALED USING THE PRE-TAGGED RESULTS OF THE MATRIX METHOD.

The W+jets MC sample has contributions from W+light, Wb(b) and Wc(c) events as shown in Equation 7.5:

$$N_{W+jets} = N_W^{b(b)} + N_W^{c(c)} + N_W^{light}. \quad (7.5)$$

The total number of W+jets in the jet multiplicity bin = 2, is obtained from the Berends Scaling Method before tagging, as explained in section 6.2. This ensures, by construction, that the MC predictions agree with data in that bin. After tagging, we observe that the MC underestimates the data by $\approx 20\%$, as can be seen in Table XXII for events with exactly two jets and exactly one tag. We attribute this discrepancy completely to the predicted rate of Wb(b) and Wc(c) events and assign the same correction, denoted W heavy flavor scale factor W_{HF}, for both Wb(b) and Wc(c) contributions, as shown in Equation 7.6:

$$N_{W+jets} = W_{HF} \times N_W^{b(b)} + W_{HF} \times N_W^{c(c)} + N_W^{light}. \quad (7.6)$$

The correction W_{HF} which restores the agreement between data and MC after tagging in the =2 jets, =1 tag bin is measured to be: $W_{HF} = 2$. We assigned a 50% uncertainty to this value, which covers the difference between the MC prediction and estimation from data. This scale is applied to the normalization of Wb(b) and Wc(c) events with 3 or more jets.

Sample	$W_{HF} = 1$	$W_{HF} = 2$
Data	360	360
QCD	55.3	55.3
Wbb	14.3	28.6
Wcc	56.7	113.5
Wlight	66.8	66.8
Z+jets	6.6	6.6
WW	2.4	2.4
single top	37.2	37.2
Total background	239.3	310.3
$t\bar{t}$	55.5	55.5
Total Monte Carlo	294.7	365.7
$\frac{\text{Data}}{\text{MC predictions}}$	1.22	0.98

TABLE XXII

NUMBER OF EVENTS AFTER TAGGING IN THE = 2 JETS, = 1 TAG BIN. LEFT COLUMN SHOWS THAT THE MC PREDICTIONS UNDERESTIMATE THE DATA BY $\approx 20\%$ AFTER TAGGING, EVEN THOUGH THE MC YIELDS WERE NORMALIZED TO DATA BEFORE TAGGING BY CONSTRUCTION. THIS DISCREPANCY IS CORRECTED BY APPLYING A $W_{HF} = 2$ CORRECTION TO THE YIELDS FOR $Wb(b)$ AND $Wc(c)$ EVENTS, AS CAN BE SEEN IN THE RIGHT COLUMN.

Table XXIII shows the number of Wb(b), Wc(c) and W+light events for different jet multiplicity bins with at least one tagged jet. The total number of events is normalized to the results obtained with the Berends Scaling Method (section 6.2) and the amount of heavy flavor was corrected by the measured heavy flavor scale factor $W_{HF} = 2$.

N jets	$Wb(b)^{\geq 1tag}$	$Wc(c)^{\geq 1tag}$	$W + light^{\geq 1tag}$
$= 1$	28.9	329.0	129.9
$= 2$	28.6	113.5	66.8
$= 3$	17.5	30.3	23.4
≥ 4	7.2	8.6	11.7

TABLE XXIII

NUMBER OF $Wb(b)$, $Wc(c)$ AND *W+LIGHT* EVENTS OBTAINED AFTER TAGGING. THE TOTAL NUMBER OF EVENTS IS NORMALIZED TO THE RESULTS OBTAINED WITH THE BERENDS SCALING METHOD. THE AMOUNT OF HEAVY FLAVOR WAS CORRECTED BY THE MEASURED HEAVY FLAVOR SCALE FACTOR $W_{HF} = 2$

7.3 Composition of the Tagged Data Sample

Table XXIV shows the results after tagging for events with ≥ 3 jets and ≥ 4 jets, with ≥ 1 tag. The theoretical prediction for the $t\bar{t}$ is also shown for comparison. Figure 48 shows the number of events for all jet multiplicity bins for events with at least 1 tag. From the figure it can be observed that the normalization data / MC predictions is preserved after tagging only for $= 2$ jet bin. A heavy flavor factor $= 2$ applied on both b and c jets

for =1 jet bin results in a larger MC predictions than data, due to the large amount of c-jets in this jet multiplicity bin. The $t\bar{t}$ cross section is extracted using the events from = 3 and ≥4 jet bins.

Sample	$\geq$ 3 jets, $\geq$ 1 tag	$\geq$ 4 jets, $\geq$ 1 tag
Data	471	240
QCD	32.1	10.6
Wb(b)	24.7	7.2
Wc(c)	38.9	8.6
Wlight	35.1	11.7
Z+jets	3.9	1.1
WW	1.3	0.3
single top	33.8	11.0
Total background	169.7	50.5
$t\bar{t}$	307.6	181.5
Total prediction	477.3	232.1
$\frac{Data}{MCprediction}$	1.01	0.97

TABLE XXIV

NUMBER OF EVENTS AFTER TAGGING IN THE $\geq$ 3 JET BIN AND $\geq$ 4 JET BIN, WITH $\geq$ 1 TAG. THE AMOUNT OF QCD IS ESTIMATED USING THE MATRIX METHOD BEFORE TAGGING. THE PRETAGGED W+JETS WAS SCALED USING THE BERENDS SCALING METHOD. THE AMOUNT OF Wb(b) AND Wc(c) AFTER TAGGING WAS ALSO SCALED BY THE HEAVY FLAVOR SCALE FACTOR $W_{HF} = 2$. THE SMALL RATE ELECTROWEAK BACKGROUNDS ARE NORMALIZED TO THE CORRESPONDING THEORETICAL CROSS SECTIONS. THE THEORETICAL PREDICTION FOR THE $t\bar{t}$ IS ALSO SHOWN FOR COMPARISON.

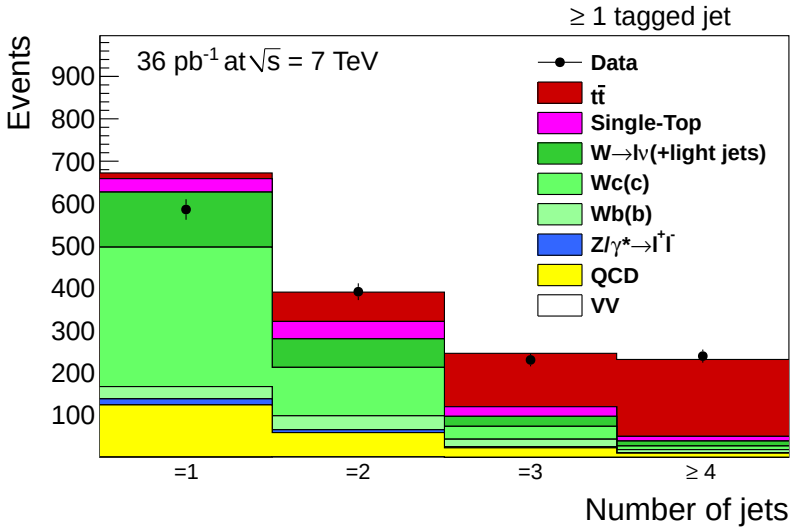

Figure 48. Number of events with at least 1 tag for data and MC samples for each jet multiplicity bin. QCD normalization in each bin is taken from the Matrix Method prediction before tagging. W+jets normalization is also taken from the Matrix Method for events with 1 and 2 jets, and from the Berends' scaling for higher jet multiplicity bins. In this plot, the Wb(b) and Wc(c) contributions were scaled by a heavy flavor scale factor $W_{HF} = 2$.

