Search for high-mass resonant $t\bar{t}$ production in electron+jets events in 7 TeV pp collisions

BY

SAMVEL KHALATIAN
M.S., NRNU MEPhI, Russia, 2004

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, 2013

Chicago, Illinois

Defense Committee:

Cecilia Gerber, Chair and Advisor
Richard Cavanaugh
Mikhail Stephanov
Zhenyu Ye
Michael Schmitt, Northwestern Univ.

To my Parents and Family

ACKNOWLEDGEMENTS

I would like to thank my advisor Cecilia Gerber for all support she gave to me during my research. Cecilia would always listen to my ideas and correct them if these sound too wild. She gave me so much freedom in my research to try new approaches and implement my own ideas of doing the analysis especially in computing software. It was a nice and enjoing exploration of the world of physics analysis.

The PIRE program let me to perform research with CMS Pixels group in Switzerland. It gave me great experience in physics and hardware. Because of PIRE program I did a radiation hardness study that was used later in the Pixels detector upgrade. I met a lot of very kind and nice people. Special thanks to Prof. Alice Bean, Prof. Roland Horisberger and rest of the CMS Pixels group at PSI.

I thank the UIC group at CMS, Prof. Mark Adams, Prof. Nikos Varelas, Prof. Richard Cavanaugh for their help, assistance and many discussions of various topics. Prof. Leonard Apanasevich has been always open for discussion of any kind of questions I had in my research. He gave me so many valuable advices that I can not even count them. The analysis would be impossible without Dr. Derek Strom and all his help with edits. I'd like to thank Dr. Victor Bazterra for great advices and sharing his experience in the research.

My lovely family and relatives were always next to me no matter how hard it was and what difficulties I was going through. Their support and love was giving me strength every single day. I dedicate this thesis to my parents, brother and sister, the ones I love most in my life.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

CHAPTER **PAGE**

LIST OF TABLES

LIST OF FIGURES

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
AOD	**A**nalysis **O**bject **D**ata
ATLAS	**A** **T**oroidal **L**HC **A**pparatu**S** Experiment
BSM	**B**eyond the **S**tandard **M**odel
CERN	European Organization for Nuclear Research
C.L.	**C**onfindence **L**evel
CKM	**C**abibbo-**K**obayashi-**M**askawa
CMS	**C**ompact **M**uon **S**olenoid Experiment
CMSSW	**CMS** **S**oft**W**are
CSC	**C**athode **S**trip **C**hamber
CSV	**C**ombined **S**econdary **V**ertex
CSVHEM	**C**ombined **S**econdary **V**ertex **H**igh **E**fficiency **M**edium
DT	**D**rift **T**ube
EB	**E**CAL **B**arrel
ECAL	**E**lectromagnetic **C**alorimeter
EDM	**E**vent **D**ata **M**odel
EE	**E**CAL **E**ndcap

LIST OF ABBREVIATIONS (Continued)

ES	**E**CAL Pre**S**hower
EWSB	**E**lectro**W**eak **S**ymmetry **B**reaking
JEC	**J**et **E**energy **C**orrection
JER	**J**et **E**energy **R**esolution
JES	**J**et **E**energy **S**cale
JP	**J**et **P**robability
JPT	**J**et **P**lus **T**rack
HB	**H**CAL **B**arrel
HCAL	**H**adron **C**alorimeter
HE	**H**CAL **E**ndcap
HF	**H**CAL **F**orward
HLT	**H**igh **L**evel **T**rigger
HO	**H**CAL **O**uter
IP	**I**mpact **P**arameter
KK	**K**aluza **K**lein
L1 Trigger	**L**evel-1 **Trigger**
LHCb	**L**arge **H**adron **C**ollider **B**eauty Experiment
LINAC2	**Lin**ear **Ac**celerator

LIST OF ABBREVIATIONS (Continued)

LHC	**L**arge **H**adron **C**ollider
LO	**L**eading **O**rder
MC	**M**onte-**C**arlo
ME	**M**atrix **E**lement
NNLO	**N**ext-to-**N**ext-to-**L**eading **O**rder
PAT	**P**hysics **A**nalysis **T**oolkit
PDF	**P**arton **D**istribution **F**unction
PF	**P**article **F**low
PS	**P**roton **S**ynchrotron
PSB	**P**roton **S**ynchrotron **B**ooster
PU	**P**ile **U**p
RAW	**RAW** Data Format
RECO	**RECO**nstructed Data Format
RMS	**R**oot **M**ean **S**quare
RPC	**R**esistive **P**late **C**hamber
SM	**S**tandard **M**odel
SPS	**S**uper **P**roton **S**ynchrotron
SSM	**S**equential **S**tandard **M**odel

LIST OF ABBREVIATIONS (Continued)

SSV	**S**imple **S**econdary **V**ertex
SSVHE	**S**imple **S**econdary **V**ertex **H**igh **E**fficiency
SSVHP	**S**imple **S**econdary **V**ertex **H**igh **P**urity
TC	**T**rack **C**ounting
TEC	**T**racker **E**nd **C**ap
TIB	**T**racker **I**nner **B**arrel
TID	**T**racker **I**nner **D**isk
TOB	**T**racker **O**uter **B**arrel

SUMMARY

In this thesis we present a model-independent search for the production of heavy resonances with mass greater than 1 TeV decaying to top quark pairs. Using data samples corresponding to 5.0 fb^{-1} of integrated luminosity of pp collision data recorded with the Compact Muon Solenoid experiment in 2011 at $\sqrt{s} = 7$ TeV, we select events containing one electron and at least two jets and look for excess above Standard Model background prediction in the top quark pair invariant mass spectrum. The high transverse momenta of the top quarks originating from such decays result in an event topology which requires a dedicated event selection and reconstruction of the invariant top quark pair mass. We use a $\chi 2$ method in the reconstruction and selection of top quark pairs and apply b-tagging to improve sensitivity. In the absence of evidence for a signal, we evaluate 95% C.L. upper limits on $\sigma\left(pp \to Z' \to t\bar{t}\right) \cdot \text{BR}$ as a function of the invariant mass of the resonance.

CHAPTER 1

THEORETICAL INTRODUCTION

1.1 The Standard Model

The Standard Model (SM) is the physical theory of elementary particles which is currently adopted to provide a quantitative description of three of the four interactions in nature: electromagnetism, weak interaction, and the strong force. The gravitational force is not included in this model because it is by far weaker than the other fundamental interactions. A large number of experiments tested and confirmed the SM predictions with high precision in the past.

The Standard Model is built around the gauge group $\mathrm{SU(3)_C \times SU(2)_L \times U(1)_Y}$. The smaller subgroup $\mathrm{SU(3)_C \times U(1)_{EM}}$ manifests at low energies. Therefore an electroweak symmetry breaking (EWSB) should be present, the process of reducing the $\mathrm{SU(2)_L \times U(1)_Y}$ to $\mathrm{U(1)_{EM}}$. Each subgroup in the SM generates one of the known interactions to date.

The $\mathrm{SU(3)_C}$ group is the symmetry that generates the strong interaction. Its conserved quantum number is the color, which is equivalent to the electromagnetic charge for strong interactions. There are three color charges available: red, green, blue, and corresponding anticolors. Each quark carries one color and antiquark carries anticolor. All naturally occurring particles are colorless, that is either the total amount of each color is zero or all three colors are present in equal amounts. Because of this rule the only possible colorless combinations are meson ($q\bar{q}$), baryon (qqq), or antibaryon ($\bar{q}\bar{q}\bar{q}$). The strong interaction is mediated by colored

particles, gluons. Gluons carry one positive unit of color and one negative unit, anticolor. Because of the color only quarks participate in the strong interaction.

The $SU(2)_L \times U(1)_Y$ group is the symmetry that generates the electromagnetic and weak interactions. It is described by a unified electroweak theory, the Glashow-Salam-Weinberg model (3). The mediators generated by this symmetry are the vector bosons $W^{\pm}$, Z, and γ. The photon is massless but all mediators of the weak interactions are massive particles. All quarks and leptons participate in the weak interaction, while only electrically charged particles interact electromagnetically. The $SU(2)_L$ group symmetry gives rise to a conserved quantum number called weak isospin L, while the $U(1)_Y$ group has hypercharge Y quantum number. These quantum numbers are connected to the electrical charge by the Gell-Mann-Nishijima formula:

$$Q = L^3 + \frac{1}{2}Y \tag{1.1}$$

where L^3 is the third component of isospin.

The theoretical description of the EW interaction does not contain the mass terms for the $W^{\pm}$ and Z bosons, in contrast with the experimental observations. This problems is resolved by enlarging the particle content of the SM with the introduction of a new field, the Higgs field. Its potential is invariant under the local gauge transformations. The new field introduces a spontaneous symmetry breaking that is called the Higgs Mechanism, which leads to the inclusion of the mass terms for the weak gauge bosons in the theory but with addition of a new particle in the SM, the Higgs boson.

In the SM the elementary particles composing matter are spin 1/2 fermions. These are leptons and quarks with their respective antiparticles. There are three generations, or families, of fermions, shown in Figure 1. Each generation is a doublet of particles associated to the isospin quantum number. All the particles of three families have been discovered and no experimental evidence of the existence of a fourth generation has been found so far. Each next generation contains heavier particles compared to the previous family.

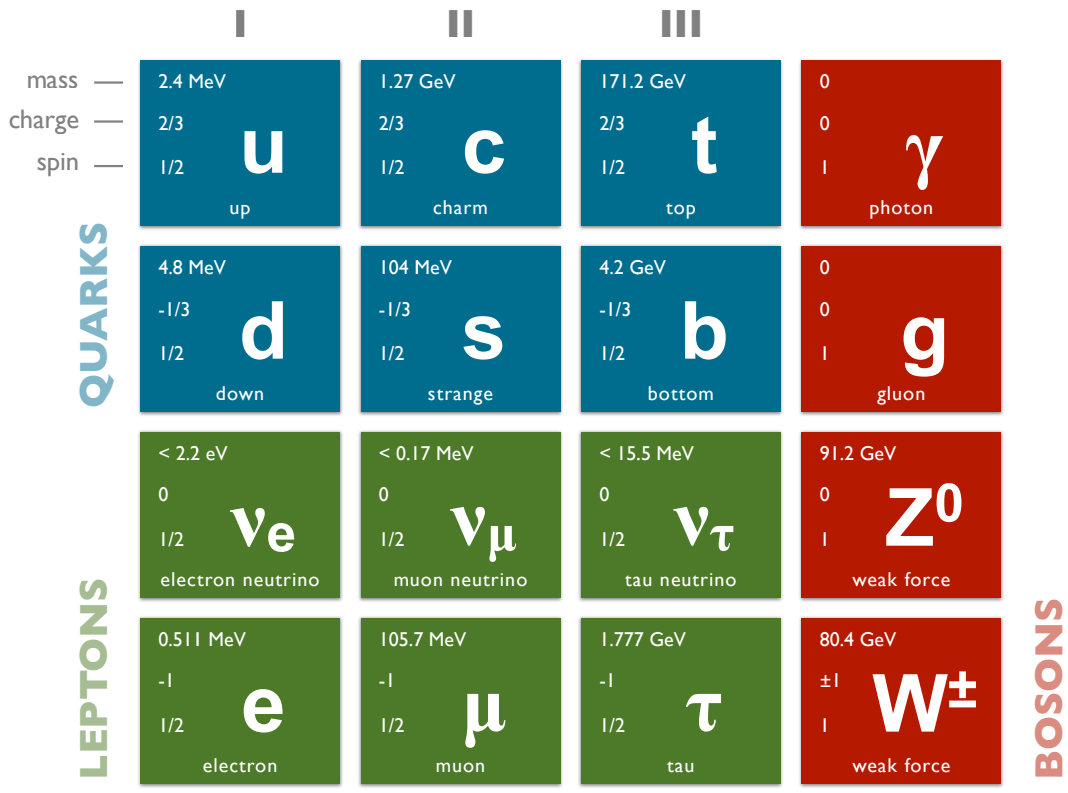


Figure 1: A schematic view of the Standard Model of particle physics.

The lepton generations include a particle with charge $-1e$, the electron, muon, tau, and a neutral particle, neutrino of corresponding flavor, e.g. electron neutrino, muon neutrino, and tau neutrino. Neutrinos are assumed to be massless within the SM.

The quarks are observed only in bound states called hadrons of two types: mesons and baryons. Each quark generation contains a particle with $+2/3e$ charge, called an up-type quark, and a particle with charge $-1/3e$, a down-type quark.

Leptons and quarks interact with each other through three fundamental interactions, which are mediated by spin 1 bosons. The force carriers are shown in Figure 1 and represented by: 8 gluons for the strong interaction, $W^{\pm}$, Z bosons for the weak force, and photon for the electromagnetic interaction. The number of mediator bosons for an interaction is determined by the symmetry group associated with it.

1.1.1 Top Quark

The top quark is the heaviest of the six quarks in the SM. It has a charge of $2/3e$, transforms as a color triplet under the $\mathrm{SU(3)_C}$ group of the strong interactions and is the weak isospin partner of the bottom quark. The top quark was discovered in 1995 at the Fermilab Tevatron $p\bar{p}$ collider at the center-of-mass energy $\sqrt{s} = 1.8$ TeV (4). Its mass has been precisely measured in a large number of experiments and the combined value according to the Review of Particle Physics by the Particle Data Group in 2012 is $\mathrm{M_{top}} = 173.5 \pm 0.6(\mathrm{stat}) \pm 0.8(\mathrm{syst})$ GeV (5).

Top quarks can be produced singly or in pairs in pp collisions at the Large Hadron Collider (LHC).

Figure 2 shows the leading-order (LO) Feynman diagrams for $t\bar{t}$ production via quark-antiquark annihilation $q\bar{q} \rightarrow t\bar{t}$ and gluon-gluon fusion $gg \rightarrow t\bar{t}$. The q and $\bar{q}$ in the quark-antiquark annihilation process represent the same type quarks, for example u and $\bar{u}$. The gluon-gluon interactions take place in about 85% of the $t\bar{t}$ process at the LHC, while quark-antiquark interactions represent only 15%.

Figure 3 shows the LO Feynman diagrams for the single top production in the s-, tW-, and t-channels. In the s- and t-channels the q and q' are up- and down-type quarks due to electromagnetic charge conservation law, for example u and d.

Within the SM the top quark decays via the weak interaction to a W boson and a down-type quark: d, s, or b. In 1963, in attempt to explain absence of the strangeness-changing weak interaction, Cabibbo proposed an idea which was further developed by Kobayashi and Maskawa, that the observed quarks, represented by the strong interaction eigenstates, were composed of a linear combination of the quarks in the weak eigenstates. Thus, instead of:

$$\begin{pmatrix} u \\ d \end{pmatrix} , \begin{pmatrix} c \\ s \end{pmatrix} , \begin{pmatrix} t \\ b \end{pmatrix} \tag{1.2}$$

the weak force couples:

$$\begin{pmatrix} u \\ d' \end{pmatrix} , \begin{pmatrix} c \\ s' \end{pmatrix} , \begin{pmatrix} t \\ b' \end{pmatrix} \tag{1.3}$$

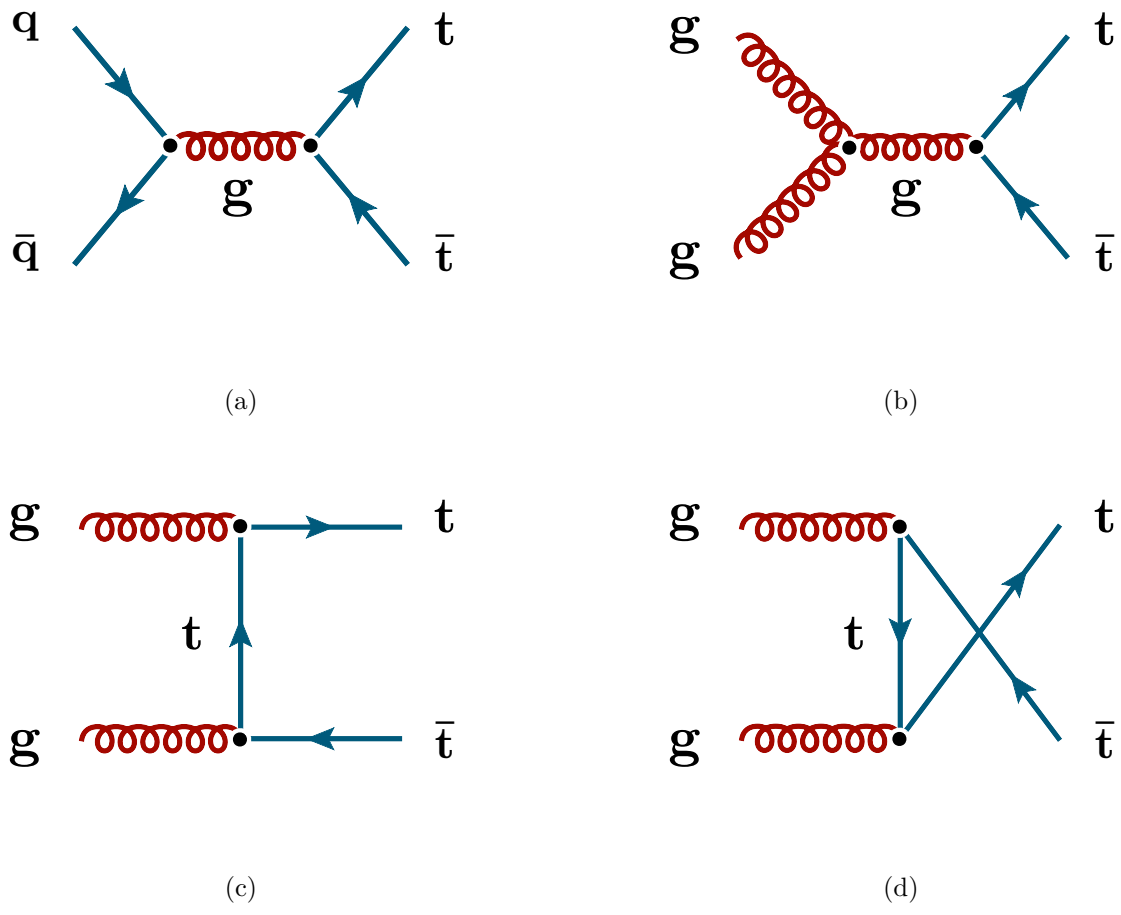


Figure 2: Leading-order Feynman diagrams for $t\bar{t}$ production via (a) quark-antiquark annihilation, and (b-d) gluon-gluon fusion. The q and $\bar{q}$ in the quark-antiquark annihilation (a) are the same type quarks, for example u and $\bar{u}$.