CHAPTER 8

$t\bar{t}$ PRODUCTION CROSS SECTION MEASUREMENT

8.1 Systematic Uncertainties

The $t\bar{t}$ cross section ($\sigma_{t\bar{t}}$) measurement is affected by several systematic uncertainties, listed below.

8.1.1 Systematic uncertainty from the Berends-Giele Scaling Method

The systematic uncertainty that is assigned to the W+jets normalization is obtained from the difference between the constant value C(1) and the C(n) values obtained from the fit. The systematic uncertainty estimated for this method is $^{+0.0}_{-8.3}$ pb. Note that this uncertainty is nonzero only in the "$-$" direction, due to the fact that the values C(n) obtained from the fit are always larger than the constant value C(1) used in the measurement of $\sigma_{t\bar{t}}$.

8.1.2 Systematic uncertainty from the Matrix Method

The uncertainties assigned to the QCD background in the pretagged data are shown in section 6.1.2 as a function of the number of jets. For the case of $\geq$ 3jets, the combined uncertainty is obtained as a quadratic sum of the uncertainties assigned to events with = 3 jets and $\geq$ 4 jets and corresponds to $\pm$ 36.3 %. To be more conservative, we took this variation to be $\pm$50%. This variation results in a systematic uncertainty on the $\sigma_{t\bar{t}}$ of $^{+0.9}_{-1.2}$ pb.

8.1.3 Systematic uncertainty from the normalization of signal and background

The theoretical cross sections for top pair and single top production were used to subtract the signal from data when normalizing the W+jets contribution (Equation 6.3). In addition, single top was also normalized to theory in the $\sigma_{t\bar{t}}$ extraction (Equation 8.2). The systematic uncertainty obtained by varying $\sigma_{t\bar{t}}$ by 30% and $\sigma_{\text{single Top}}$ by 50%. The uncertainty due to the variation of the theoretical cross section for the $Z+jets$ is $^{+1.2}_{-1.2}$ pb. The combined uncertainty from these variations on the $\sigma_{t\bar{t}}$ is $^{+4.3}_{-4.4}$ pb.

8.1.4 Systematic uncertainty due to Factorization Scale

To account for the systematic uncertainty due to the factorization scale, different MC samples were simulated with MADGRAPH by modifying the scale μ_F by a factor of 0.5 (down) and 2.0 (up) with respect to the nominal value. $t\bar{t}$, Z+jets and a combined sample of W+jets and VQQ, all scaled in the same direction (up or down), were used to study the difference in the $\sigma_{t\bar{t}}^{\geq 3\text{jets},\geq 1\text{tag}}$ with respect to the central values.. The systematic uncertainty resulting from this variation on the $\sigma_{t\bar{t}}$ is $^{+9.9}_{-3.6}$ pb.

8.1.5 Systematic uncertainty due to ME to PS Matching

We investigated the impact of varying the ME to PS jet matching threshold for $t\bar{t}$, W+jets and Z+jets events.The threshold used for generating the $t\bar{t}$ MC sample was changed to 10 GeV/c (down) and 40 GeV/c (up) from the nominal value of 20 GeV/c. In the case of the W+jets and Z+jets samples, the threshold was changed to 5 GeV/c (down)

and 20 GeV/c (up) from the nominal value of 10 GeV/c. The systematic uncertainty resulting from this variation on the $\sigma_{t\bar{t}}$ is $^{+2.0}_{-0.01}$ pb.

8.1.6 Systematic uncertainty from b-tagging efficiency

The uncertainty assigned for the b-tagging efficiency is directly related to the tagged jets scale factors SF_x defined in Equation 7.1. We vary the scale factors SF_x as follows:

1. For b-jets and c-jets we vary the SF_x by $\pm 15\%$.
2. For light-jets, the uncertainty is parametrized as a function of the energy of the jet E_t^{jet} and the jet pseudorapidity $|\eta^{\text{jet}}|$, varying up to 15% (65).

The variation of the tagged jets scale factors SF_x is also directly propagated into the expression of the non-tagged jets $\bar{SF}_x$, as expressed in Equation 7.3. The systematic uncertainty assigned to $\sigma_{t\bar{t}}$ due to these variations is : $^{+18.7}_{-13.2}$ pb.

8.1.7 Systematic uncertainty due to the Heavy Flavor Scale Factor

As specified in section 7.2, the heavy flavor scale factor used to scale the amount of Wbb and Wcc in our tagged sample is 2. We assigned a $\pm 50\%$ uncertainty to this factor and we recalculated the cross section. The systematic uncertainty resulting from this variation on the $\sigma_{t\bar{t}}$ is $^{+15.2}_{-15.6}$ pb.

8.1.8 Systematic uncertainty due to Jet Energy Scale

The uncertainty to the JES correction is applied to each jet four-momentum and includes:

1. 1σ uncertainty according to jet transverse momentum and jet pseudorapidity
2. for b-jets: 0.02 for $50\text{GeV} < \text{p}_\text{T} < 200\text{GeV}$ and 0.03 for values outside this range
3. additional corrections that account for CMSSW release differences, calibration changes and L2 and L3 trigger corrections (66).

All the corrections are added in quadrature and applied to jet-by-jet four-momentum. The difference in energy obtained from the scaling up and down of the jets energy was also propagated into the missing transverse energy:

$$\Delta \overrightarrow{\not{E}}_\text{T} = -\sum \Delta \overrightarrow{\text{p}}_{\text{T,jet}}. \tag{8.1}$$

The systematic uncertainty resulting from these variations on $\sigma_{t\bar{t}}^{\geq 3\text{jets},\geq 1\text{tag}}$ is $^{+7.8}_{-8.2}$ pb.

8.1.9 Systematic uncertainty due to Jet Energy Resolution

The measurement of the jet asymmetry suggests that the jet energy resolution, denoted as **JER**, measured in data is about 10% lower that the measured values on Monte Carlo simulations (67). To account for this effect, all jets in the MC simulated samples have to be scaled accordingly and the effect is propagated to the missing transverse energy as shown in Equation 8.1. The overall JER uncertainty assigned to the measured $t\bar{t}$ cross section is $^{+0.9}_{-0.9}$ pb.

8.1.10 Systematic Uncertainty due to Initial and Final State Radiation

A pp collision at a hadron collider can produce gluon radiation from both incoming and outgoing partons. ISR/FSR variations on MC $t\bar{t}$ result on a systematic uncertainty on $\sigma_{t\bar{t}}^{\geq 3\text{jets},\geq 1\text{tag}}$ of $^{+4.8}_{-0.2}$ pb.

8.1.11 Systematic Uncertainty due to Parton Distribution Function

The limited knowledge of the PDFs of colliding protons represents another source of systematic uncertainty for the measured $t\bar{t}$ cross section. This systematic uncertainty is evaluated following the re-weighting method described in (68), using the alternative PDF set CTEQ6.6 (69), via the LHAPDF library (70). The event weight is applied to the $t\bar{t}$ and W+jets simulated samples using positive and negative variations of all of the 22 eigenvectors of the CTEQ6.6 PDF, resulting in a total of 44 different templates for each sample. The resulting variations on the measured $t\bar{t}$ cross section is $^{+3.1}_{-3.1}$ pb .

8.1.12 Systematic Uncertainty due to Pile-up

The effect of the pile-up for this data set is very small, due to the fact that only a small number of events are found to have two primary vertices. We found that there is a negligible effect on our $\sigma_{t\bar{t}}^{\geq 3\text{jets},\geq 1\text{tag}}$ measurement.

8.1.13 Luminosity

To the measured $t\bar{t}$ cross section is assigned an uncertainty of 4% of the total integrated luminosity, as described in (29).

8.1.14 Summary of the systematic uncertainties

The systematic uncertainties are summarized in Table XXV. The combined systematic uncertainty is obtained from the quadratic sum of all the systematic uncertainties, except the luminosity, which is quoted separately. The dominant uncertainties are the ones corresponding to the b-tagging efficiency and the heavy flavor scale factor.

Systematic	$\sigma_{t\bar{t}}$		fractional changes (%)	
W+jets estimation	+0.0	−8.3	+0.0	−5.4
QCD estimation	+0.9	−1.2	+0.5	−0.8
Theoretical cross sections	+4.3	−4.4	+2.8	−2.9
Factorization Scale	+9.9	−3.6	+6.4	−2.3
ME-PS Matching	+2.0	−0.0	+1.3	−0.01
b-tagging efficiency	+18.7	−13.2	+12.2	−8.6
HF scale factor	+15.2	−15.6	+9.9	−10.1
JES	+7.8	−8.2	+5.1	−5.3
JER	+0.9	−0.9	+0.6	−0.6
ISR/FSR	+4.8	−0.2	+3.1	−0.1
PDF CTEQ6.6	+3.1	−3.1	+2.0	−2.0
Total systematics	+28.2	−24.4	+18.3	−15.8

TABLE XXV

SYSTEMATIC UNCERTAINTIES FOR THE TOP PAIR CROSS SECTION MEASURED FOR ≥ 3 JETS, ≥ 1 TAG. THE DOMINANT UNCERTAINTIES ARE THE ONES CORRESPONDING TO THE B-TAGGING EFFICIENCY AND THE HEAVY FLAVOR SCALE FACTOR, SHOWN IN BOLD.

8.2 $t\bar{t}$ Cross Section Extraction Method

The $t\bar{t}$ production cross section is measured using 36 pb^{-1} of proton-proton collisions at $\sqrt{s} = 7$ TeV. The data was recorded with the CMS detector at the Large Hadron Collider. As shown in section 7.3, the $t\bar{t}$ signal dominates the preselected sample with three or more jets after a lifetime b-tagging algorithm that explicitly reconstructs the secondary vertices is applied.

The $t\bar{t}$ production cross section $\sigma_{t\bar{t}}$ is extracted by counting the number of tagged events, $N_{data}^{\geq 1tag}$, and subtracting the estimated number of tagged background events, $N_{background}^{\geq 1tag}$, as shown in Equation 8.2:

$$\sigma_{t\bar{t}} = \frac{N_{data}^{\geq 1tag} - N_{background}^{\geq 1tag}}{BR \times \mathcal{L} \times \epsilon_{presel} \times P^{\geq 1tag}}, \tag{8.2}$$

where BR represents the branching ratio of the $\mu + jets$ final state of the $t\bar{t}$ production, $\mathcal{L}$ represents the integrated luminosity, ϵ_{presel} is the $t\bar{t}$ preselection efficiency and $P^{\geq 1tag}$ is the probability for a $t\bar{t}$ event to have one or more jets identified as b jets, also called the **event tagging probability**.

The total efficiency of the preselection is measured in the $t\bar{t}$ simulated sample and corrected by the following data-to-MC SFs: trigger efficiency, detector acceptance efficiency and the efficiency to select an isolated muon in the sample with three or more jets. The tagging probability was measured in the $t\bar{t}$ simulated sample and corrected as described in section 7.1.

The measured $t\bar{t}$ production cross section in the $\mu + jets$ channel for events with at least three jets and at least one jet tagged as a b-jet is:

$$\sigma_{t\bar{t}} = 153.7 \pm 9.0(\text{stat.})^{+28.2}_{-24.4}(\text{syst.}) \pm 6.1(\text{lumi}) \text{ pb.} \tag{8.3}$$

The measured value is found to be in good agreement with the NLO theoretical prediction of $\sigma_{t\bar{t}} = 157.5^{+23.3}_{-24.4}$ pb and NNLO theoretical prediction of $\sigma_{t\bar{t}} = 165^{+11}_{-16}$ pb (12) for a top mass of 173 GeV/c^2 (1).

Figure 49 − Figure 53 show various kinematic distributions of the tagged events in data compared with the sum of the signal and the background. The $t\bar{t}$ distributions were normalized to the measured cross section. The plotted variables are:

- M_3 - the reconstructed mass of the hadronically decaying top quark. For the hadronically decaying top, there are three jets reconstructed in the final state, and one of them should be identified as a b-jet. These jets are expected to have a very large transverse momentum. Hence, for reconstructing the top quark decaying hadronically, the sum of the two leading untagged jets (ordered according to their transverse momentum) and one b-jet is performed, as shown in Equation 8.4:

$$p_{t_{hadronic}} = \sum_{i=1,2} p_{\text{untagged jet}_i} + p_{\text{tagged jet}}. \tag{8.4}$$

All the possible jets combinations are taken into account and the sum with the largest transverse momentum is assigned to be $p_{t_{hadronic}}$. The mass of the hadronically decaying top quark is extracted from the reconstructed top quark four vector $p_{t_{hadronic}}$.

- H_T - the scalar sum of the transverse energies of the four leading jets, ordered according to their transverse momentum.
- $p_T^{\text{leading Jet}}$ - the transverse momentum of the leading jet.
- p_T^{μ} - the transverse momentum of the muon.
- $\not{E}_T$ - missing transverse energy.

Figure 49 and Figure 50 shows M_3 and H_T, respectively, for the events with ≥ 3 jets and at least 1 tag and events with ≥ 4 jets and at least 1 tag. Figure 51 and Figure 52 shows the $p_T^{\text{leading Jet}}$ and p_T^{μ}, respectively, for events with ≥ 3 jets and at least 1 tag and events with ≥ 4 jets and at least 1 tag. Figure 53 shows the $\not{E}_T$ for events with ≥ 3 jets and at least 1 tag and events with ≥ 4 jets and at least 1 tag. On all the figures, only the statistical uncertainties are shown. A good agreement between data and the scaled MC predictions can be observed for all the variables.