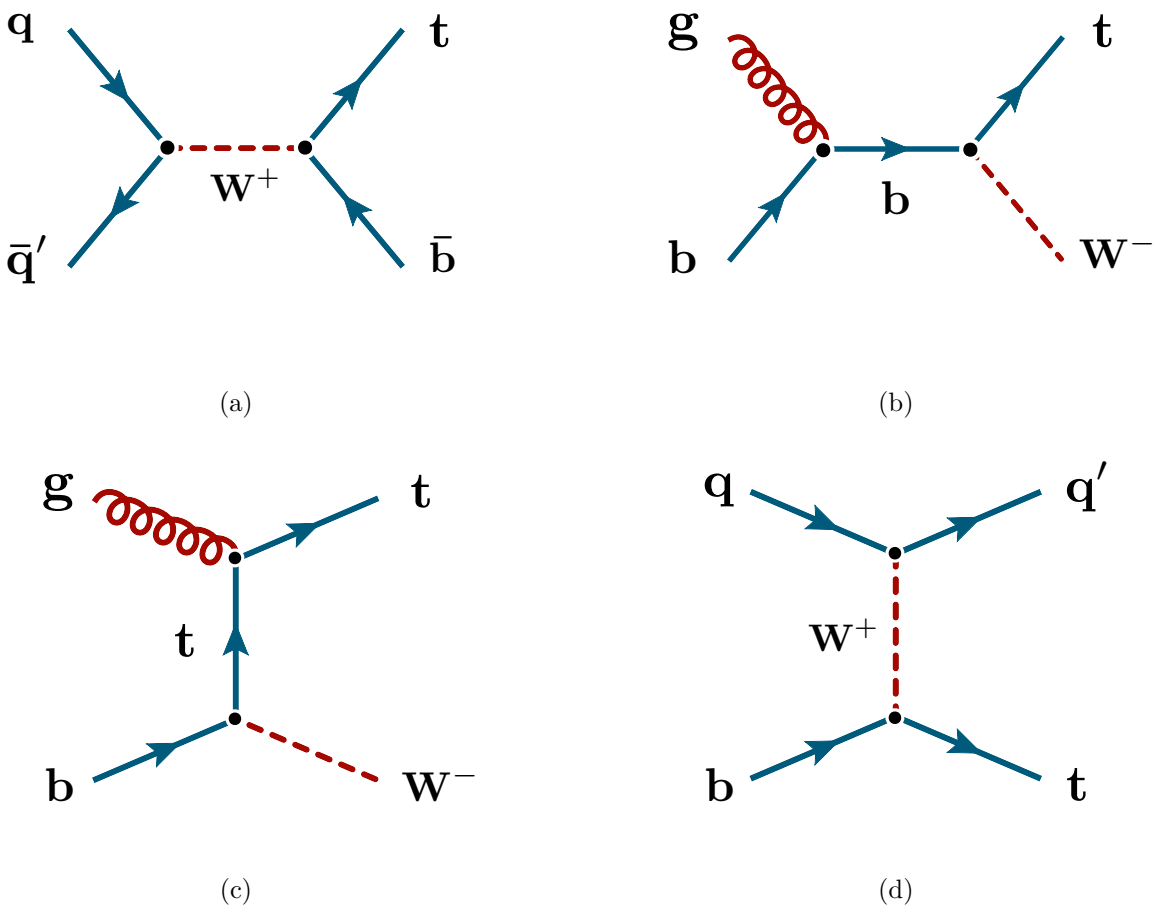


Figure 3: Leading-order Feynman diagrams for single top production in (a) s-channel, (b-c) tW-channel, and (d) t-channel. In the s- and t-channels the q and q' are up- and down-type quarks due to electromagnetic charge conservation law, for example u and d.

where d', s', and b' are linear combination of the physical quarks d, s, and b:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.4}$$

where V_{ij} specifies the coupling of the quarks i and j. The mixing matrix is called the Cabibbo-Kobayashi-Maskawa (CKM) matrix. The matrix is nearly diagonal. Most of the coupling is between the pairs (u, d), (c, s), and (t, b). The V_{tb} is much larger than V_{td} and V_{ts}. Thus, the decays to d and s quarks are highly suppressed compared to decay to b. The decay $t \to W^+b$ takes place almost 100% of the time and is the main decay mode for the top quark.

We distinguish $t\bar{t}$ channels according to the decay of the two W bosons present in the top pair decay. Figure 4 shows all possible decay channels with branching fractions given in per cent according to the theoretical predictions at tree level. This analysis focuses on the electron+jets channel which represents 15% of all $t\bar{t}$ decays.

Figure 5 shows the next-to-leading order cross sections for different processes for proton-antiproton collisions at $\sqrt{s} < 4$ TeV and proton-proton collisions at $\sqrt{s} > 4$ TeV (6). The design instantaneous luminosity for the LHC is 10^{34} cm^{-2}s^{-1}, a factor of ten higher than what is shown. The top pairs have a very large cross section, with more than 10^5 top quark pairs produced per year at the nominal luminosity in pp collisions at $\sqrt{s} = 7$ TeV. The largest background processes for the $t\bar{t}$ production originate from W + jets boson production.

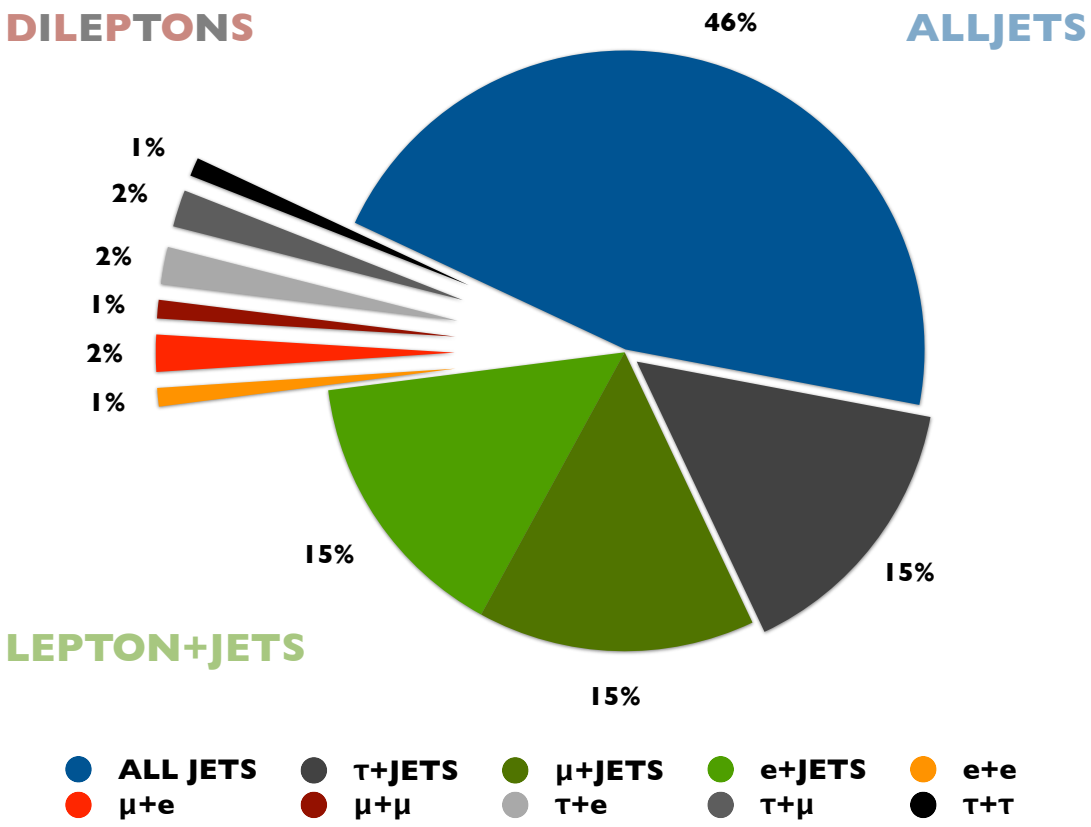


Figure 4: Branching fractions of the $t\bar{t}$ decay channels. The values are shown in per cent values according to the theoretical predictions at tree level.

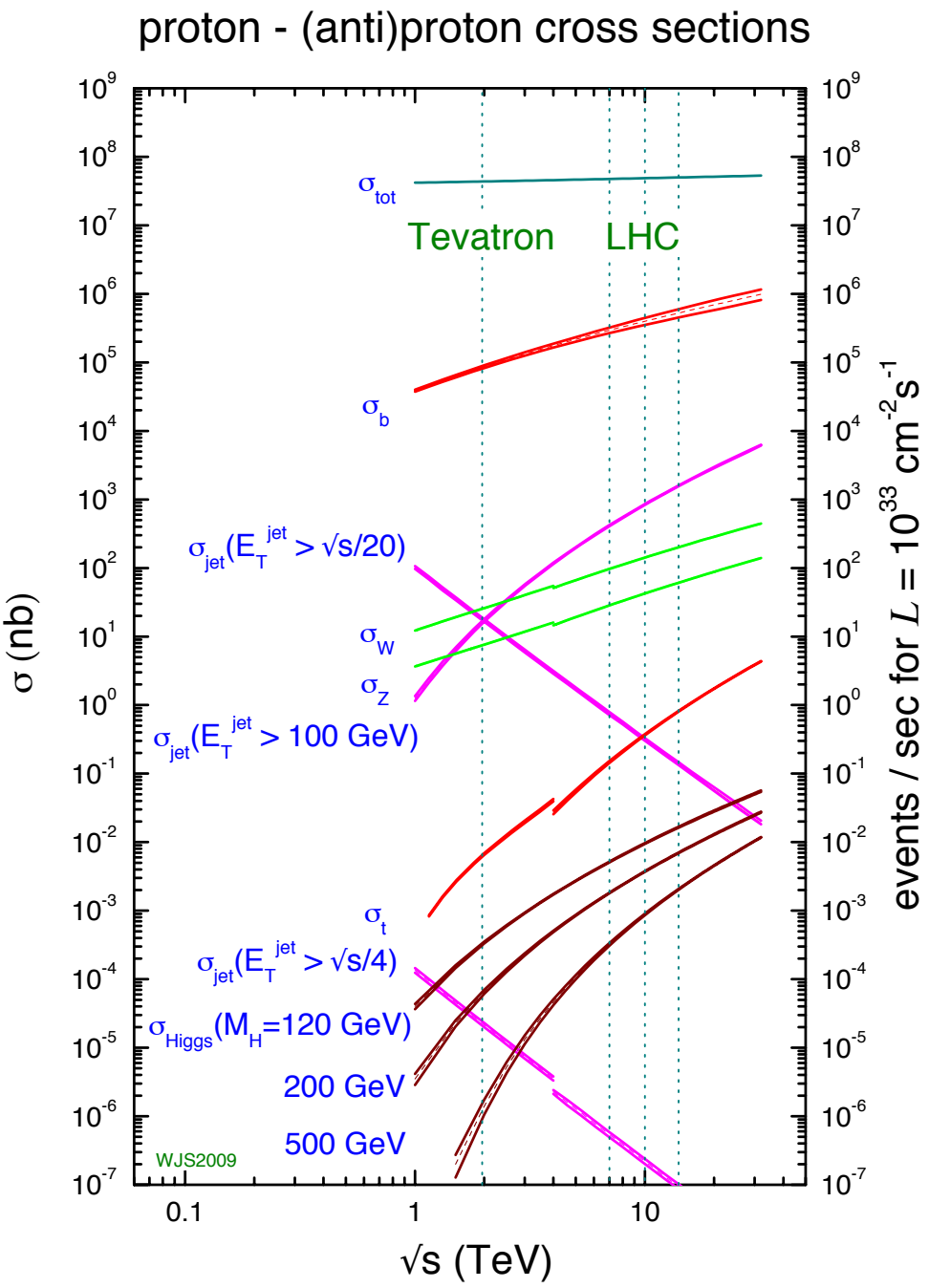


Figure 5: Next-to-leading order cross sections for different processes within the SM at the Tevatron and LHC colliders. The proton-antiproton collisions are shown at $\sqrt{s} < 4$ TeV and proton-proton collisions at $\sqrt{s} > 4$ TeV.

Because of its high mass the top quark could be the key to understand the dynamic origin of how particle masses are generated by the mechanism of electroweak symmetry breaking. In fact, the masses of the Higgs boson, the W boson and the top quark are closely related through higher order corrections to various physics processes.

1.2 Beyond the Standard Model

The SM description of particle physics is a huge success. The Higgs mechanism solves the EWSB problem. However, there are still open questions left that can not be answered by SM (7). For this reason, new physics theories Beyond the Standard Model (BSM) are proposed to address some of the questions, including:

- The Higgs boson mass parameter is extremely sensitive to quantum corrections. As a result, attempts to extrapolate the SM to energies much above the electroweak scale lead to the gauge hierarchy problem.
- Most of the matter in our Universe is composed of a non-relativistic particle species which interact only very feebly with the SM. The SM is unable to explain the dark matter in the universe.
- The asymmetry of visible matter over antimatter is another reason for new physics near the electroweak scale, which can potentially give rise to this baryon asymmetry.

There are additional shortcomings of the SM including an explanation of flavor mixing and the masses of neutrinos, a quantum description of gravity, and others. Many new physics BSM

theories were proposed to address one or more challenges from the list. Based on current data, it is not clear which of them, if any, are correct.

A fundamental scalar Higgs boson is not the only way to introduce the electroweak symmetry breaking. New strong interactions can also play this role.

Several of the BSM theories predict the existence of heavy resonances that manifest themselves as an additional resonant component to the Standard Model top quark pair production. These include models with massive color-singlet Z-like bosons in the extended gauge theories (8; 9; 10), colorons (11), or axigluons (12; 13), models in which a pseudoscalar Higgs may couple strongly to top quarks (14), and models with extra dimensions, such as Randall-Sundrum (15), and ADD models (16), where TeV-scale gravitons decay preferentially to top quark pairs.

We perform our search for the resonant production of top quark pairs independently of the exact model by analyzing the reconstructed $t\bar{t}$ invariant mass spectrum. The limits can be extracted as a mass-dependent upper limit on the resonance production cross section times branching fraction or as a limit on the mass of the searched resonance. In the latter case, a specific model that predicts the cross sections for the resonance must be chosen.

The analysis presented in this thesis uses the topcolor Z' model as a benchmark (17) that was chosen for the searches for heavy resonances at the Tevatron in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. We use the predicted cross sections by that model for the LHC energies for comparison (1).

In this model the QCD gauge group $\mathrm{SU(3)_C}$ is embedded into a larger structure $\mathrm{SU(3)_1} \times \mathrm{SU(3)_2}$ with couplings h_1 and h_2, respectively. The $\mathrm{SU(3)_2}$ couples only to the third generation of quarks, while the $\mathrm{SU(3)_1}$ couples only to the first and second families of quarks. The coupling

constants must satisfy a condition $h_2 \gg h_1$. This choice of couplings results in the topgluons coupling strongly to top quarks and weakly to the light quarks. The unbroken $\mathrm{SU(3)_C}$ symmetry guarantees that the massless bosons couple universally to all of the quarks. The symmetry $\mathrm{SU(3)_1 \times SU(3)_2}$ has to be broken to the Standard Model $\mathrm{SU(3)_C}$. The breaking of the new symmetry as well as the EWSB is achieved by topgluons which couple mainly to $b\bar{b}$ and $t\bar{t}$. To enhance the formation of the $t\bar{t}$ condensate, while blocking the formation of the $b\bar{b}$ condensate, a new mechanism is required, so that the b-quark is light while top is heavy. For this a neutral gauge boson Z' is introduced with an attractive interaction between $t\bar{t}$ and a repulsive interaction between $b\bar{b}$. The Z boson of the Standard Model does precisely this and could itself provide the new mechanism. However, the SM coupling constant g_1 is so small that fine-tuning would be required to achieve the mechanism in the presence of a large h_2. For this reason a new Z' boson is introduced. The predicted values of the cross section times branching ratio into $t\bar{t}$ as a function of the resonance mass depend on the resonance width (1) as shown in Figure 6.