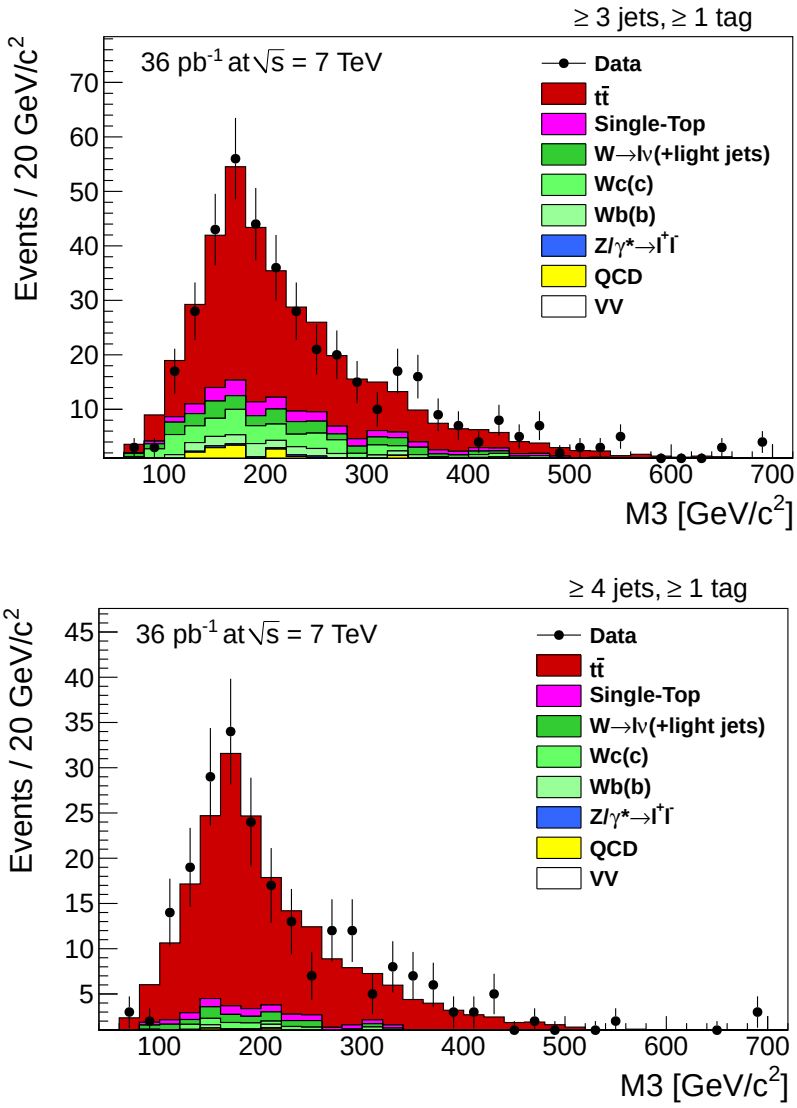

Figure 49. M_3 of the leading jet distributions for the events with ≥ 3 jets and at least 1 tag (top) and events with ≥ 4 jets and at least 1 tag (bottom). The M_3 values are obtained by taking the largest p_T from the vectorial sum of the two untagged jets and one tagged jet in each event. Only statistical uncertainties are shown.

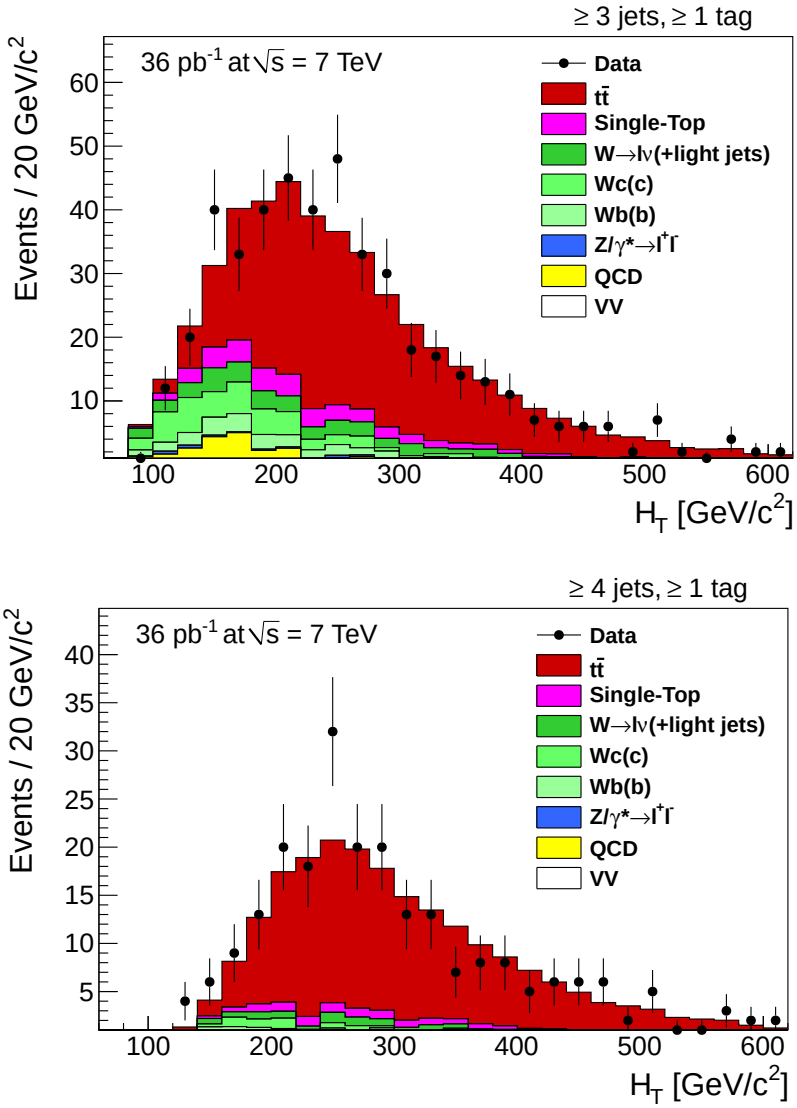

Figure 50. H_T of the leading jet distributions for the events with ≥ 3 jets and at least 1 tag (top) and events with ≥ 4 jets and at least 1 tag (bottom). Only statistical uncertainties are shown.

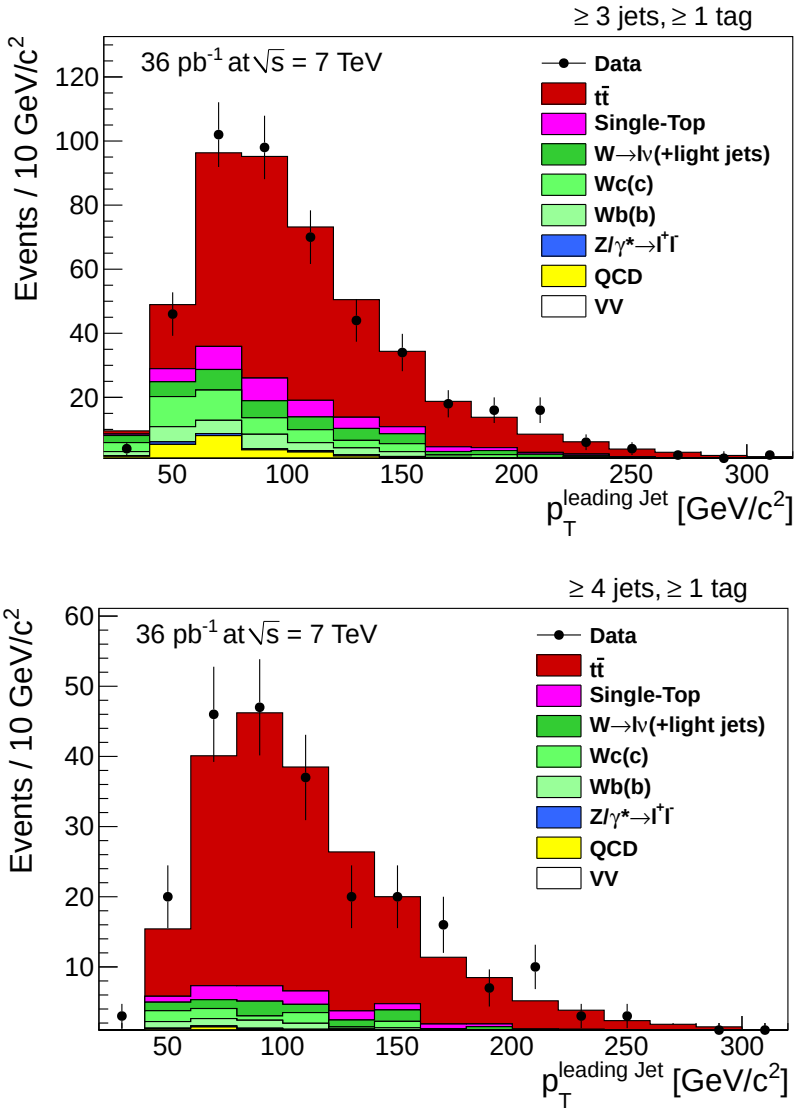

Figure 51. Transverse momentum of the leading jet for the events with ≥ 3 jets and at least 1 tag (top) and events with ≥ 4 jets and at least 1 tag (bottom). Only statistical uncertainties are shown.