The upper limits on the production cross section of narrow resonances decaying to $t\bar{t}$ pairs for Z' masses below 900 GeV at 95% Confidence Level (C.L.) are set by searches performed at the Tevatron collider (18; 19; 20; 21; 22; 23; 24). We focus on a search for the BSM heavy particle with mass greater than 1 TeV decaying to top quark pairs in the electron+jets topology $t\bar{t} \to (W^+b)(W^-\bar{b}) \to (e^+\nu_e b)(q_1\bar{q}_2\bar{b})$ (or charge conjugate) wherein one W boson decays to an electron and neutrino, and the other W boson decays hadronically as shown on Figure 7. Because of the large BSM particle mass we expect the top quark decay products to be boosted, do not require the lepton to be isolated, and expect large neutrino energy. We also combine

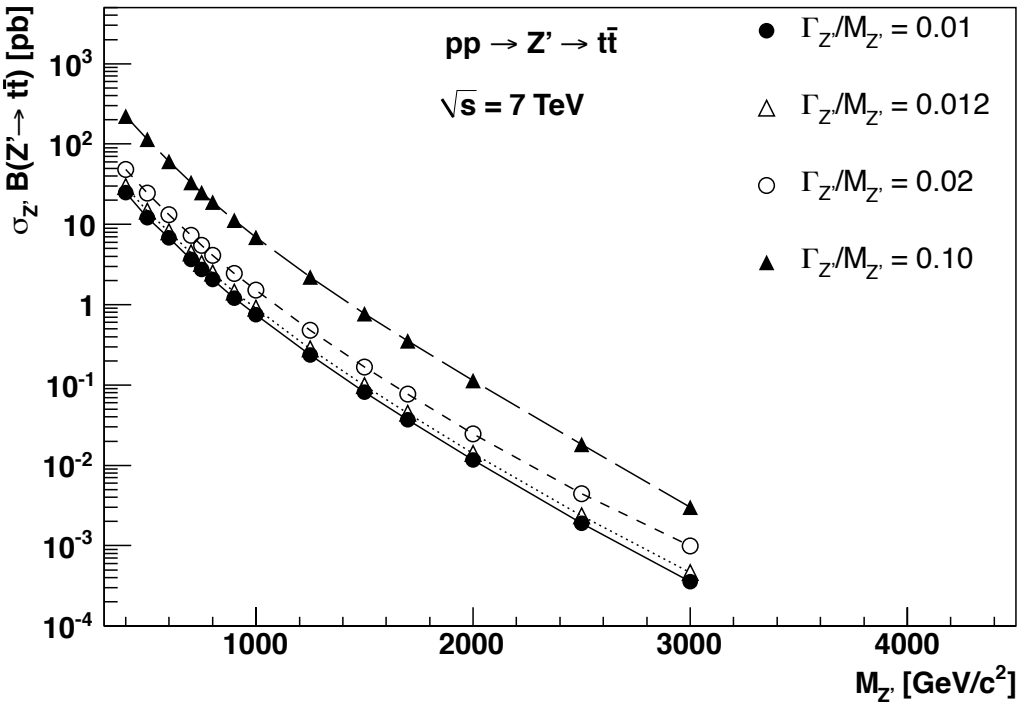


Figure 6: Cross sections at the LHC at $\sqrt{s} = 7$ TeV, for $\sigma_{Z'} \cdot B(Z' \to t\bar{t})$, with different choices of the resonance width.

the results with the Compact Muon Solenoid (CMS) complimentary analysis in the muon+jets channel (25), and the search close to the $t\bar{t}$ mass threshold in the lepton+jets channel (26). Finally, we compare the results to all hadronic (27) and di-lepton channels (28). The latter analysis focuses on the $t\bar{t}$ mass in $0.5 - 1$ TeV range.

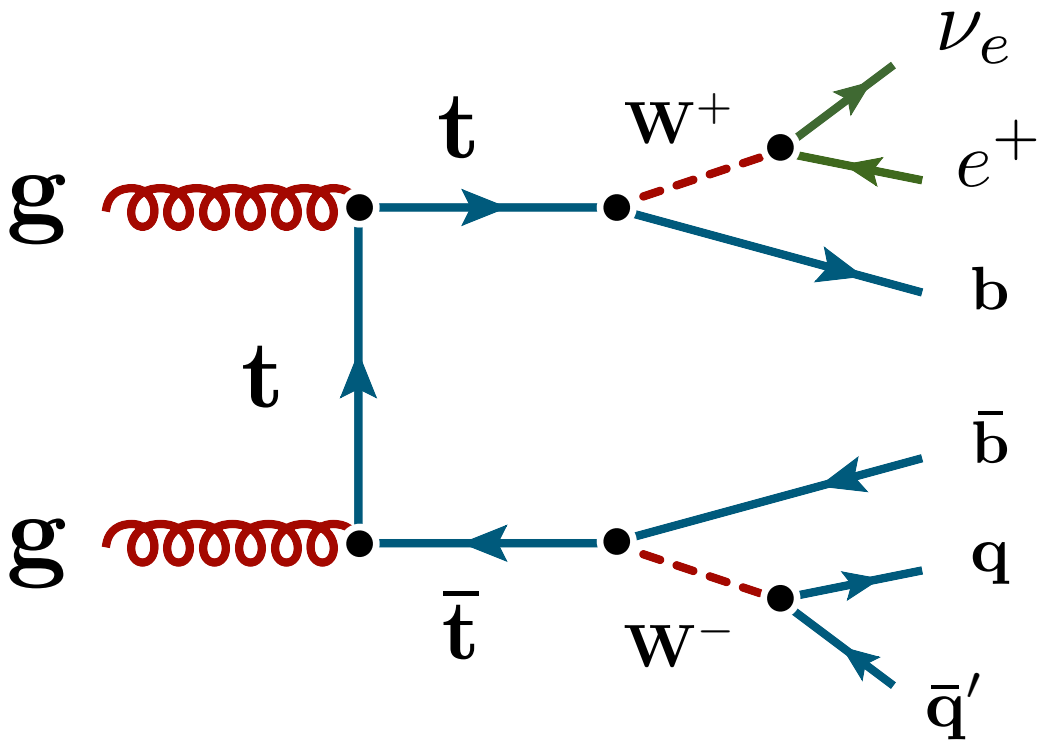


Figure 7: Production of the $t\bar{t}$ pair with subsequent decay of the top quarks in the electron+jets channel.

CHAPTER 2

THE LARGE HADRON COLLIDER AND THE CMS DETECTOR

2.1 Overview

The Large Hadron Collider (29) is currently the world's highest energy and largest particle collider. Located at CERN near Geneva, Switzerland, it accelerates protons at a design center-of-mass energy of 14 TeV that collide at four interaction points, located in the center of four detectors: CMS (30; 31), ATLAS (32; 33), ALICE (34), and LHCb (35).

2.2 Large Hadron Collider

The LHC consists of a ring-shaped tunnel of 27 km circumference about 100 m below the surface. More than one thousand magnets are installed inside the tunnel to bend the proton beams around the ring.

Figure 8 shows the accelerator chain. The protons are separated from electrons in hydrogen atoms and then accelerated in four cycles before they enter the LHC ring. The first step is a Linear Accelerator (LINAC2). It accelerates protons up to 50 MeV and sends them to a Proton Synchrotron Booster (PSB). The PSB accelerates the beams to 1.4 GeV and injects protons into a Proton Synchrotron (PS). The beams are accelerated up to 25 GeV in the PS and are forwarded to the last step before the LHC, a Super Proton Synchrotron (SPS). The SPS increases the beam energy up to 450 GeV.

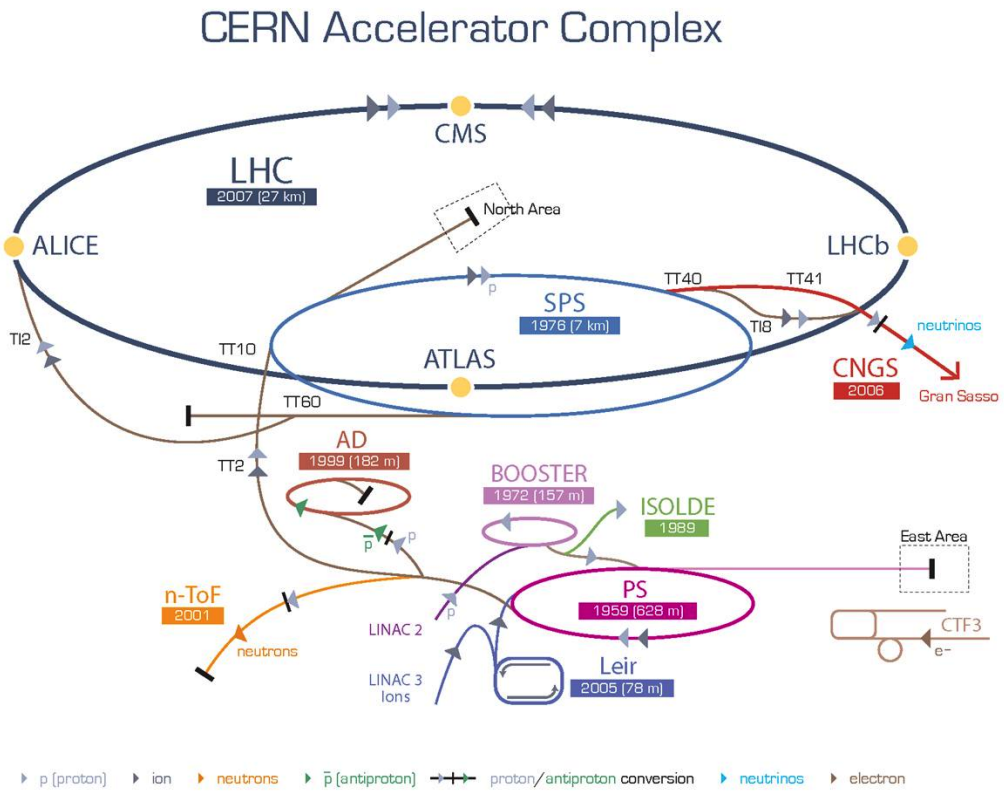


Figure 8: A schematic overview of the LHC accelerator chain. The key components are: the Linear Accelerator - LINAC2, the Proton Synchrotron Booster - PSB, the Proton Synchrotron - PS, the Super Proton Synchrotron - SPS, and the Large Hadron Collider ring - LHC.

After the SPS, proton beams are finally transferred to the LHC ring (both in a clockwise and an anticlockwise directions). The LHC was operating at the center-of-mass energy of 7 TeV in 2011 with 50 ns bunch separation distance. The collider will reach the design energy of $\sqrt{s} = 14$ TeV in 2015 after the shutdown of $2013 - 2014$. In addition to the center-of-mass energy, a luminosity is another important quantity to measure the performance of a collider. The instantaneous luminosity is the measure of the collision rate per unit area per unit time, usually measured in $\mathrm{cm}^{-2}\mathrm{s}^{-1}$, $\mathcal{L} = fb\frac{n_1 n_2}{4\pi\sigma_x\sigma_y}$, where f is the revolution frequency in Hertz, b is the number of bunches, n_i is the number of protons per bunch, and σ_i is the root mean square (RMS) beam size in the transverse plane to the direction of the beam line at the interaction point. The integrated luminosity is the integral of the instantaneous luminosity with respect to time and is usually measured in cm^{-2} or inverse barns, b^{-1}. One inverse barn is equal to $10^{-24}\ \mathrm{cm}^{-2}$.

2.3 Compact Muon Solenoid Detector

Figure 9 shows the Compact Muon Solenoid detector, one of the two general-purpose detectors at the LHC. Compared to ATLAS, CMS is two times smaller and weighs about two times more. For this reason it is called compact. A detailed description of the CMS detector can be found elsewhere (36).

One of the main goals of the CMS scientific program is the discovery of the Higgs boson. This imposes a set of requirements, driven by a search, and can be summarized as follows:

- excellent muon identification over a large range of momenta and angles,

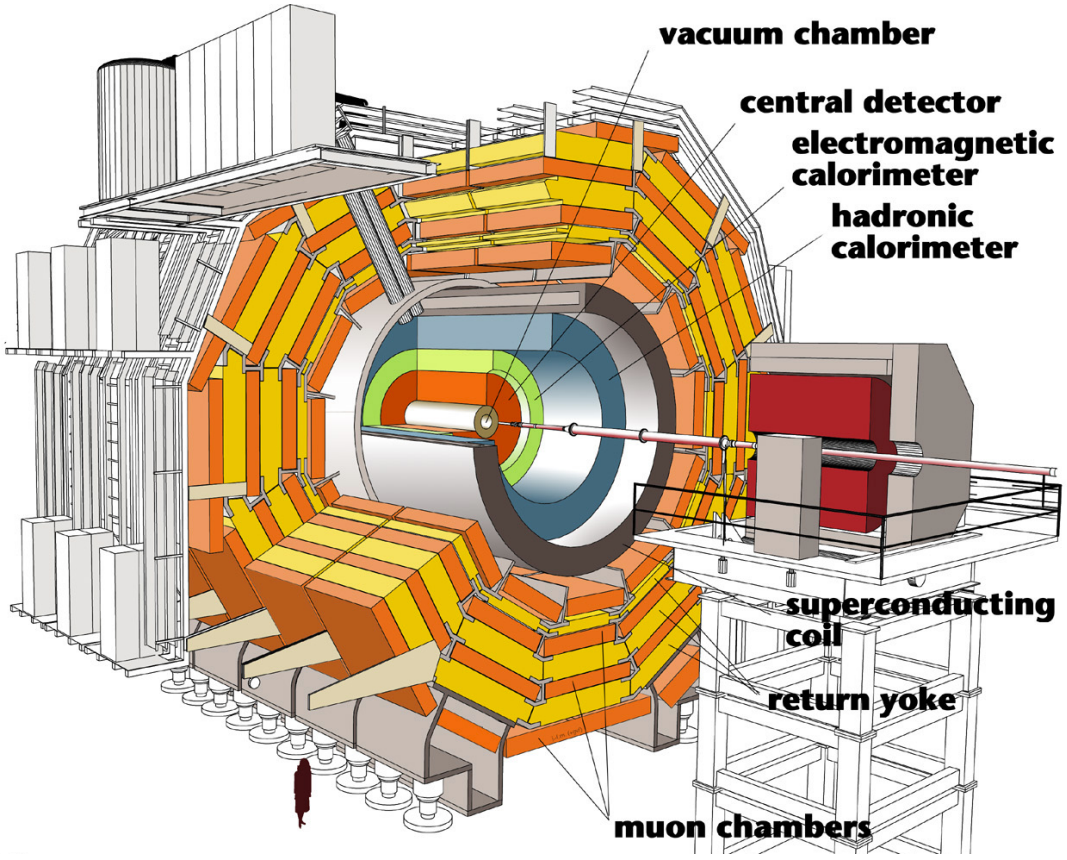


Figure 9: The CMS detector at the LHC.

- excellent reconstruction efficiency and momentum resolution for charged particles in the inner Tracker,
- outstanding electromagnetic energy resolution, diphoton and dielectron mass resolution with high angular coverage, and efficient photon and lepton isolation at high luminosities,
- excellent missing transverse energy and dijet mass resolution.

The CMS detector has a cylindrical shape, symmetrical around the beam line, with a radius of 7.5 m, a total length of 22 m, and weighs about 12 500 tons. It is divided into a central section, called barrel, and two endcaps. The barrel is made of several layers coaxial to the beam axis. The endcaps close the barrel at the ends by two hermetic discs orthogonal to the beam.

As can be seen in Figure 9, the innermost part of the detector is the silicon Tracker. Moving outwards from the beam line, next is a crystal electromagnetic calorimeter, a hadron calorimeter, and the super conducting solenoid magnet, in the return yoke of which the muon drift chambers are inserted.

CMS uses a Cartesian coordinate system with origin centered in the nominal collision point at the center of the detector. The x axis points towards the center of the LHC ring, the z direction coincides with the CMS cylinder axis, and the y axis points upwards, towards the surface.

The cylindrical symmetry of the apparatus drives the use of a pseudoangular reference system, defined by (r, ϕ, η), where r is the radial distance from the beam axis, ϕ is the azimuthal angle with respect to the x axis, and η is the pseudorapidity, $\eta = -\ln\tan\frac{\theta}{2}$, where θ is the polar angle with respect to the z axis. The pseudorapidity is a good approximation for the

rapidity y of the particle at high energies, $y = \frac{1}{2}\ln\left(\frac{E+p_l}{E-p_l}\right)$, where E is the energy of the particle, and p_l is the projection of its momentum on the beam axis, and is equal to p_z in the CMS coordinate system. The rapidity is an invariant variable under Lorentz boosts along the beam axis. According to the definition, the pseudorapidity is equal to zero at $\theta = \pi/2$, and increases in absolute value when approaching the beam pipe, asymptotically reaching infinity at $\theta = 0$.

We also use transverse energy, $E_T = E\sin\theta$, and transverse momentum of a particle, $p_T = \sqrt{p_x^2 + p_y^2}$. The spatial separation of two particles, $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$, represents another Lorentz boost invariant variable, where $\Delta\phi$ and $\Delta\eta$ are the differences of the corresponding variables of two particles.

Figure 10 illustrates a slice of the CMS detector and some high-energy particles passing through. The hermetic design of the detector allows the detection of almost all particles generated in the beam collision.

2.3.1 Magnet

One of the most important parts in the detector design is the configuration and parameters of the magnetic field for the charged particle momenta measurement.