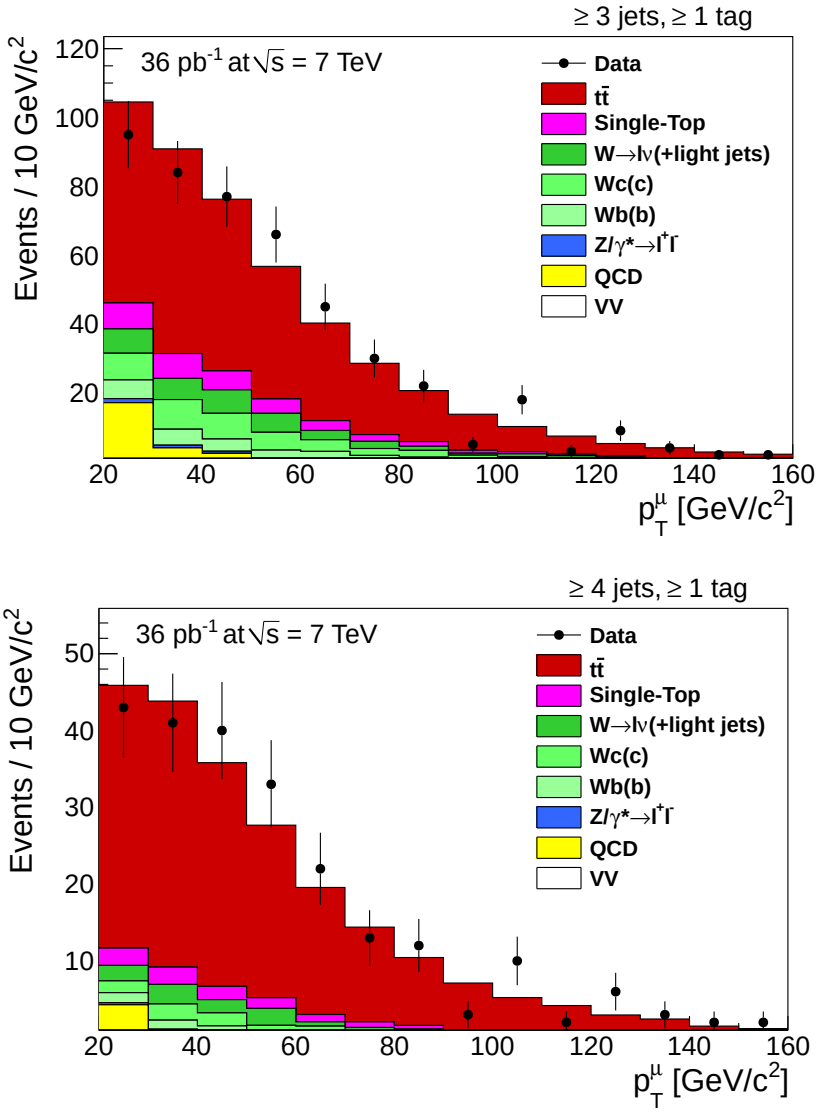

Figure 52. Transverse momentum of the muon for the events with ≥ 3 jets and at least 1 tag (top) and events with ≥ 4 jets and at least 1 tag (bottom). Only statistical uncertainties are shown.

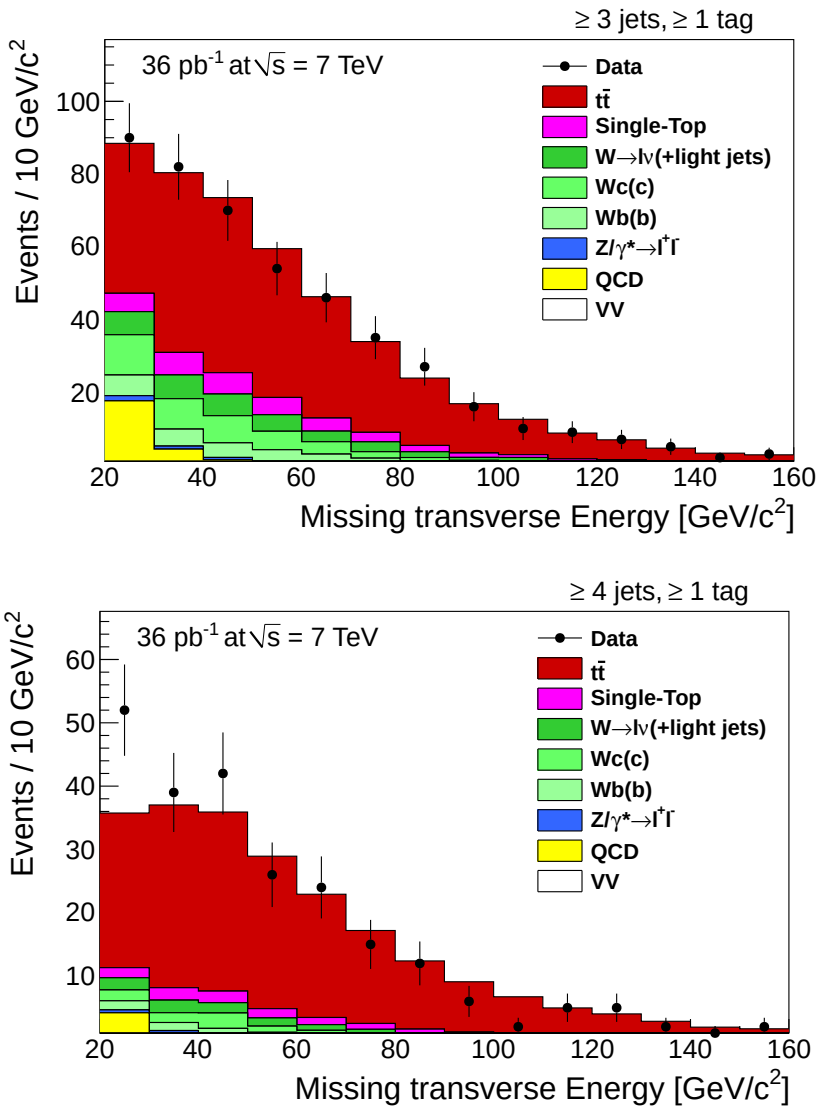

Figure 53. The missing transverse energy distributions for the events with ≥ 3 jets and at least 1 tag (top) and events with ≥ 4 jets and at least 1 tag (bottom). Only statistical uncertainties are shown.

CHAPTER 9

SUMMARY AND CONCLUSIONS

We measured the top pair production cross section using events with a high p_T isolated muon, missing transverse energy, and at least three jets, of which at least one is identified as originating from the hadronization of a b-quark. Before requiring the presence of a b-jet, we use a data-driven method to estimate the amount of QCD multijet background separately for events with different number of jets. This method also allows us to obtain the overall normalization of the W+jets background for events with ≥ 1 and ≥ 2 jets after subtracting the contributions from small electroweak backgrounds and top signal as predicted by theoretical predictions.

The Berends-Giele scaling is used to determine the number of W events with ≥ 3 jets from the scale measured in events with ≥ 1 and ≥ 2 jets. We required that at least one jet is tagged in the selected data and the MC samples using a simple secondary vertex algorithm which selects b-jets with an average efficiency of 60%, and light jets at an average mis-rate of 1%. After tagging, we observe that the MC prediction underestimates the data by approximately 20% for events with $= 2$ jets, and $= 1$ tag. We attribute this discrepancy to an underestimate on the relative rate of W+bb and W+cc processes in our W+jets MC sample, and derive a W heavy flavor scale factor of 2 ± 1 to restore the normalization. This same factor is also applied to the MC W+bb and W+cc processes

with at least 3 jets. The top pair production cross section is then extracted from the number of events above the background prediction with ≥ 3 jets and ≥ 1 tags. The main source of systematic uncertainty comes from the b-tagging efficiency and heavy flavor scale factor.

The measured top quark pair cross section for a top quark of mass $m_t = 173\ GeV/c^2$ (1),

$$\sigma_{t\bar{t}} = 153.7 \pm 9.0(\text{stat.})^{+28.2}_{-24.4}(\text{syst.}) \pm 6.1(\text{lumi})\ \text{pb},$$

is in good agreement with theoretical predictions, $\sigma^{NLO}_{t\bar{t}} = 157^{+23}_{-24}$ pb and $\sigma^{NNLO}_{t\bar{t}} = 165^{+11}_{-16}$ pb (12). Figure 54 shows that it also agrees well with the measurements performed with different techniques (e.g. extraction of the $t\bar{t}$ cross section using likelihood fit (71)), combined different channels (muon+jets, electron+jets (72) and dilepton (73)) or the measurements performed on data collected with other experiments (ATLAS (74)).

Figure 55 (71) shows the measurements of the $t\bar{t}$ production cross section in pp collisions at the LHC and the ones performed in $p\bar{p}$ collisions at the Tevatron (75; 76; 77; 78). The approximate NNLO theory predictions as a function of center of mass energy is also shown. The error band corresponds to the variation resulting from changing the Q^2 scale up and down by a factor of 2. We find good agreement between the data and the theory in all cases.

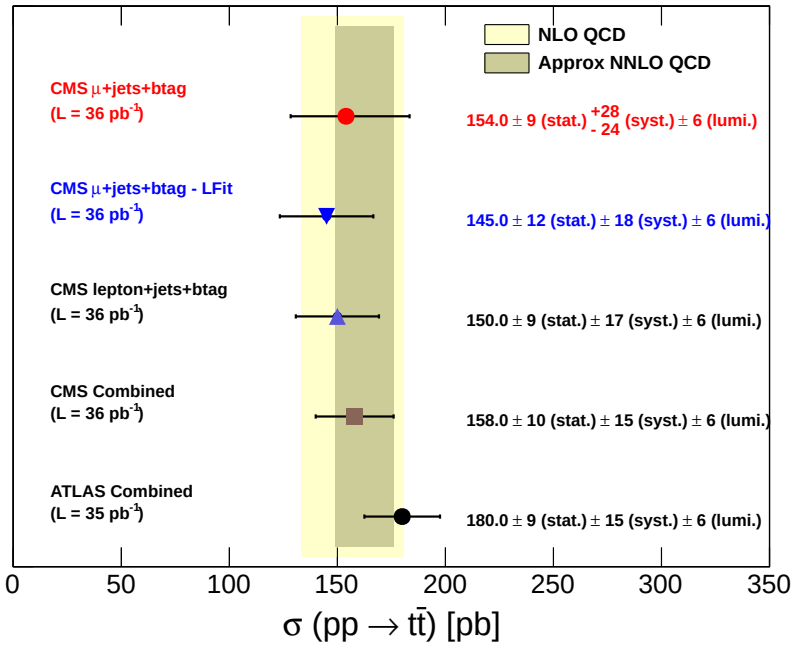

Figure 54. Comparison of the measured top quark production cross section in muon+jets+missing transverse energy events at CMS with measurements in other channels at both CMS and ATLAS experiments. The NLO QCD and approximate NNLO QCD theory prediction is also shown.

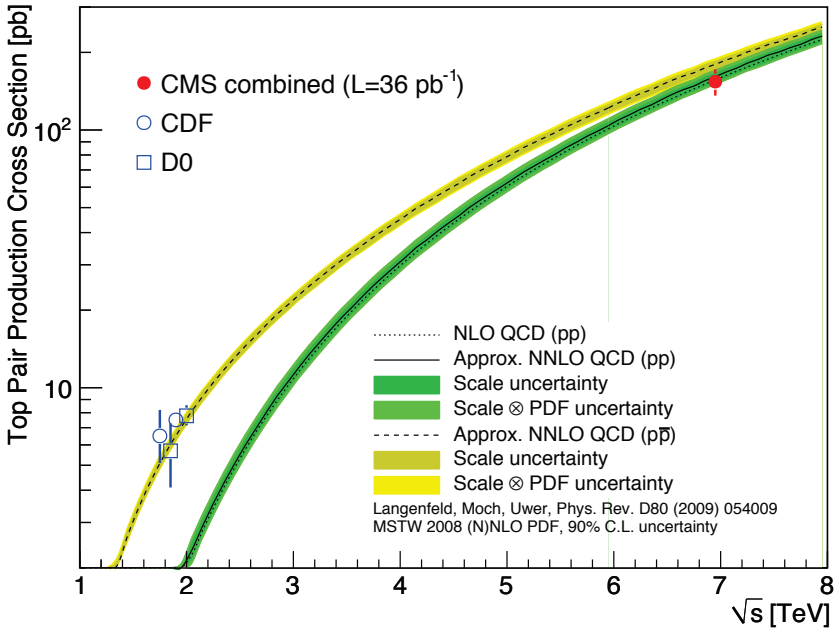

Figure 55. Comparison of the CMS and Tevatron results for the $t\bar{t}$ production cross section with theoretical predictions as a function of $\sqrt{s}$.

CITED LITERATURE

1. Group, T. E. W.: Combination of CDF and Dzero Results on the Mass of the Top Quark. FERMILAB TM 2427 E, May 2009. http://arxiv.org/abs/0903.2503.

2. Fermilab: Discoveries at Fermilab - Discovery of the Bottom Quark. Press Release, August 1977.

3. Collaboration, P. D. G.: Particle Physics Group, volume 37. Journal of Physics G, 2010.

4. Quigg, C.: History of the Top quark mass. Physics Today, 50N5(20), April 1997.

5. Francis Halzen, A. M.: Quarks and Leptons. John Wiley And Sons, 1984.

6. Sakurai, J.: Modern Quantum Mechanics. Addison-Wesley, revised edition edition, 1994.

7. Michael Peskin, D. S.: An Introduction to Quantum Field Theory. Westview Press, 1995.

8. Yao, W. et al.: Review of Particle Physics. Journal of Physics G, 33, 2006.

9. Quadt, A.: Top Quark Physics at Hadron Colliders. The European Physical Journal C, 48:835–1000, 2007.

10. J. Pumplin et al.: New Generation of Parton Distributions with Uncertainties from Global QCD Analysis. Journal of high Energy Physics, 07(012), 2002.

11. S. Moch, P. U.: Heavy-quark pair production at two loops in QCD. Nuclear Phys. Proc. Suppl., 183:75, 2008.

12. N. Kidonakis: Next-to-next-to-leading soft-gluon corrections for the top quark cross section and transverse momentum distribution. Physics Review D, 82(114030), 2010. arXiv:1009.4935.