CMS has a superconducting solenoid magnet, a 13 m long superconducting cylindrical Niobium-Titanium coil, with a diameter of 5.9 m. It provides a uniform magnetic field of 3.8 T at its center, carrying a current of 18 kA and a total stored magnetic energy of 2.4 GJ. The magnetic flux is returned by a saturated iron yoke, which also works as a mechanical support structure of the detector.

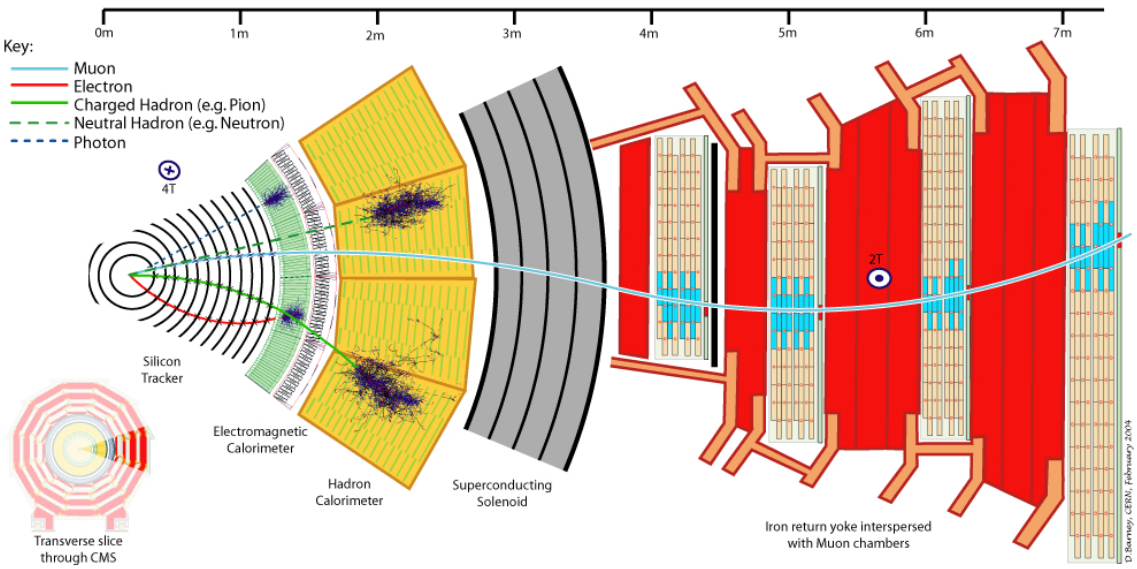


Figure 10: A slice of the CMS detector showing the particles interacting with sub-detectors.

The solenoid design of the magnet in the CMS detector has a number of advantages over toroid design:

- the field is parallel to the beams and the bending of the charged particle track is in the transverse plane only. The position of the interaction point, vertex, in this plane is known to a precision of 20 μm. The strong bending in the transverse plane facilitates the task of triggers based on tracks pointing back to the vertex,
- a particle momentum measurement starts at $r = 0$ cm, while for a toroid it starts after the absorber, typically after a few meters.

2.3.2 Tracker

The closest system of the CMS detector to the interaction point is the Tracker. Its design goal is to reconstruct tracks of the charged particles. The sign of the charge and the momentum of the particle can be derived from the curvature of the track inside the magnetic field. In addition, it is possible to reconstruct primary and secondary vertices in the event.

A schematic overview of the Tracker is shown in Figure 11. It consists of 1 440 pixel and over 15 000 strip detector modules with an active area of about 200 m^2. The CMS Tracker covers the pseudorapidity range $|\eta| < 2.5$.

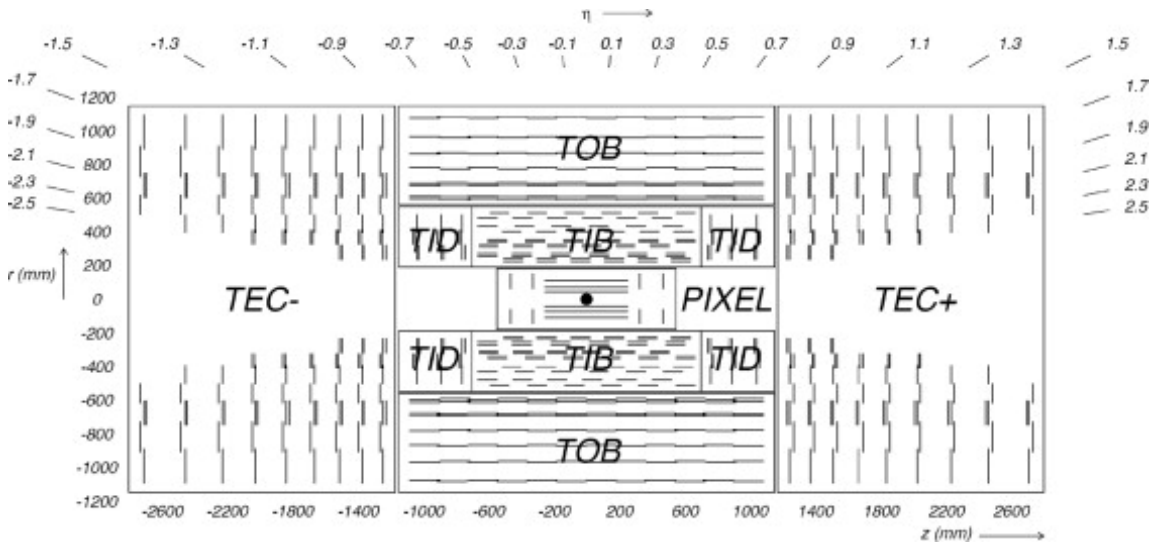


Figure 11: A schematic cross section through the CMS Tracker. Each line represents a detector module and double lines indicate stereo modules.

The CMS Tracker consists of two parts: pixel and silicon microstrip detectors. Each part has its own design for the barrel and the endcap.

Because of the large magnetic field, the low momentum charged particles can not escape the magnetic field and remain in the very first layers of the Tracker. This contributes to high particle flux and high track multiplicity. The high collision rate demands for a fast response.

The Tracker provides robust, efficient, and precise reconstruction of the charged particle trajectories inside the magnetic field. The nominal momentum resolution is 0.7% at 1 GeV, 5.0% at 1000 GeV in the central region and the impact parameter resolution for high-momentum tracks is 10 μm (37).

2.3.2.1 Pixel Detector

The Pixel detector consists of 3 barrel layers and 2 endcap disks on each side. The 3 barrel layers are located at mean radii of 4.4 cm, 7.3 cm and 10.2 cm, and have a length of 53 cm. The 2 endcap disks extend from 6 cm to 15 cm in radius, and are placed on each side at $|z| = 34.5$ cm and $|z| = 46.5$ cm. Each pixel has dimension of 100×150 μm^2.

The design ensures that each charged particle produced in $|\eta| < 2.2$ gives rise to at least two hits in the Pixel detector as it passes through the detector. In total, the Pixel detector has 65 million channels. One read-out chip controls a group of 52×80 pixels. The read-out is only carried out if the event contains interesting information from the physics point of view, which is triggered by electronics.

2.3.2.2 Silicon Microstrip Detector

The Silicon Microstrip detector starts from $r = 20$ cm to $r = 116$ cm. Different silicon sensors for inner and outer parts, for the barrel and the endcap are used. These can be seen in Figure 11.

The inner most part, the Tracker Inner Barrel (TIB), consists of four layers of microstip modules with thickness of 320 μm and a strip pitch ranging from 80 μm to 120 μm. The first two layers provide a stereo measurement in order to be able to measure in both $r - \phi$ and $r - z$ coordinates. The last layer of TIB is located at $r = 60$ cm. The TIB is complemented by three disk layers at each end, Tracker Inner Disk (TID). Some of the layers in the endcap are also equipped with stereo modules.

The TIB and the TID are surrounded by the Tracker Outer Barrel (TOB), which consists of 6 layers and extends to $r = 116$ cm. Lower tracks multiplicity allows the use of thicker sensors with width of 500 μm. The strip pitch in this region varies from 120 μm to 180 μm. Two layers with stereo modules in the TOB help to improve the position measurement.

Beyond the z range of the TOB, $|z| \leq 118$ cm, the Tracker End Caps (TEC+ and TEC-), cover the region 124 cm $\leq |z| \leq 282$ cm. The sign in the name denotes the location along the z axis. Each side has 9 disks and the thickness of the 3 innermost rings is 320 μm and 500 μm for the 6 outer disks.

2.3.3 Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) surrounds the Tracker of the CMS as seen in Figure 12. It is split into a barrel and endcaps. The barrel (EB) extends up to $|\eta| < 1.4442$ and

the endcap (EE) part of the ECAL covers the region $1.556 < |\eta| < 3.0$, as shown in Figure 13. A gap between $1.4442 < |\eta| < 1.556$ in the endcap is not instrumented.

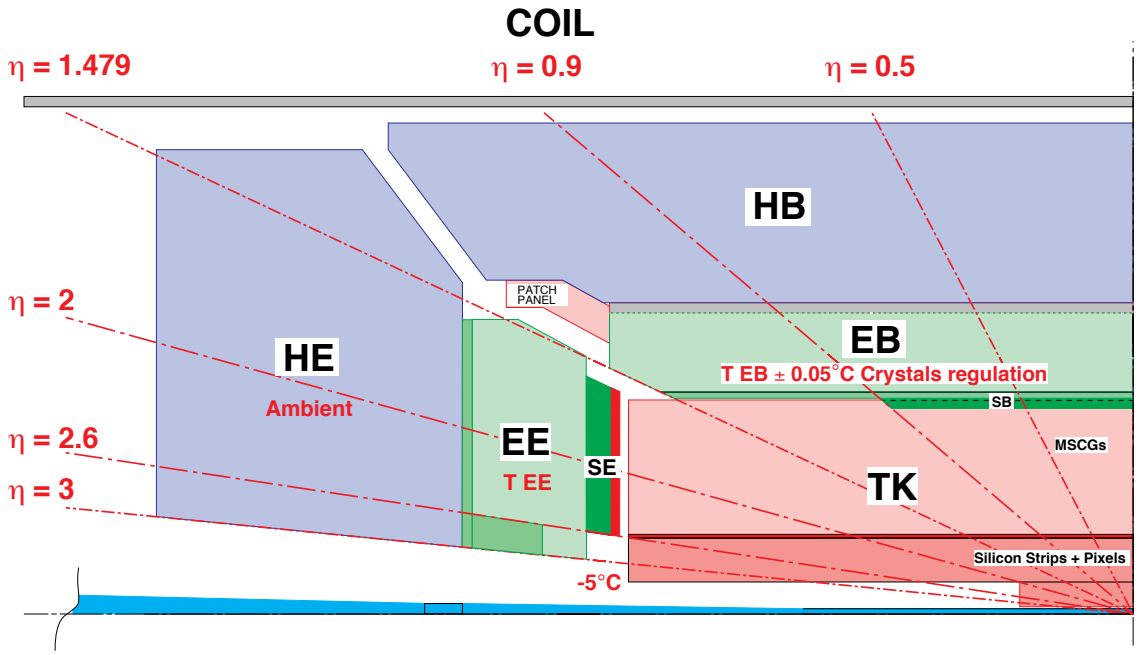


Figure 12: A schematic view of a quarter of the CMS Tracker, the Electromagnetic and the Hadron Calorimeters.

A preshower (ES) is placed in front of the crystal calorimeter over much of the endcap pseudo rapidity range. The active elements of this device are two planes of silicon strip detectors with

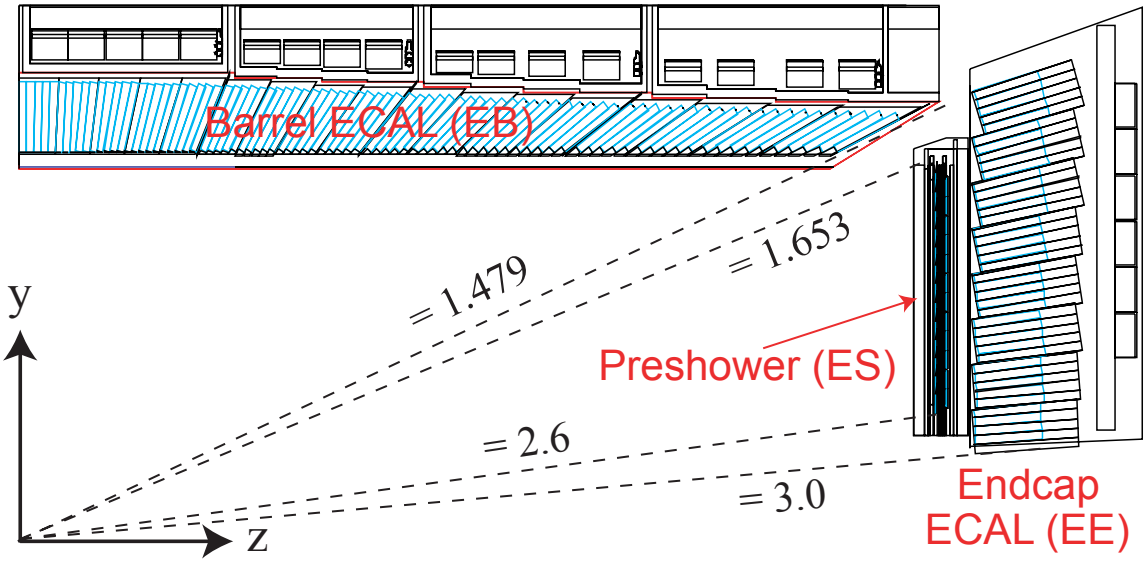


Figure 13: A schematic view of a quarter of the electromagnetic calorimeter of the CMS.

a pitch of 1.9 mm, which lie behind disks of lead absorber at depths of $2X_0$ and $3X_0$, where X_0 is the radiation length.

The ECAL is made of lead tungstate, $PbWO_4$ crystals. There are 61 200 modules mounted in the Barrel and 7 324 in each of the two endcaps. These crystals have a short radiation length because of high density, $X_0 = 0.89$ cm, are fast, 80% of the light is emitted in the 25 ns LHC bunch crossing interval, and radiation hard. A low light yield requires use of photodetectors with intrinsic gain that can operate in the magnetic field.

The radius of the ECAL Barrel is 129 cm and all the crystals point toward the interaction point. The size of the crystals is $22 \times 22 \times 230$ mm^3. The front face cross section is equal to $\Delta\phi = \Delta\eta = 0.0174$ and the length of 23 cm, which is equivalent to $26X_0$.

The Endcaps are located at $|z| = 315$ cm. The crystals are arranged in a rectangular $x - y$ grid. All crystals are identical and have a front face cross section of 28.6×28.6 mm^2 and a length of 220 mm $= 25X_0$. Each endcap is divided into two halves. The crystals in both the EE and the EB are grouped into 5×5 units, called supercrystals.

The Preshower detector is designed to identify neutral pions in the Endcaps within $1.653 < |\eta| < 2.6$. It also supports the correct identification of electrons and improves position resolution of photons and electrons. The Preshower is a sampling calorimeter with two alternating types of layers: lead layers initiate showers from incoming high-energy electrons and photons, while layers of silicon strips measure the deposited energy.

The ES is 20 cm thick and consists of a lead layer of $2X_0$ thickness, followed by the first sensor layer, and another lead layer of thickness $1X_0$ before reaching the second sensor plane. This way, 90% of incident photons produce electron-positron pairs before they reach the second sensor layer.

The ECAL energy resolution is given by

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \tag{2.1}$$

where the first term is the stochastic term, the second term is the noise term, which includes both electronics noise and pile-up energy, and the last term is the constant term C^2. The value of the S and C coefficients are determined by the material used in the ECAL. With the lead tungstate $S = 2.8\%$ and $C = 0.3\%$. The value of N is obtained using the test beams, and

is expected to be of the order of 12%. For example, an electron with $E = 120$ GeV can be measured with a relative energy resolution better than 0.45%.

2.3.4 Hadron Calorimeter

The Hadron Calorimeter (HCAL) is a sampling calorimeter: there are alternating layers of absorbers and scintillators that are used to measure the energy, position and arrival time of the particles. It is used to measure the energy of hadron jets, and plays important role in measuring the missing transverse energy, $\not{E}_T$, which is calculated using the information from the ECAL and the HCAL. The fluorescent scintillator materials produce a rapid light pulse when the particle passes through. Special optic fibers collect this light and feed it into the readout boxes where photodetectors amplify the signal. Figure 12 shows a schematic view of a quarter of the CMS Tracker and Calorimeters.

The hadron barrel (HB) part of the HCAL has 32 slices each with $\Delta\eta = 0.087$ and covers the region $|\eta| < 1.4$. The barrel is divided at $z = 0$ cm into two halves. Each slice consists of 72 sectors with $\Delta\phi = 0.087$. In total, the HB contains 2 304 towers. The absorber layers are 50 mm thick and are made of a composition of 70% Cu and 30% Zn, brass. The radiation length is $X_0 = 1.5$ cm and the nuclear interaction length is $\lambda_I = 16.4$ cm. The total absorber thickness at $\eta = 0$ is $5.8\lambda_I$ and increases to $10.6\lambda_I$ at $|\eta| = 1.3$. The ECAL adds $1.1\lambda_I$ of the material.

The hadron outer (HO) detector is located outside of the magnet coil and inside the Muon Barrel system. It is divided into 5 sectors and covers the region $|\eta| < 1.26$. The tiles are grouped into 30° sectors, matching the ϕ segmentation of the drift tube chambers. They sample the

energy from penetrating hadron showers leaking through the rear of the calorimeters and so serve to catch the tails of the hadron showers, using the solenoid coil as absorber material. The HB has the minimal absorber depth at $\eta = 0$. Therefore, an additional absorber in the form of a 19.5 cm thick iron ring is installed. This way, the minimal absorber thickness is extended to $11.8\lambda_I$ except in the boundary region between the Barrel and the Endcap.

The HCAL Endcaps (HE) are designed to detect the hadronic particles in the region $1.3 < |\eta| < 3$. Each Endcap consists of 13 η towers with 5° ϕ segmentation. For the 5 innermost towers, at smaller η, the ϕ segmentation is 5° and the η segmentation is 0.087. For the 8 outermost towers the ϕ segmentation is 10°, while the η segmentation varies from 0.09 to 0.35 at the highest η. The total number of HE towers is 2 304. In the HE region, the total absorber thickness is $10\lambda_I$, including the ECAL.

The last part of the HCAL is the hadron forward (HF) detector. It covers the range $3.0 < |\eta| < 5.0$ and its front face is located at 11.2 m from the interaction point. The absorber depth is 1.65 m. HF was designed to withstand the large particle fluxes in the forward η regions. The signal is generated by Cerenkov light and it is mostly sensitive to the charged component of the hadron showers. It has a coarser $\eta - \phi$ segmentation of 0.175×0.175 compared to other parts.

The performance of the CMS Hadron Calorimeter with cosmic ray muons and the LHC beam data was measured in 2008, (38). Parts of the various HCAL subsystems were exposed to beams of electrons, pions, protons and muons, to measure their characteristics and to obtain a reference calibration. An ECAL module was also included in the test beam setup. The

hadronic energy resolution of the barrel HCAL and ECAL combination is being parametrized as $\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus b$, where a corresponds to a stochastic term and b to a constant term. One measures $a = 0.847 \pm 0.016 \sqrt{\text{GeV}}$ and $b = 0.074 \pm 0.008$, with E measured in GeV. The energy resolution in the Endcaps is similar to that in the Barrel. The corresponding values for HF are $a = 1.98 \sqrt{\text{GeV}}$ and $b = 0.09$.

2.3.5 Muon System

Figure 14 shows a schematic view of a quarter of the Muon system of the CMS detector. Similar to other subdetectors, the Muon system consists of barrel and endcap systems.

The Barrel detector covers a region $|\eta| < 1.2$ and the Endcap extends up to $|\eta| < 2.4$. There is an overlap region between the Barrel and Endcap at $0.9 < |\eta| < 1.2$.

The Muon system uses three different technologies in sub-detectors: Drift Tubes (DT) in the Barrel region, Cathode Strip Chambers (CSC) in the Endcap, and Resistive Place Chambers (RPC) in both the Barrel and the Endcap. The return yoke absorbs the return flux of the 3.8 T magnetic field.

The muons rate in the Barrel is expected to be low, less than 10 particles per second per square cm. The DT's are the best choice for a precise measurement of the position of the tracks in such conditions. The CSC's are faster and more suitable for the Endcap, where the magnetic field and the rate of particles are higher. The RPC's provide a lower spatial resolution but faster timing signal and have a different sensitivity to backgrounds. The RPC's are placed in the Barrel and the Endcap in order to provide a trigger complementarity.

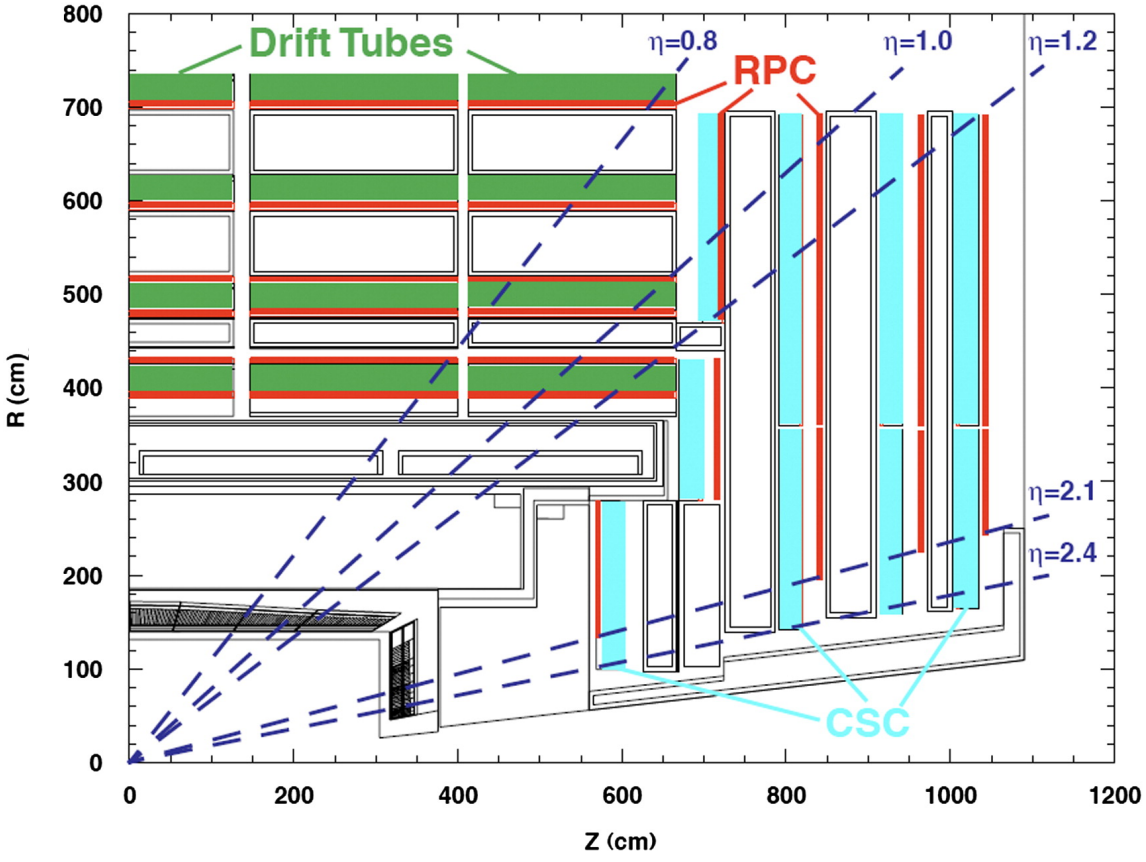


Figure 14: A schematic view of a quarter of the Muon system of the CMS detector.

The performance of the Muon system was studied using data collected in pp collisions at $\sqrt{s} = 7$ TeV at the LHC in 2010 (39). The relative p_T resolution is between 1.3% to 2.0% for muons in the Barrel and better than 6% in the Endcaps, and was measured using the momentum measurements provided by the Tracker. The transverse momentum resolution of tracks reconstructed using information from only the Muon detectors is better than 10% in the Barrel region up to 1 TeV.

2.3.6 Trigger

The purpose of the Trigger system is to reduce the event rate to a manageable level of a few hundred Hz. Because of the high bunch crossing frequency of 40 MHz, the event filtering has to be done at a very early stage in order to reduce the data rate from the detector. Therefore, the first trigger level (L1 Trigger) is implemented in hardware. Its diagram is shown in Figure 15.

The L1 Trigger has only very limited capabilities of combining information of different elements. A second trigger step is introduced, the High-Level Trigger (HLT), which is implemented in software and runs on usual computer hardware. It uses similar or same algorithms as the ones that are used for the full event reconstruction.

Figure 15 shows the CMS L1 Trigger diagram. In this figure, the very top elements represent the L1 Trigger inputs which are based on the energy deposits in the calorimeter trigger towers and hit patterns in the muon chambers. These trigger primitives are propagated to the regional triggers, which combine the information of several trigger primitives into calorimeter and muon objects. These are passed to the global muon and calorimeter triggers, which rank the muon and calorimeter objects and pass the high-rated objects to the Global Trigger. The latter element

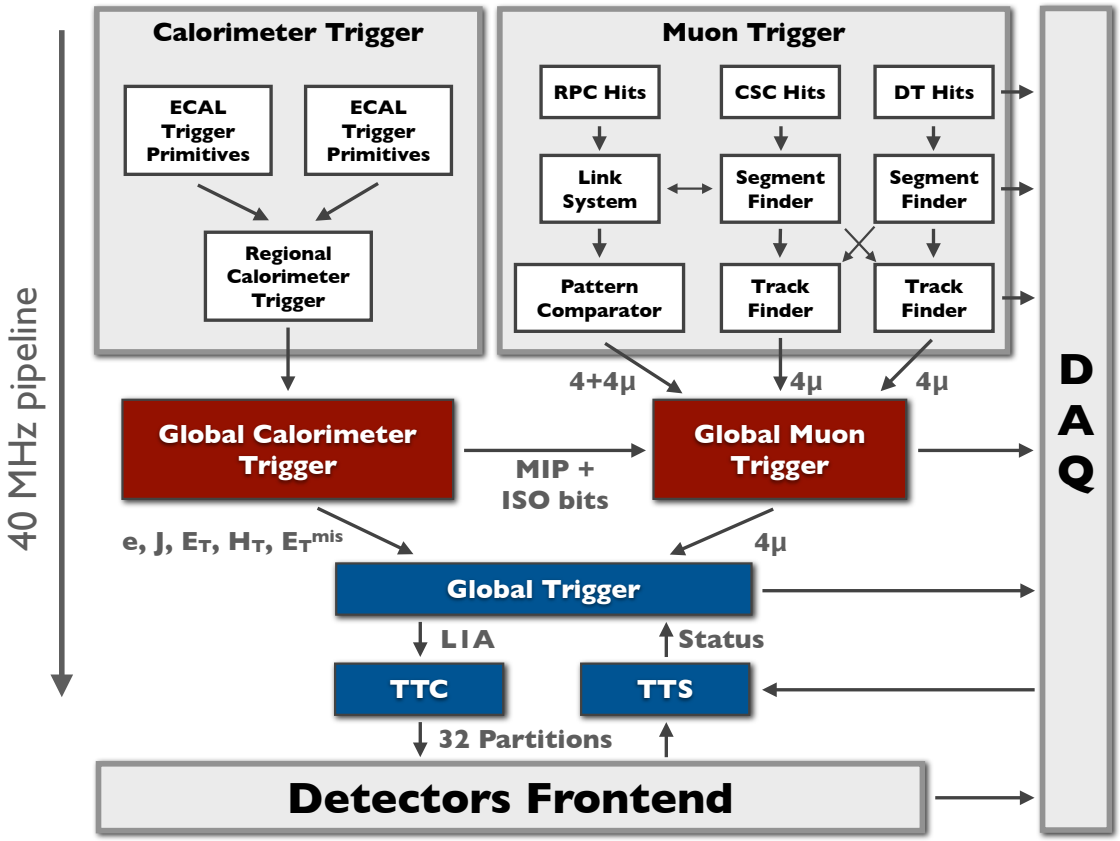


Figure 15: A schematic view of the CMS L1 Trigger.

decides whether to read out and pass the event to the High-Level Trigger or to reject the event. The L1 Trigger output rate is below 50 kHz, while the HLT output rate is kept below 150 Hz.

2.3.7 Computing and Software

CMS adopts a distributed network of computing centers that are used to reconstruct and analyze the data. These centers are distributed around the world and are organized in three layers: Tier-0, Tier-1, and Tier-2.

Every event is replicated in different data format, which defines the role of the specific Tier layer in the data processing chain. The RAW data format saves the low level detector information such as individual pixel hits, signals. The reconstructed (RECO) format contains all the reconstructed objects. It is too large for frequent transfers and contains much more information than most analyses will need. The analysis object data (AOD) format is a subset of the RECO and holds only the information that is sufficient for most analyses.

Recently, a new plan was introduced to add one more level, called Physics Analysis Toolkit (PAT). It uses data from AOD as input, does basic data processing that is common to most analyses, and saves physics objects in its own format.

The data formats and the reconstruction algorithms are implemented in the CMS software framework (`CMSSW`). The framework is written with a modular approach and is based on the Event Data Model (EDM), which treats every event independently.

The Tier-0 center is the first level of data processing and is located at CERN. It accepts RAW data from the detector and does a first calibration and reconstruction. The events are divided into primary datasets based on the trigger information and are archived to the Tier-0

storage. Data are then distributed to at least one Tier-1 center. This way each event is saved two or more times. There are currently eight Tier-1 centers in the world. These centers also provide extensive CPU power for re-reconstructions, calibration, AOD extraction and other common processing steps. The Tier-1 centers redistribute further the events to smaller Tier-2 centers.

Tier-2 centers provide CPU and storage resources for individual analysis, including the one described in this thesis.

CHAPTER 3

OBJECT IDENTIFICATION

Proton-proton collisions produce a large number of particles that interact with CMS sub-detectors by producing electronic signals and pulses. These raw signals are then analyzed to estimate particle's energy, momentum, and trajectory. This process is called reconstruction and is done for all physics objects such as electrons, jets, muons, and missing transverse energy. The reconstruction algorithms used in this analysis are described below.

3.1 The Particle Flow Reconstruction

The Particle Flow (PF) (40) is an event reconstruction technique that aims at the reconstruction and identification of all stable particles in the event: electrons, muons, photons, charged hadrons and neutral hadrons by combining the information from all CMS sub-detectors for an optimal determination of their direction, energy and type. The reconstruction uses the list of individual particles, as if it came from a Monte-Carlo (MC) event generator, to build jets, to determine the missing transverse energy $\not\!\!E_T$, to reconstruct and identify tau leptons from their decay products, to quantify charged lepton isolation with respect to other particles, to tag b jets, and more. The missing transverse energy gives an estimate of the direction and energy of the neutrinos and other particles that are invisible by the detector.

The CMS detector is ideal for the particle flow event reconstruction. With large silicon Tracker immersed in a uniform axial magnetic field, charged particle tracks can be reconstructed

with large efficiency and small fake rate down to a transverse momentum p_T of 150 MeV and for pseudo-rapidities as large as ± 2.6.

Photons are reconstructed with an excellent energy resolution by an essentially hermetic ECAL surrounding the Tracker. Due to the CMS ECAL granularity the photons can be separated from charged-particle energy deposits even in jets with p_T of several hundreds of GeV. Charged and neutral hadrons deposit their energy in the HCAL, installed inside the coil and surrounding the ECAL, with a similar pseudo-rapidity coverage. The granularity of the HCAL is 25 times coarser than that of the ECAL, which would not allow charged and neutral hadrons to be spatially separated in jets with a transverse momentum much above 100 GeV. The hadron energy resolution in the combined ECAL-HCAL system is of the order of 10% at 100 GeV. Electrons are reconstructed by a combination of a track and a several energy deposits in the ECAL, from the electron itself and from possible Bremsstrahlung photons radiated by the electron in the Tracker material on its way to the ECAL. Muons are reconstructed and identified, in isolation as well as in jets, with very large efficiency and purity from a combination of the Tracker and muon chamber information. The presence of neutrinos and other weakly interacting particles can be detected by transverse missing energy, defined as the modulus of the vector sum of the transverse momenta of all reconstructed particles $\not{E}_T = |\sum_i \mathbf{p}_{T,i}|$.

Most stable particles produced in proton-proton collisions have a rather low p_T, even in collisions involving a large momentum transfer. For example, the average p_T carried by the stable constituent particles is of the order of a few GeV in jets with a total p_T below 100 GeV, which is typical of the decay chains of heavy exotic particles. It is therefore essential to accurately

reconstruct and identify as many of the final stable particles as possible, even with small p_T's and energies, to disentangle the production of the exotic particles from the dominating SM background processes.

The particle reconstruction and identification is performed with a combination of the information from each CMS sub-detector, under the form of charged-particle tracks, calorimeter clusters, and muon tracks. These building elements of the particle-flow event reconstruction must be delivered with a high efficiency and a low fake rate, even in high-density environments.

3.1.1 Iterative Tracking

The momentum of charged hadrons is measured in the Tracker with a high resolution than that of the calorimeters for p_T up to several hundreds of GeV. The Tracker provides a precise measurement of the charged-particle direction at the production vertex. About two thirds of the jet energy is on average carried by charged particles, and the Tracker is fundamental for the particle-flow event reconstruction.

Each charged hadron missed by the tracking algorithm would be only detected by the calorimeters. For this reason the tracking efficiency must be as close to 100% as possible. At the same time the tracking fake rate must be kept small because of fake tracks, with a randomly distributed momentum, would lead to potentially large energy excesses.

An iterative-tracking strategy (41) has been developed to achieve both high efficiency and low fake rate. First, tracks are seeded and reconstructed with very tight criteria, leading to a moderate tracking efficiency, but a negligibly small fake rate. The next steps proceed by removing hits unambiguously assigned to the tracks found in the previous iteration, and by

progressively loosening track seeding criteria. The softer seeding criteria increase the tracking efficiency, while the hit removal allows to keep the fake rate low due to the reduced combinatorics. In the first three iterations, tracks originating form within a thin cylinder around the beam axis are found with an efficiency of 99.5% for isolated muons in the Tracker acceptance, and larger than 90% for charged hadrons in jets.