13. J. M. Campbell, R. Frederix, F. Maltoni et al.: Next−to−Leading−Order Predictions for t−Channel Single−Top Production at Hadron Colliders. Phys. Rev. Lett., 102(182003), 2009. arXiv:0903.0005. doi:10.1103/PhysRevLett.102.182003.

14. J. M. Campbell and F. Tramontano: Next-to-leading order corrections to Wt production and decay. Nuclear Physics B, 726(109), 2005. arXiv:0506289.

15. J. M. Campbell, R. K. Ellis, and F. Tramontano: Single top production and decay at next-to-leading order. Physics Review D, 70(094012), 2004. arXiv:hep-ph/0408158.

16. Evans, L. and Bryant, P.: LHC Machine. JINST, 3(S08001), August 2008.

17. The ATLAS Collaboration et al: The ATLAS Experiment at the CERN Large Hadron Collider. JINST, 3(S08003), August 2008.

18. The CMS Collaboration et al: The CMS experiment at the CERN LHC. JINST, 3(S08004), August 2008.

19. The ALICE Collaboration et al: The ALICE experiment at the CERN LHC. JINST, 3(S08002), August 2008.

20. The LHCb Collaboration et al: The LHCb Detector at the LHC. JINST, 3(S08005), August 2008.

21. The TOTEM Collaboration et al: TOTEM Technical Design Report. Technical Design Report CERN-LHCC-2004-002, CERN/LHCC, January 2004.

22. The LHCf Collaboration et al: LHCf Experiment: Technical Design Report. Technical Design Report CERN-LHCC-2006-004, CERN/LHCC, 2006.

23. The CMS Collaboration et al: The CMS magnet project. Technical Design Report CERN-LHCC-97-10, CERN, May 1997.

24. The CMS Collaboration et al: The CMS tracker system project. Technical Design Report CERN-LHCC-98-6, CERN, April 1998.

25. Grupen, C. and Shwartz, B.: Particle Detectors. Cambrige University Press, Second edition, 2008.

26. The CMS Collaboration et al: The CMS electromagnetic calorimeter project. Technical Design Report CERN-LHCC-97-33, CERN, December 1997.

27. CMS Collaboration: Jet Performance in pp Collisions at $sqrt(s) = 7$ TeV. CMS PAS JME-10-003, CERN, 2010.

28. The CMS Collaboration et al: The CMS hadron calorimeter project. Technical Design Report CERN-LHCC-97-31, CERN, June 1997.

29. CMS Collaboration: Measurement of the CMS Luminosity. CMS PAS EWK-10-004, CERN, 2010.

30. The CMS Collaboration et al: The CMS muon project. Technical Design Report CERN-LHCC-97-32, CERN, December 1997.

31. The CMS Collaboration et al: The CMS trigger and data-acquisition project. Technical Design Report CERN-LHCC-2002-26, CERN, December 2002.

32. The CMS Collaboration et al: The CMS TriDAS project. Technical Design Report CERN-LHCC-2000-38, CERN, December 2000.

33. Eck, C., et al.: LHC Computing Grid. Technical Design Report. Technical Design Report 024, CERN-LHCC, 2005.

34. Brun, R. and Rademakers, F.: ROOT An object oriented data analysis framework. Nuclear Instruments and Methods in Physics Research A, 389(1-2):81–86, April 1997.

35. Kovalskyi, D., et al.: Fireworks: A Physics Event Display for CMS. Journal of Physics: conference Series, 219(032014), 2010.

36. CMS Collaboration: Particle−Flow Event Reconstruction in CMS and Performance for Jets, Taus, and Missing e_T. CMS PAS PFT-09-001, CERN, 2009.

37. Adam, Speer, Mangano and Todorov: Track reconstruction in the CSM tracker. CMS Note 041, CERN, 2006.

38. T. Speer et al.: Track reconstruction in the CMS tracker. Nuclear Instruments and Methods in Physics Research A, 559(1):143–147, December 2006.

39. Ian Tomalin et al: Description and Performance of the CMS track reconstruction. CMS Note 172, CERN, 2011.

40. CMS Collaboration: Tracking and Primary Vertex Results in First 7 TeV Collisions. CMS PAS TRK-10-005, CERN, 2010.

41. R. Fruewirth, W. Wltenberger and P. Vanlaer: Adaptive Vertex Fitting. CMS Note 008, CERN, 2007.

42. G.Abbiendi et al.: Muon Reconstruction in the CMS Detector. CMS Note 097, CERN, 2008.

43. CMS Collaboration: Performance of muon reconstruction and identification in pp collisions at $sqrt(s) = 7$ TeV. CMS PAS MUO-10-004, CERN, 2010.

44. CMS Collaboration: Electromagnetic physics objects commissioning with first LHC data. CMS PAS EGM-10-001, CERN, 2010.

45. CMS Collaboration: Commissioning of the Particle flow Event Reconstruction with the first LHC collisions recorded in cms detector. CMS PAS PFT-10-001, CERN, 2010.

46. CMS Collaboration: Electron reconstruction and identification at $sqrt(s) = 7$ TeV data. CMS PAS EGM-10-004, CERN, 2010.

47. Cacciari, M. and Salam, G. P.: The anti-kt jet clustering algorithm. JHEP, 04(063), April 2008.

48. CMS Collaboration: Determination of the Jet Energy Scale in CMS with pp Collisions at $sqrt(s) = 7$ TeV. CMS PAS JME-10-010, CERN, 2010.

49. CMS Collaboration: Commissioning of b−jet identification with pp collisions at $sqrt(s) = 7$ TeV. CMS PAS BTV-10-001, CERN, 2010.

50. A. Rizzi, F. Palla and G. Segneri: Track impact parameter based b−tagging with CMS. CMS NOTE 019, CERN, 2006.

51. C. Gerber, S. Khalatyan, F Yumiceva et al.: Measurement of the b−tagging efficiency using the System8 Method. CMS NOTE 195, CERN, 2011.

52. CMS Collaboration: Missing Transverse Energy Performance in Minimum−Bias and Jet Events from Proton−Proton Collisions at $sqrt(s) = 7$ TeV. CMS PAS JME-10-004, CERN, 2010.

53. CMS Collaboration: CMS MET Performance in Events Containing Electroweak Bosons from pp Collisions at $sqrt(s) = 7$ TeV. CMS PAS JME-10-005, CERN, 2010.

54. J. Alwall et al.: MadGraph/MadEvent v4: The New Web Generation. JHEP, 09(028), 2007. arXiv:0706.2334. doi:10.1088/1126-6708/2007/09/028.

55. T. Sjostrand, S. Mrenna, and P. Z. Skands: PYTHIA 6.4 Physics and Manual. JHEP, 05(026), 2006. arXiv:hep-ph/0603175. doi:10.1088/1126-6708/2006/05/026.

56. J. Allison et al.: Geant4 developments and applications. IEEE Trans. Nucl. Sci., 53(270), 2006. doi:10.1109/TNS.2006.869826.

57. J. Campbell and R. Ellis: MCFM for the Tevatron and the LHC. Technical report, FERMILAB-Conf-10-244-T, 2010. arXiv:hep-ph/1007.3492.

58. A. D. Martin, W. J. Stirling, R. S. Thorne et al.: Uncertainties on α_s in global PDF analyses and implications for predicted hadronic cross sections. Eur. Phys. J., C(64):653–680, 2009. arXiv:0905.3531. doi:10.1140/epjc/s10052-009-1164-2.

59. H.-L. Lai et al.: Uncertainty induced by QCD coupling in the CTEQ-TEA global analysis of parton distributions. http://arxiv.org/abs/1004.4624.