The fourth and fifth iterations have relaxed constraints on the origin vertex, which allows the reconstruction of secondary charged particles originating from photon conversions and nuclear interactions in the Tracker material and from the decay of long-lived particles such as K_0 or Λ. With this iterative technique, charged particles with as little as three hits, a p_T as small as 150 MeV and an origin vertex more than 50 cm away from the beam axis, are reconstructed with a fake rate at the per cent level.

3.1.2 Calorimeter Clustering

The purpose of a clustering algorithm in the calorimeters is to:

- detect and measure the energy and direction of stable neutral particles such as photons and neutral hadrons,
- separate these neutral particles from energy deposits from charged hadrons,
- reconstruct and identify electrons and all accompanying Bremsstrahlung photons,
- help the energy measurement of charged hadrons for which the track parameters were not determined accurately, which is the case for low-quality, or high-p_T tracks.

A specific clustering algorithm has been developed for the particle-flow event reconstruction, with the aim of a high detection efficiency even for low-energy particles, and towards a separation of close energy deposits, as illustrated in Figure 16. The clustering is performed separately in each sub-detector: ECAL barrel, ECAL endcap, HCAL barrel, HCAL endcap, Pre-Shower first layer and Pre-Shower second layer.

(a) (b)

Figure 16: The particles are seen in the detector as tracks and energy deposits (a). The particle-flow algorithm links all the information coming from the different sub-detectors to reconstruct and identify all the particles coming from the pp collisions (b).

The algorithm consists of three steps:

1. Identify cluster seeds as local calorimeter-cell energy maxima above a given energy.

2. Grow topological clusters from the seeds by aggregating cells with at least one side in common with a cell already in the cluster, and with an energy in excess of a given threshold. These thresholds are at two standard deviations of the electronics noise in the ECAL and amount to 800 MeV in the HCAL. A topological cluster gives rise to as many particle-flow clusters as seeds.

3. The calorimeter granularity is exploited by sharing the energy of each cell among all particle-flow clusters to the cell-cluster distance, with and iterative determination of the cluster energies and positions

3.1.3 The Link Algorithm

In general, a given particle is expected to give rise to several particle-flow elements in the various CMS sub-detectors: one charged-particle track, several calorimeter clusters, or muon tracks. Most of the possible cases are shown in Figure 16.

These elements must be connected to each other by a link algorithm to fully reconstruct each single particle, while getting rid of any possible double counting from different detectors. The link algorithm is performed for each pair of elements in the event and defines a distance between any two linked elements to quantify the quality of the link. The algorithm then produces blocks of elements linked directly or indirectly. Because of the granularity of the CMS detector, blocks typically contain only one, two or three elements, and constitute simple inputs for the particle reconstruction and identification algorithm. The smallness of the blocks ensures the algorithm performance to be essentially independent of the event complexity. For example, jets made of

one to four blocks, turn out to have the same energy response and resolution as regular jets found in QCD events, which typically feature a larger number of blocks, but of the same size.

A link between a charged-particle track and a calorimeter cluster proceeds as follows. The track is first extrapolated from its last measured hit in the Tracker to the two layers of the PS, then to the ECAL, at a depth corresponding to the expected maximum of a typical longitudinal electron shower profile. After the ECAL the track is extrapolated to the HCAL, at a depth corresponding to one interaction length, typical of a hadron shower. The track is linked to any given cluster if the extrapolated position in the corresponding calorimeter is within the cluster boundaries. This cluster envelope can be enlarged by up to the size of a cell in each direction, to account for the presence of gaps between calorimeter cells, cracks between calorimeter modules, for the uncertainty on the position of the shower maximum and for the effect of multiple scattering for low-momentum charged particles. The link distance is defined as the distance in the (η, ϕ) plane between the extrapolated track position and the cluster position.

In order to collect the energy of all Bremsstrahlung photons emitted by electrons, tangents to the tracks are extrapolated to the ECAL from the intersection points between the track and each of the Tracker layers. A cluster is linked to the track as a potential Bremsstrahlung photon if the extrapolated tangent position is within the boundaries of the cluster.

Similarly, a link between two calorimeter clusters, either between an HCAL and an ECAL cluster or between an ECAL and a PS cluster, is established when the cluster position in the more granular calorimeter is within the cluster envelope in the less granular calorimeter. This

envelope can be slightly enlarged as is done for the track-cluster link. The link distance is also defined in the (η, ϕ) plane as the distance between the two cluster positions.

Finally, a link between a charged-particle track in the Tracker and a muon track in the muon system is established, global muon, when a global fit between the two tracks returns an acceptable value of $\chi 2$. Only the global muon that returns the smallest $\chi 2$ is retained if several global muons can be fit with a given muon track and several Tracker tracks. In this case the $\chi 2$ defines the link distance.

3.1.4 Particle Reconstruction and Identification

The reconstruction and identification of a set of particles from each block of elements is finally performed by the particle-flow algorithm. The resulting list of reconstructed particles constitutes a global description of each event, available for subsequent physics analysis.

For each block, the algorithm proceeds as follows. First, each global muon gives rise to a particle-flow muon, if its combined momentum is compatible with that determined from the sole Tracker within three standard deviations. The corresponding track is removed from the block. An estimate of the energy deposited in the HCAL and ECAL, used at a later stage in the algorithm, was measured with cosmic rays to be 3 (0.5) GeV, with an uncertainty of $\pm 100\%$.

Electron reconstruction and identification follows. Each track of the block is submitted to a pre-identification stage which exploits the Tracker as a pre-shower: electrons tend to give rise to short tracks, and to lose energy by Bremsstrahlung in the Tracker layers on their way to the calorimeter. Pre-identified electron tracks are refit with a Gaussian-Sum Filter (42) in an attempt to follow their trajectories all the way to the ECAL. A final identification is performed

with a combination of a number of tracking and calorimetric variables. Each identified electron gives rise to a particle-flow electron. The corresponding track and ECAL clusters (including all ECAL clusters identified as Bremsstrahlung photons) are removed from further processing of the block.

Tighter quality criteria are applied to the remaining tracks: it is required that the relative uncertainty on the measured p_T be smaller than the relative calorimetric energy resolution expected for charged hadrons. In hadronic jets 0.2% of the tracks are rejected by this requirement. While about 90% of those are fake tracks, the energy of the 10% tracks originating from real particles is not lost for the particle-flow reconstruction, as it is measured independently with more precision, in this case, by the calorimeters. The remaining elements may give rise to charged hadrons, photons or neutral hadrons, and more rarely to additional muons.

A track can be directly connected to a number of ECAL and HCAL clusters. The detection of the neutral particles in the block (photons and neutral hadrons) involves a comparison between the momentum of the tracks and the energy detected in the calorimeters. Several tracks can be linked to the same HCAL cluster. In this case the sum of their momenta is compared to the calibrated calorimetric energy. On the other hand, if a track is linked to several HCAL clusters, only the link to the closes cluster is kept for the comparison. A track can also be linked to more one ECAL clusters, and the link to the closest cluster is kept too.

In rare cases, the total calibrated calorimetric energy is still smaller than the total track momentum by a large amount. When the difference is larger than three standard deviations, a relaxed search for muons and for fake tracks is performed. First, all global muons, not already

selected by the algorithm and for which an estimate of the momentum exists with a precision better than 25%, are treated as particle-flow muons. The redundancy of the measurements in the Tracker and the calorimeters thus allows a few more muons to be found without increasing the fake-muon rate. This redundancy is further exploited by progressively removing tracks from the block, ordered according to their measured p_T uncertainty. The process stops either when all tracks with a p_T uncertainty in excess of 1 GeV have been examined, or when the removal of a track would render the total track momentum smaller than the calibrated calorimetric energy. Less than 0.3 per mil of the tracks are concerned by this procedure.

Each of the remaining tracks in the block gives rise to a particle-flow charged hadron, the momentum and energy of which are taken directly from the track momentum, under the charged pion mass hypothesis. If the calibrated calorimetric energy is compatible with the track momentum within measurements uncertainties, the charged-hadron momenta are redefined by a fit of the measurements in the Tracker and the calorimeters, which reduces to a weighted average if only one track is present. This combination is relevant at very high energies and/or large pseudo-rapidities, for which the track parameters are measured with degraded resolutions.

It may well be that the calibrated energy of the closest ECAL and HCAL clusters linked to the track(s) be significantly larger than the total associated charged-particle momentum. If the relative energy excess is found to be larger than the expected calorimeter energy resolution, it gives rise to a particle-flow photon, and possible to a particle-flow neutral hadron. Specifically, if the excess is large than the total ECAL energy, a photon is created with this ECAL energy and a neutral-hadron is created with the remaining part of the excess. It is observed that in

jets 25% of the jet energy is carried by photons, and neutral hadrons leave only 3% of the jet energy in the ECAL. Therefore, the precedence is given to photons in the ECAL over neutral hadrons.

The remaining ECAL and HCAL clusters, either originally not linked to any track or for which the link was disabled, give rise to particle-flow photons and particle-flow neutral hadrons, respectively.

3.2 Jet Reconstruction

There are four types of jet reconstruction techniques at CMS, which differently combine individual contributions from subdetectors to form the input to the jet clustering algorithm: Calorimeter jets, Jet Plus Track (JPT) jets, Particle Flow jets, and Track jets. The PF jets have the best resolution and for this reason are used for the full reconstruction and offline analysis.

Jets are reconstructed using the $anti-k_T$ (43) clustering algorithm with the size parameter R = 0.5. The algorithm defines a distance d_{ij} between entities (particles, pseudo-jets) i and j and d_{iB} between entity i and the beam B. The clustering proceeds by identifying the smallest of the distances. If it is a d_{ij}, then entities i and j are recombined, while if it is d_{iB} then the entity i is called a jet and removed from the list of entities. The distances are recalculated every time the entities are merged or new jet is found, and the procedure is repeated until no entities are left.

The extension relative to the k_T (44) and Cambridge/Aachen (45) algorithms lies in the definition of the distance:

$$d_{ij} = \min\left(k_{Ti}^{2p}, k_{Tj}^{2p}\right)\frac{\Delta_{ij}^2}{R^2} \qquad \text{and} \qquad d_{iB} = k_{Ti}^{2p} \tag{3.1}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and k_{Ti}, y_i, and ϕ_i are the transverse momentum, rapidity and azimuth angle of particle i, respectively. In addition to the radius parameter R, the algorithm has a parameter p to manage the relative power of the energy versus geometrical scales ij. For $p = 1$ one recovers the inclusive k_T algorithm. The case $p = 0$ corresponds to the inclusive Cambridge/Aachen algorithm, and $anti - k_T$ jets clustering algorithm is defined by $p = -1$.

The performance of the $anti - k_T$ algorithm can be demonstrated with a simple example. Consider an event with a few well separated high k_T and many low k_T particles. Suppose, that there is a low k_T particle i, that is separated by ΔR from another low k_T particle j . Assume that there is another high k_T particle k which is separated from particle i by the same distance ΔR. According to the definition of d_{ij}, we have $d_{ij} > d_{ik}$. Therefore the low k_T particle i will be recombined with the high k_T particle k. This way low-k_T particles will tend to cluster with high-k_T ones long before they cluster among themselves. This way low-k_T particles do not modify the shape of the jet.

It has been verified with simulated events that the algorithm gives rise to a more regular shape of the reconstructed jets with respect to other algorithms, as can be seen in Figure 17. Due to this fact, the $anti - k_T$ algorithm is more robust than the other ones with respect to

non-perturbative effects like hadronization and underlying event contamination, improves the momentum resolution and calorimeter performance. The $anti - k_T$ clusterization is also faster than the other algorithms.

All particles reconstructed with the particle-flow algorithm, without distinction of type and without any energy threshold, are clustered into particle-flow jets.

3.3 Identification of b jets

The top quark decay products include b jets that arise from bottom quark hadronization and decay. It becomes important to accurately identify b jets in order to reduce the overwhelming background in the analysis, which mainly comes from processes that involve jets from gluons and u, d, s quark fragmentation, and from c quark hadronization.

The main property of the b quark, that is used in the identification, is the long lifetime. Typically it travels several millimeters before decaying.

There are several b-tagging algorithms in CMS. These include Track Counting (TC), Jet Probability (JP), Simple Secondary Vertex (SSV), and Combined Secondary Vertices (CSV). Each produces a numerical discriminator as output, which can be used to select jets according to operating point that is best for the analysis. Lifetime b-tagging algorithms rely on the presence of a reconstructed secondary vertex shown in Figure 18, or tracks with large impact parameter significance, $\frac{IP}{\sigma_{IP}}$, where IP is the impact parameter of the track with respect to the primary vertex, i.e. the distance between the track and the vertex at the point of closest approach,and σ_{IP} is its error.

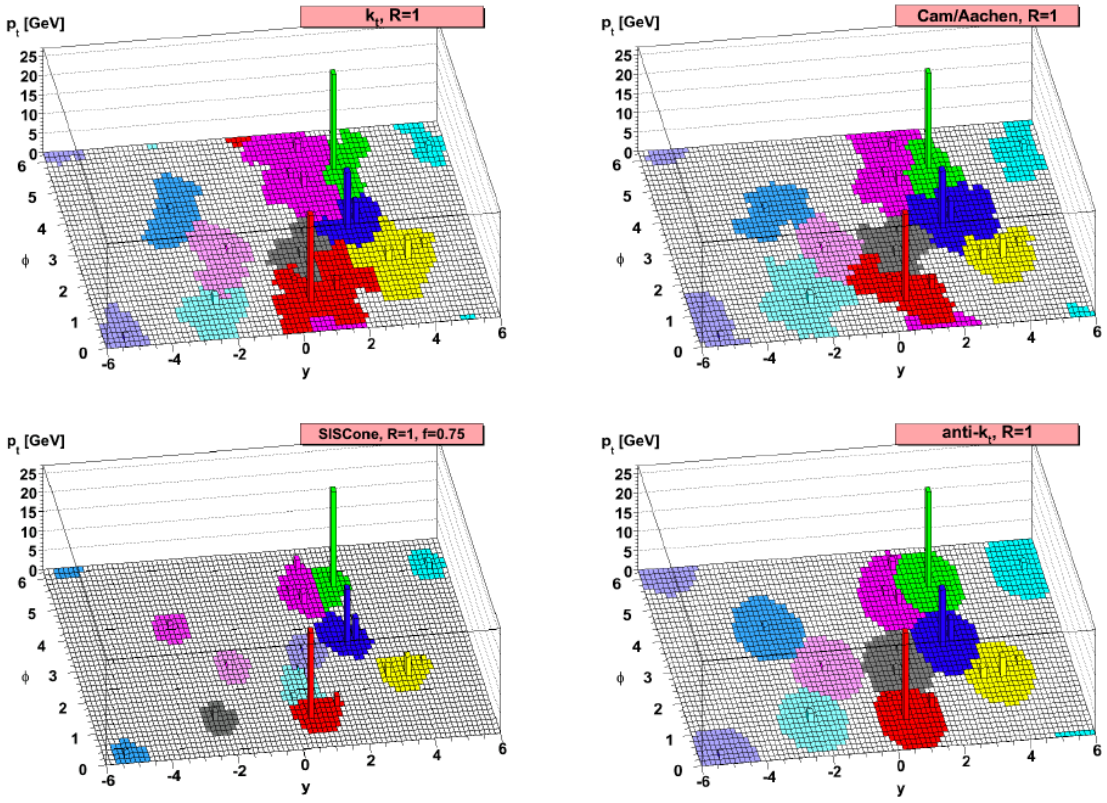


Figure 17: A simulated event of a few high energetic jets together with many low k_T particles. The jets are clustered with four different algorithms. The $anti - k_T$ jets have a more regular shape.