60. F. Demartin, S. Forte, E. Mariani et al.: The impact of PDF and α_s uncertainties on Higgs Production in gluon fusion at hadron colliders. Phys. Rev. D, 82(014002), 2010. arXiv:1004.0962. doi:10.1103/PhysRevD.82.014002.

61. Hung-Liang Lai et al.: Uncertainty induced by QCD coupling in the CTEQ global analysis of parton distributions. http://wwwhep.ucl.ac.uk/pdf4lhc/.

62. K. Melnikov and F. Petriello: Electroweak gauge boson production at hadron colliders through $O(\alpha(s)^2)$. Phys. Rev. D., 74(114017), 2006. arXiv:hep-ph/0609070. doi:10.1103/PhysRevD.74.114017.

63. CMS Collaboration: FlavorHistory: A Heavy Flavor Overlap Tool. CMS NOTE 147, CERN, 2009.

64. Aurelio Juste et al.: The Matrix Method and its Error Calculation. DØ NOTE 4564, DØ , 2004.

65. Daniel Bloch et al.: Study of b-tag mistag rates using negative taggers in early CMS data. CMS NOTE 147, CERN, 2010.

66. CMS Collaboration: Determination of the Jet Energy Calibration and Transverse Momentum Resolution in CMS. CMS PAS JME-10-011, CERN, 2010.

67. CMS Collaboration: Jet Resolution Determination at $sqrt(s) = 7$ TeV. CMS PAS JME-10-014, CERN, 2010.

68. CMS Collaboration: Parton Distribution Uncertainty Determination within cmssw. CMS NOTE 048, CERN, 2009.

69. H.L. Lai et al.: Uncertainty induced by QCD coupling in the CTEQ-TEA global analysis of parton distributions. Technical report, CTEQ, 2010.

70. M. R. Whalley, D. Bourilkov, and R. C. Group: The Les Houches Accord PDFs (LHAPDF) and Lhaglue.

71. CMS Collaboration: Measurement of the t t-bar Production Cross Section in pp Collisions at 7 TeV in Lepton + Jets Events Using b-quark Jet Identification. http://arxiv.org/abs/1108.3773, August 2011.

72. CMS Collaboration: Measurement of the Top-antitop Production Cross Section in pp Collisions at $\sqrt{s} = 7$ TeV using the Kinematic Properties of Events with Leptons and Jets. http://arxiv.org/abs/1106.0902, June 2011.

73. 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. http://arxiv.org/abs/1105.5661, May 2011.

74. ATLAS Collaboration: A combined measurement of the top quark pair production cross-section using dilepton and single-lepton final states. ATLAS-CONF 040, ATLAS, March 2011.

75. DØ Collaboration: Measurement of the top quark pair production cross section in the lepton+jets channel in proton-antiproton collisions at *sqrts* = 1.96 tev. Physics Review D, 84(012008), 2011. arXiv:1101.0124. doi:10.1103/PhysRevD.84.012008.

76. DØ Collaboration: t anti-t production cross-section in p anti-p collisions at *sqrts* = 1.8 tev. Physics Review D, 67(012004), 2003. arXiv:0205019. doi:10.1103/PhysRevD.67.012004.

77. CDF Collaboration: Combination of cdf top quark pair production measurements with up to 4.6 fb^{-1}. CDF Note 9913, CDF, 2009.

78. CDF Collaboration: Measurement of the t anti-t production cross-section in p anti-p collisions at *sqrts* = 1.8 tev. Physics Review D, 64(032002), 2001. arXiv:0101036.

VITA

Name	Ioana Maria Anghel
Education	Ph.D., Physics University of Illinois at Chicago, Chicago, Illinois 2011
	M.S. Graduate Classes, Physics University of Bucharest, Bucharest, Romania 2003 - 2004
	B.Sc., Applied Physics and Engineering Technical University of Cluj-Napoca, Cluj-Napoca, Romania 2002
Summer Schools	The 5^{th} CERN-Fermilab Hadron Collider Physics Summer School Fermilab, Batavia, Illinois 2010
	The 16^{th} Coordinated Theoretical-Experimental Project on QCD Summer School University of Wisconsin, Madison, Wisconsin 2009
Awards	Stipendium des Deutschen Akademischen Austauschdienstes Technical University of Chemnitz, Chemnitz, Germany 2002
	Award of Merit Technical University of Cluj-Napoca, Cluj-Napoca, Romania 1998 - 2002
Scholarships	Socrates Erasmus Scholarship Technical University of Chemnitz, Chemnitz, Germany 2002

Teaching

Teaching Assistant
University of Illinois at Chicago, Chicago, Illinois
2005

Experience

Research Assistant (CMS Experiment)
University of Illinois at Chicago, Chicago, Illinois
2007 - 2011

Research Assistant (D-Zero Experiment)
University of Illinois at Chicago, Chicago, Illinois
2005 - 2006

Engineer
National R&D Institute for Cryogenic and Isotopic Technologies, Ramnicu Vilcea, Romania
2002 - 2004

Intern
Technical University of Chemnitz, Chemnitz, Germany
2002

Presentations

Measurement of the top quark pair production cross section in pp collisions at 7 TeV center of mass energy using secondary vertex b-tagging (Talk)
APS April Meeting, Anaheim, California, USA
2011

Measurement of the charge asymmetry in top quark pair production at 7 TeV (Talk)
XIX International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS2011), Newport News, Virginia, USA
2011

Measurement of the top quark pair production cross section at CMS (Talk)
Physics Fest, University of Illinois at Chicago, Illinois, USA
2011

Status of the CMS Silicon Strip Tracker and commissioning results (Poster)
Position Sensitive Detectors (PSD08), Glasgow, UK
2008

CMS Tracker Detector Performance Studies (Poster)
APS April Meeting, St Louis, Missouri, USA
2008

Tagging Jets at D0 (Talk)
Physics Fest, University of Illinois at Chicago, Illinois, USA
2007

Conference Proceedings

I. Anghel: Measurement of the charged asymmetry in top quark pair production at 7 TeV. , DIS 2011

I. Anghel: Status of the CMS Silicon Strip Tracker and commissioning results. , Nuclear Instruments and Methods in Physics Research Section A, 2009, doi:10.1016/j.nima.2009.01.183

Publications

CMS Collaboration: Measurement of the top-quark pair-production cross section in the lepton+jets channel with the use of b-tagging. Physical Review D, 2011, arXiv:1108.3773

CMS Collaboration: Measurement of the top-quark pair-production cross section in the lepton+jets channel without the use of b-tagging. European Physical Journal C, 2011, http://arxiv.org/pdf/1106.0902

CMS Collaboration: Commissioning and Performance of the CMS Silicon Strip Tracker with Cosmic Ray Muons. Journal of Instrumentation and Detectors, 5:T03008, 2010, http://iopscience.iop.org/1748-0221/5/03/T03008/

CMS Collaboration: Stand-alone Cosmic Muon Reconstruction Before Installation of the CMS Silicon Strip Tracker. Journal of Instrumentation and Detectors, 4:P05004, 2009, http://iopscience.iop.org/1748-0221/4/05/P05004/

CMS Collaboration: Performance studies of the CMS Strip Tracker before installation. Journal of Instrumentation and Detectors, 4:P06009, 2009, http://iopscience.iop.org/1748-0221/4/06/P06009/

M. Anghel, DRT Zahn et al.: VASE and IR spectroscopy: excellent tools to study biaxial organic molecular thin films: DiMe-PTCDI on S-passivated GaAs(100). Thin Solid Films, Vol. 455-456 (1 May 2004), pp. 586-590