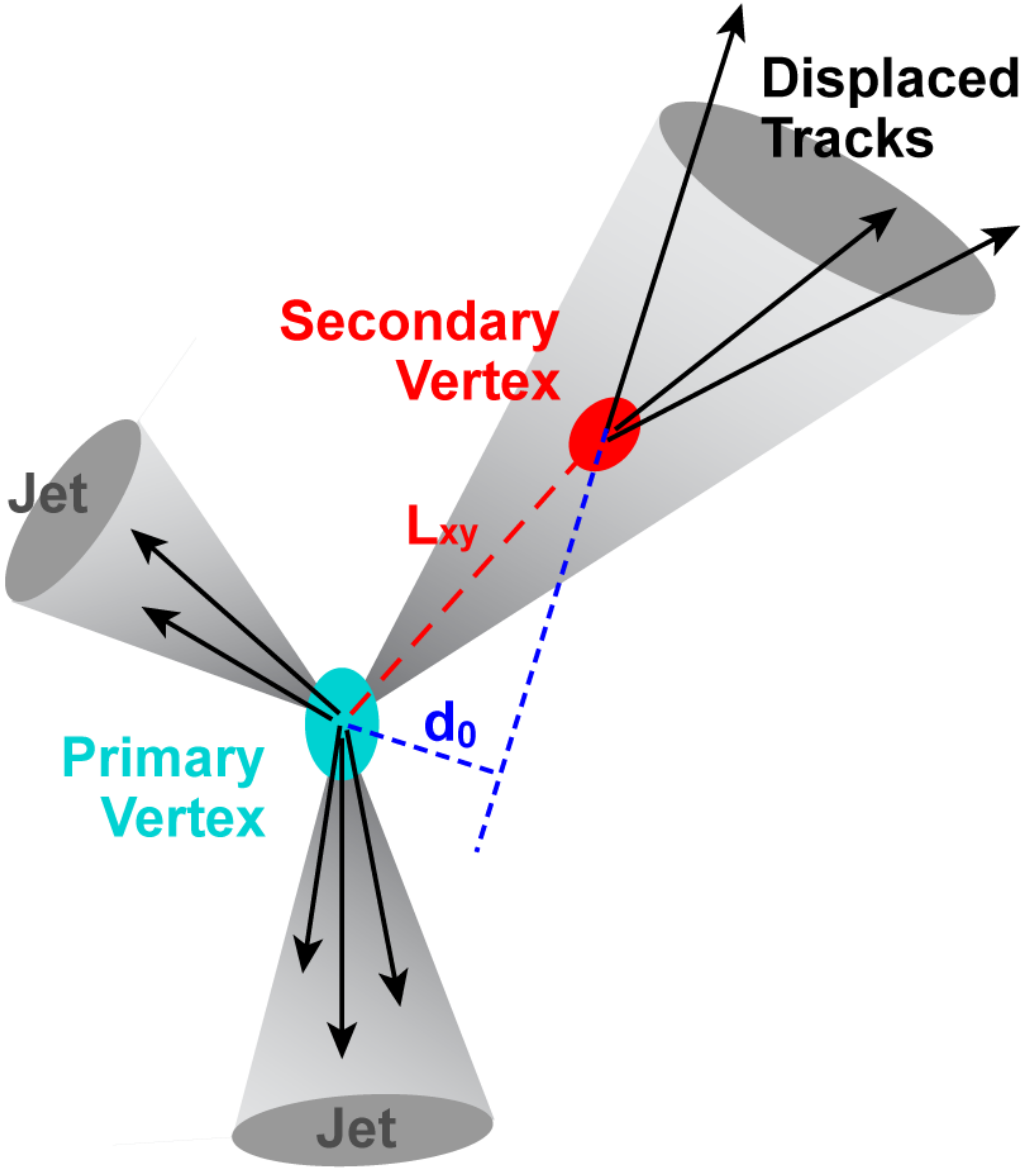


Figure 18: A schematic representation of the impact parameter of a track with respect to the primary vertex.

Impact parameters can be signed as positive/negative if the associated tracks are produced downstream/upstream with respect to the primary interaction vertex. The reconstructed decay length between the primary and secondary vertices can be signed in a similar way.

Due to the experimental resolution, the distributions of the impact parameter and decay length are symmetric around zero for genuine tracks from the primary interaction vertex, as in case of light flavor jets.

In the Track Counting b-tagging algorithm, a jet is b-tagged if there are at least N tracks with the impact parameter significance larger than a given threshold. The tracks are ordered by significance in descending order in TC tagger, and the N^{th} track significance is used as discriminator. The algorithm defines two modes: high efficiency when second track significance is used as discriminator, and high purity with discriminator given by the third track significance.

The Jet Probability tagger computes how likely it is for a set of tracks to have originated from the primary vertex, based on the impact parameter significance of each track in a jet.

The Simple Secondary Vertex algorithm is based on the reconstruction of at least one secondary vertex. The discriminator is computed from the three dimensional decay length $L3D = \log\left(1 + \frac{|L_{3D}|}{\sigma_{L_{3D}}}\right)$, which is signed as the L_{3D} value. The tagger has two operating modes: high efficiency (SSVHE) with two tracks associated to the secondary vertex, and high purity (SSVHP) with three tracks associated to the secondary vertex.

We used the Combined Secondary Vertex tagger, which is similar to SSV. A discriminating variable is built by combining different topological and kinematic variables with the track impact parameter significances. The algorithm first reconstructs secondary vertices in an inclusive way

inside the jet using the Trimmed Kalman Vertex Finder (46). This algorithm begins by using all tracks in the jet and subsequently rejects outliers, which then are used to reconstruct additional vertices. The algorithm requires the presence of a secondary vertex and uses topological and kinematic variables related to the vertex. Therefore, it is desirable to reconstruct as completely as possible the decay vertex in order to increase the discriminating power of these topological and kinematic variables.

Each vertex candidate has to fulfill the following requirements:

- the distance between the primary vertex and the secondary vertex in the transverse plane has to exceed 100 μm and be below 2.5 cm,
- the distance between the primary vertex and the secondary vertex in the transverse plane divided by its error has to be $\frac{L_t}{\sigma_{L_t}} > 3$,
- the invariant mass of the charged particles associated to the vertex must not exceed 6.5 GeV,
- the vertex must not be compatible with a K^0_S decay: vertices with two oppositely charged tracks are rejected if their mass is within a window of 50 MeV around the nominal K^0_S mass.

The CSV algorithm defines three categories after the secondary vertex reconstruction and selection steps:

- in the Reco Vertex category at least one secondary vertex candidate must be reconstructed which satisfies the selection criteria,

- if no reconstructed secondary vertex is found, then a Pseudo Vertex is created using charged particle tracks that are not compatible with the primary vertex,
- if neither of the first two categories are fulfilled then No Vertex category is used.

Depending on the category different variables are used to build up the discriminator.

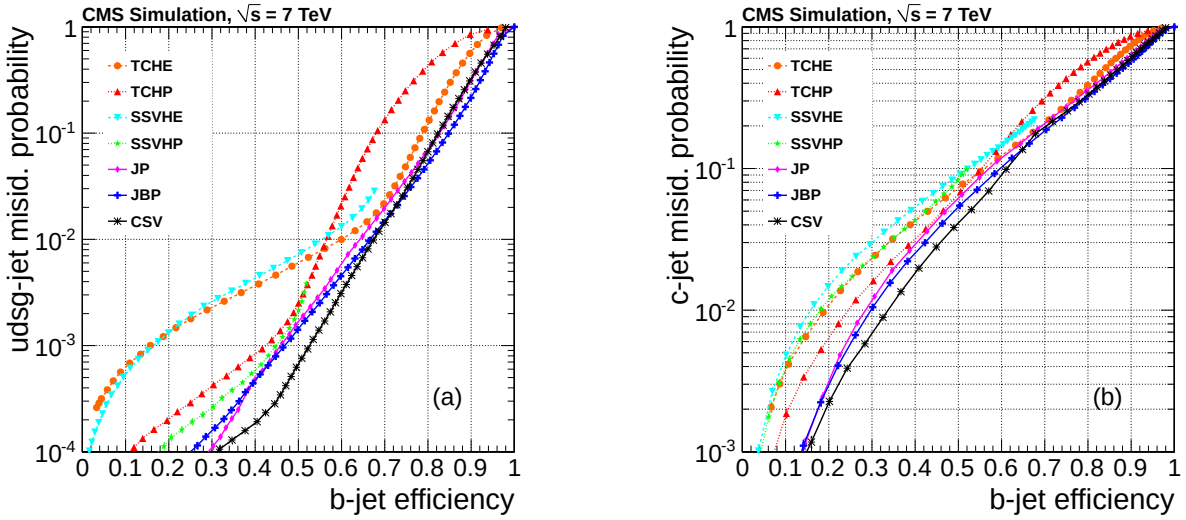
The largest discriminating power is found in the Reco Vertex category. The following variables are used:

- the invariant mass of charged particles associated to the secondary vertex: for the secondary vertices in b-jets, the vertex mass can be significantly above the mass of charm hadrons,
- the multiplicity of charged particles associated to the secondary vertex: b-hadron decays show a significantly larger track multiplicity than charm hadron decays,
- the flight distance, distance between the primary vertex and the secondary vertex in the transverse plane divided by its error,
- the energy of charged particles associated to the secondary vertex divided by the energy of all charged particles associated to the jet: this quantity is sensitive to the hard fragmentation function of b and c quarks,
- the rapidities of charged particle tracks associated to the secondary vertex with respect to the jet direction,
- the track impact parameter significance of the first track that exceeds the charm threshold.

The charm threshold is used to improve the rejection of charm quark jets. The tracks are sorted in decreasing order by impact parameter significance and invariant mass of tracks 1 to n is computed. If this mass exceeds 1.5 GeV, the value is related to the mass of charm hadrons, the impact parameter significance of the track moving the n^{th} track mass above the threshold can be added to the discriminator. The threshold value is lower than typical charm hadron masses because of missing neutral or undetected charged particles.

The Pseudo Vertex category uses the same set of variables except the significance of flight distance. The No Vertex category does not add any additional variables to the discriminator.

The performance of described algorithms is summarized in Figure 19 where the predictions of the simulation for the misidentification probabilities (the efficiencies to tag non-b jets) are shown as a function of the b-jet efficiencies (47). Jets with $p_T > 60$ GeV in a sample of simulated multijet events are used to obtain the efficiencies and misidentification probabilities. For loose selections with 10% misidentification probability for light-parton jets a b-jet tagging efficiency of $80 - 85\%$ is achieved. For tight selections with misidentification probabilities of 0.1% the typical b-jet tagging efficiency values are $45 - 55\%$. For medium and tight selections the CSV algorithm shows the best performance. Because of the non-negligible lifetime of c hadrons the separation of c from b jets is naturally more challenging. Due to the explicit tuning of the CSV algorithm for light-parton- and c-jet rejection it provides the best c-jet rejection values in the high-purity region.



(a) (b)

Figure 19: Performance curves obtained from simulation for different algorithms. (a) light-parton- and (b) c-jet misidentification probabilities as a function of the b-jet efficiency. Jets with $p_T > 60$ GeV in a sample of simulated multijet events are used to obtain the efficiency and misidentification probability values.

CHAPTER 4

DATA AND SIMULATED SAMPLES

4.1 Data

The BSM resonances search in the electron plus jets final state uses the full data sample of proton-proton collisions at $\sqrt{s} = 7$ TeV collected with the CMS detector between March and October 2011. The CMS 2011 data taking was stopped for a short accelerator maintenance shutdown in August 2011. The data taking period before the shutdown is referred to as 2011A, while the period after the shutdown is known as 2011B. No major changes to the detector have been made in 2011 and the data sample is referred to 2011 data.

We used samples that are reconstructed with standard `CMSSW` v4.2.x software and stored in the Single Electron primary dataset, in AOD format. These samples are summarized in Table I.

All the events in the single electron primary dataset must pass standard CMS quality checks in order to remove those events that are affected by subdetector problems. For this, we applied Golden JSON filter, to keep only good luminosity sections.

After the data quality checks, the events must pass the High Level Trigger requirements. We applied different triggers to different primary dataset samples as summarized in Table II.

We chose triggers so that they do not overlap in data taking periods. The data was taken with HLT with electron p_T cut at 45 GeV during the first data taking period. Later, the

Dataset Name	**Run Range**	**L pb^{-1}**
/SingleElectron/Run2011A-May10ReReco-v1/AOD	160431 - 163869	221.7
/SingleElectron/Run2011A-PromptReco-v4/AOD	165088 - 167913	826.1
/SingleElectron/Run2011A-05Aug2011-v1/AOD	170722 - 172619	396.1
/SingleElectron/Run2011A-PromptReco-v6/AOD	172620 - 173692	706.4
/SingleElectron/Run2011B-PromptReco-v1/AOD	175860 - 180252	2251.0
total		**4401.3**

TABLE I: 2011 data samples used in the analysis. The integrated luminosity takes into account the HLT active periods.

HIGH LEVEL TRIGGER	**Comment**
`HLT_Ele45_CaloIdVT_TrkIdT`	*before run number 165970*
`HLT_Ele65_CaloIdVT_TrkIdT`	*electron p_T is increased to keep HLT rate low*

TABLE II: The High Level Triggers that we applied to 2011 data.

increase in instantaneous luminosity led to higher value of the p_T threshold, 65 GeV. Because of increased threshold the HLT rate remains at acceptable level.

Our analysis covers the full 2011 data sample corresponding to 4.4 fb^{-1} of integrated luminosity of good quality data.

4.2 Simulated Samples

In our analysis, we used simulated events, generated at the center of mass energy $\sqrt{s} = 7$ TeV using the `CTEQ6L` (48) Parton Distribution Function (PDF) parametrization of the proton. The signal and background processes are generated with `MadGraph` (49) or `POWHEG` (50) event generators and interfaced to `PYTHIA` (51) for parton showering. The QCD multijet background samples are fully simulated with `PYTHIA`. The CMS detector simulation and further event reconstructions is done with `CMSSW` v4.2.x software based on `GEANT` version 4 (52; 53). Table III summarizes the signal Monte-Carlo samples, and Table IV lists Standard Model backgrounds. More details about the generation of the samples is given next.

The SM $t\bar{t}$ background is generated with `MadGraph`, assuming a top-quark mass of 172.5 GeV, using `PYTHIA` for the parton showering. Spin correlation in the top decays is taken into account and higher order gluon and quark production is described by the Matrix Elements (ME) with up to three extra partons beyond the top pair system, matched to the parton shower using the MLM algorithm (54). The chosen threshold for the matching is 40 GeV, which ensures a smooth transition from the matrix element to the parton showering description. An additional $t\bar{t}$ sample is generated using `POWHEG` to provide a cross-check and to estimate systematic uncertainties in the modeling. The inclusive $t\bar{t}$ cross section value of 157.5 pb is used for normalization (55; 56).

Process	$\mathbf{N_{events}}$ **(x10^3)**	**Event Generator**
Z' signal with $\Gamma_{Z'}/M_{Z'} = 1.2\%$		MadGraph
$M_{Z'} = 1$ TeV	208	
$M_{Z'} = 1.5$ TeV	168	
$M_{Z'} = 2$ TeV	179	
$M_{Z'} = 3$ TeV	195	
Z' signal with $\Gamma_{Z'}/M_{Z'} = 10\%$		MadGraph
$M_{Z'} = 1$ TeV	229	
$M_{Z'} = 1.5$ TeV	195	
$M_{Z'} = 2$ TeV	209	
$M_{Z'} = 3$ TeV	221	
Kaluza-Klein Gluon signal		PYTHIA
$M_{g_{KK}} = 1$ TeV	100	
$M_{g_{KK}} = 1.5$ TeV	100	
$M_{g_{KK}} = 2$ TeV	100	
$M_{g_{KK}} = 2.5$ TeV	100	
$M_{g_{KK}} = 3$ TeV	100	

TABLE III: Monte-Carlo Z' boson and Kaluza-Klein Gluon samples. The Z' samples are generates with two mass widths: $\Gamma_{Z'}/M_{Z'} = 1.2\%$ and $\Gamma_{Z'}/M_{Z'} = 10\%$.

Process	σ **x BR [pb]**	$\mathbf{N_{event}}$ **(x10^3)**	**Event Generator**
			MadGraph
$t\bar{t}$	163	59514	
$Z/\gamma^* \to l^+l^-$	3048	36278	
$W \to l\nu$	31314	77106	
			PYTHIA
EM-enriched QCD multijet	6271310	114272	
BCtoE QCD multijet	292575	5194	
			POWHEG
Single Top, s channel, lepton decay	4.6	398	
Single Top, t channel, lepton decay	64.6	5845	
Single Top, tW channel, lepton decay	15.7	1624	

TABLE IV: Monte-Carlo SM background samples. Theoretical predictions for the cross sections are calculated at the Next-to-Next-to-Leading Order. Each QCD multijet sample, EM-enriched and BCtoE, is split into three subsamples in $\hat{p}_T$ regions: 20 to 30 GeV, 30 to 80 GeV, and 80 to 170 GeV. Every single top channel includes charge conjugate.

The W boson and Drell-Yan production in association with up to four jets are simulated with the same combination of MADGRAPH and PYTHIA as the $t\bar{t}$ sample. The matching threshold was set to 20 GeV. The Next-to-Next-to-Leading Order (NNLO) production cross sections times branching fractions into leptons (electrons, muons and taus) are used: 31.3 nb for W, and 3.05 nb for Drell-Yan production of dilepton final states with invariant mass > 50 GeV (57). The background from Drell-Yan production of dilepton final states with invariant mass < 50 GeV is negligible.

The electroweak production of single top quarks has been simulated with POWHEG. The approximate NNLO cross sections of 42 pb and 3.2 pb are used for t-channel and s-channel single top-quark production, respectively, along with the corresponding single $\bar{t}$-quark production cross sections of 23 pb and 1.4 pb. The approximate NNLO value of 7.9 pb is used for tW and $\bar{t}W$ associated production (58; 59; 60).

The simulation of QCD multijet samples is done with PYTHIA. The samples are used to validate data-driven QCD background modeling. The generated $2 \to 2$ events have a generator-level cut on $\hat{p}_T > 20$ GeV, with additional filters applied after the hadronization and the decay to enrich the sample with electrons by requiring presence of the electromagnetic object with $E_T > 20$ GeV.

As a reference model for new physics, we use the Sequential Standard Model (SSM), topcolor Z' bosons with a natural width $\Gamma_{Z'}/M_{Z'}$ equal to 1.2% (narrow width) and 10% of the Z' mass based on (17; 1; 11; 61). The signal samples are generated with MADGRAPH, assuming the same left and right-handed couplings to fermions as the SM Z boson with a range of masses

between 1 TeV and 3 TeV. The generated events include the calculation of higher-order parton radiations, as for SM $t\bar{t}$, for up to three extra partons at tree-level. The Z' decays into $t\bar{t}$ in all generated events. The simulation uses `PYTHIA` for the parton showering.

We also use Kaluza-Klein (KK) gluon samples generated with `PYTHIA`. The KK gluons decay into $t\bar{t}$ in all generated events and include higher order corrections for up to three extra partons at tree-level.

Figure 20 compares kinematic properties of the semi-leptonically decaying top quark decay products in the SM $t\bar{t}$ and Z' boson samples. The transverse momentum of all four objects becomes harder, if heavier resonances are considered.

Figure 21 compares the angular distributions of the top quark decay products for the SM $t\bar{t}$ and Z' boson samples. The quarks may not be reconstructed as separate jets if the distance between them is below the jet clustering parameter used in this analysis, $\Delta R = 0.5$. The angular separation of the hadronically decaying t quark products decreases if heavier resonances are considered and fewer number of jets can be present in the Z' boson event compared to the SM $t\bar{t}$. The angular separation between the electron and the b quark in the semi-leptonically decaying top quark decreases as resonance mass increases. The fraction of events for which this happens rises with increasing $M_{Z'}$, as expected, and the electron can not be considered isolated compared to SM $t\bar{t}$ sample.

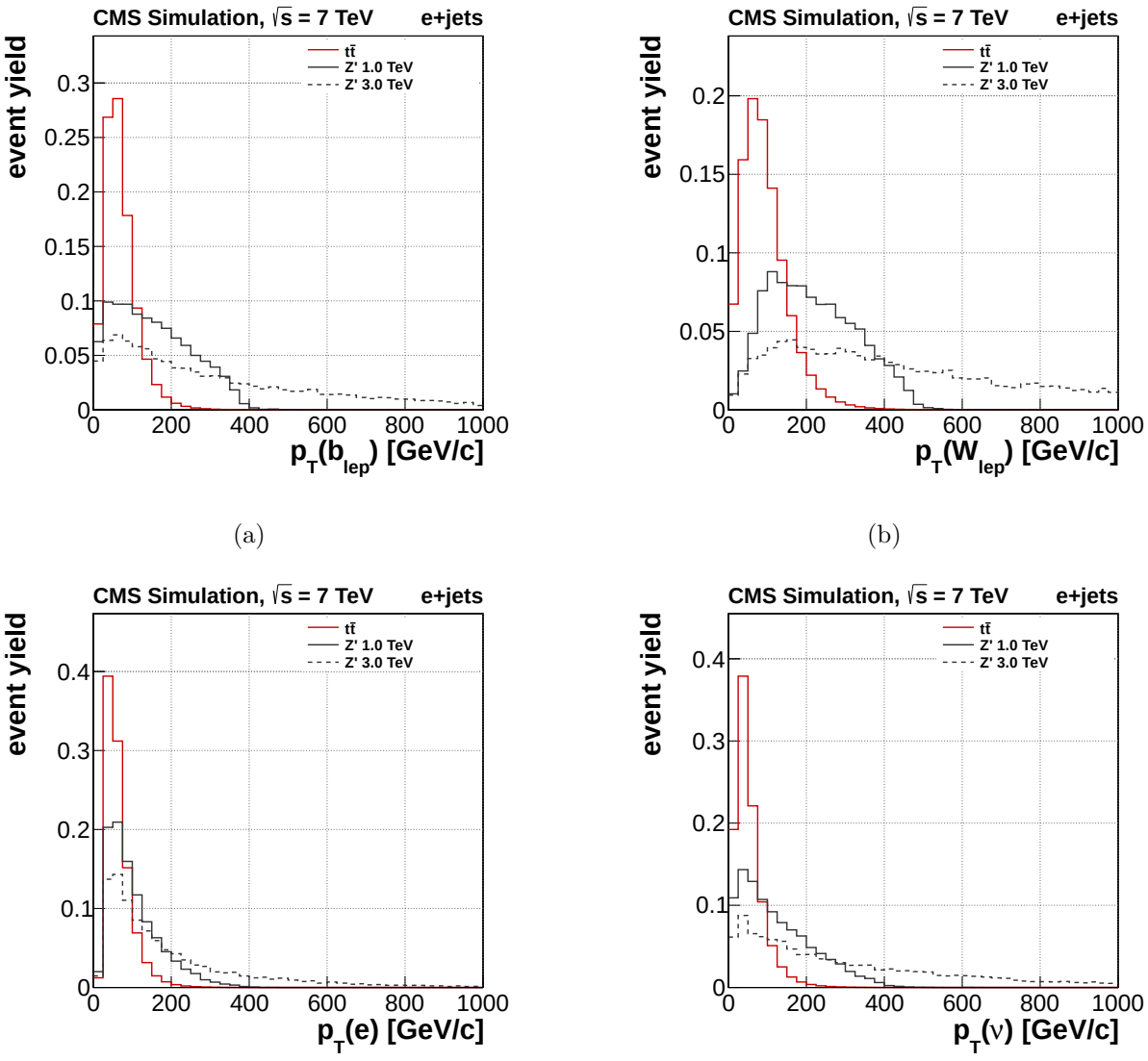
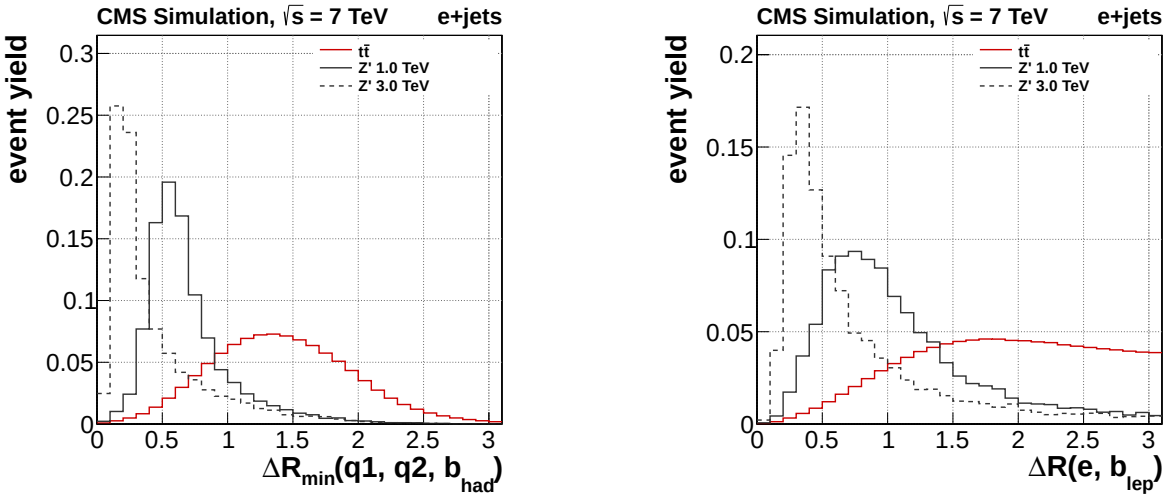


Figure 20: Kinematic properties of the semi-leptonically decaying top quark decay products for the SM $t\bar{t}$ and Z' boson samples at generator level: (a) b quark transverse momentum, (b) p_T of the W boson, (c) electron p_T, (d) p_T of the neutrino. The transverse momentum of all four objects become harder, if heavier resonances are considered. Distributions are normalized to unit area.



(a) (b)

Figure 21: Kinematic properties of SM $t\bar{t}$ and Z' boson at generator level. Minimum ΔR between any two of the three quarks of the hadronically decaying top quark (a) decreases as higher resonance masses are considered. The quarks may not be reconstructed as separate jets if the distance between them is below the jet clustering parameter used in the analysis, $\Delta R = 0.5$. The angular separation between electron and b-quark in the semi-leptonically decaying top quark (b) decreases with resonance mass. The electron can not be considered isolated compared to the SM $t\bar{t}$ sample. Distributions are normalized to unit area.

4.3 Corrections

We apply a number of different corrections to the MC samples in order to more accurately simulate CMS data-taking conditions. These include: pileup, jet energy scale (JES), jet energy resolution (JER), and b-tagging efficiency and mistag corrections as described below.

4.3.1 Pileup

Consider a total inelastic cross section of proton-proton collisions at $\sqrt{s} = 7$ TeV $\sigma_{inel} \sim$ 68 mb (62). The LHC collisions in 2011 were separated by $\tau = 50$ ns. If we use the maximum value of instantaneous luminosity from Figure 22, $L = 4.02$ Hz/nb, then the number of expected interactions per bunch crossing is:

$$N_{PU} \sim \frac{\sigma_{inel} \times L}{1/\tau} \sim \frac{68 \times 10^{-3} \times 4.02 \times 10^{9}}{(50 \times 10^{-9})^{-1}} \sim 14 \text{ interactions} \tag{4.1}$$

Typically, only one interaction has high energy and is of interest, while the others are low energy background and referred to as pileup (PU).

CMS distinguishes two types of pileup: in-time, and out-of-time. The in-time pileup is related to interactions in the current bunch crossing, and the out-of-time pileup corresponds to interactions in the previous and next bunch crossings with respect to the current one.

The origin of the out-of-time pileup lies in the detector components integration time. These interactions may leave energy or tracks in the detector for the next bunch-crossing. For example, tracker is only sensitive to in-time, while calorimeter and muon systems are affected by out-of-time pileup.

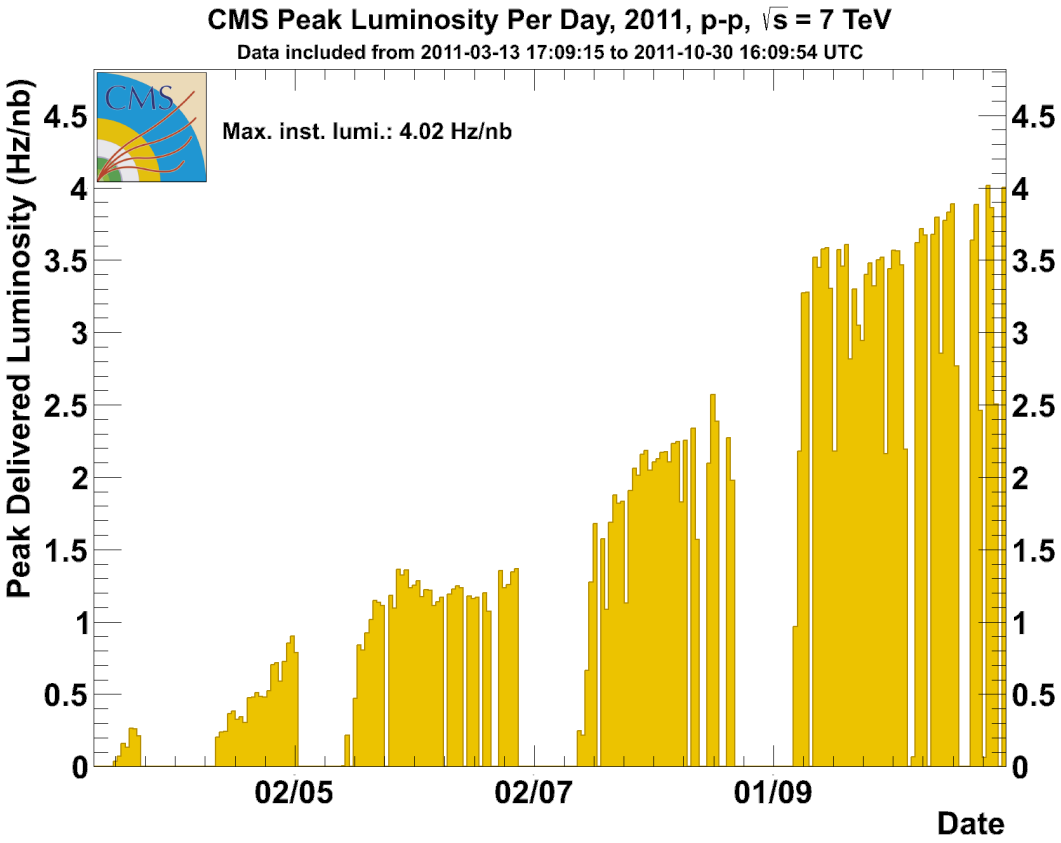


Figure 22: Peak Delivered Luminosity of proton-proton collisions at $\sqrt{s} = 7$ TeV in CMS during the 2011 data-taking period.

It becomes important to properly simulate both types of pileup in the simulated samples to match the data. Each MC sample, signal and background, includes the simulation of both types of pileup with 50 ns bunch separations.

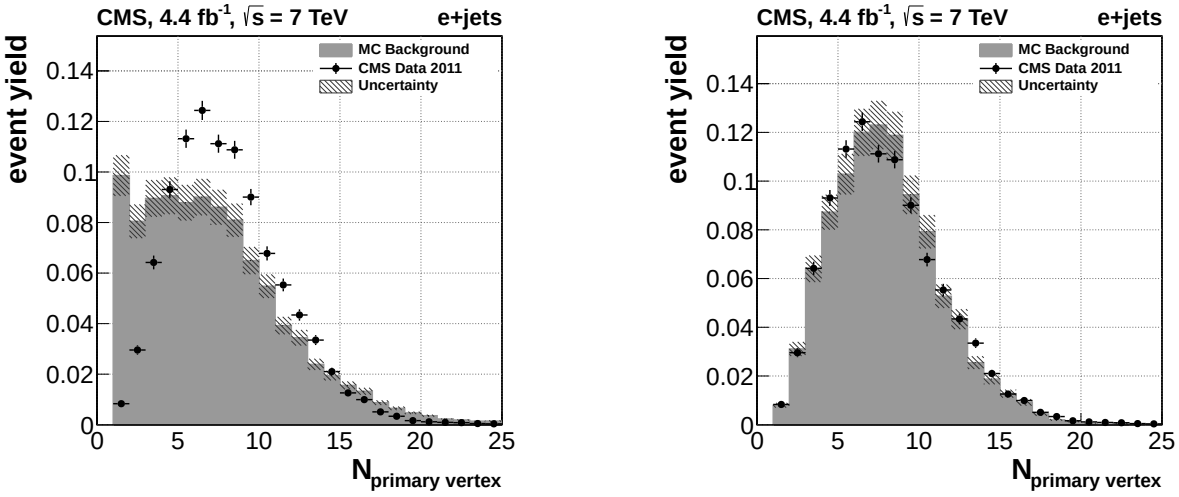
The true data pileup conditions were unknown at the time of samples generation, and a predefined pileup distribution was used: flat up to 10 pileups and Poisson beyond.

CMS adopts a 3D pileup correction method. It is a three dimensional event scales matrix which is derived from the in-time and out-of-time pileup distributions in data and MC. The scale is a function of the number of pileup interaction in the previous, current, and next bunch crossings: $SF_{PU}(N_{PV}^{prev}, N_{PV}^{curr}, N_{PV}^{next})$. We apply the correction scale factor to each generated sample. Figure 23 shows a data background comparison of the number of primary vertices before and after the pileup correction is applied. The distributions are normalized to unit area.

4.3.2 Jet Energy Scale

The calorimeter response to the jet energy is non-linear. The jet energy corrections (JEC) are a set of tools that allows the proper mapping of the measured jet energy to the lower level objects such as particles and partons.

In general, JEC is a procedure for scaling the reconstructed jet 4-momentum by a coefficient, which depends on a number of different jet properties, such as p_T and η: $\mathbf{p} \rightarrow \alpha(p_T, \eta)\mathbf{p}$. The overall JEC is split into levels and these are applied sequentially in a fixed order. Each level of correction takes care of a different effect. It is important to emphasize that the output of each step is used as input to the next level correction: $\mathbf{p} \rightarrow \mathbf{p}_1 = \alpha_1\mathbf{p} \rightarrow \mathbf{p}_2 = \alpha_2\mathbf{p}_1 \rightarrow ...$ The following jet energy corrections were used in the analysis:



(a) (b)

Figure 23: Data background comparison of the number of primary vertices before (a) and after (b) the pileup correction is applied. The distributions are normalized to the unit area.

- The L1 Pileup correction removes jet extra energy coming from the pileup interactions. This way luminosity dependency of the dataset is taken away and all other corrections become luminosity independent.
- The L2 Relative correction makes jet response flat in η. It corrects jet at any η with respect to the central region $|\eta| < 1.3$.
- The L3 Absolute flattens jet response in p_T. The corrected calorimeter jet p_T is equal on average to the generated jet p_T.
- The L2L3 Residual applies a residual calibration in data to fix small difference between data and MC. The correction depends on η and p_T, and is only applied to data.

We apply all levels of jet energy corrections to both, data and MC.

4.3.3 Jet Energy Resolution

A dijet asymmetry measurement study (63) showed that measured jet transverse momentum resolution in MC and data differ. The correction procedure is called jet energy resolution and is applied to jets after the JEC.

The jet p_T resolution is measured from two different samples in both data and MC using:

- the dijet asymmetry method, which is applied to the dijet sample,
- the photon-plus-jet balance method, which is applied to the γ+jet sample.

The dijet asymmetry method exploits momentum conservation in the transverse plane of the dijet system and is almost exclusively based on the measured kinematics of the dijet events.

The adopted strategy in CMS is to scale the reconstructed jet p_T by a value, which is based on the p_T difference between the reconstructed and matched generated jet: $p_T \rightarrow p_T + \alpha(p_T - p_T^{gen})$, where α is the resolution scaling factor, and represents the measured data/MC resolution ratio. It depends on the jet η, after the Jet Energy Corrections are applied.

4.3.4 Missing Transverse Energy

We use raw $\not\!\!E_T$ in the analysis, which is equal to the imbalance in the transverse momenta of all reconstructed particles in the event with the particle flow algorithm: $\not\!\!E_T^{raw} = -\sum_i \mathbf{p}_{T,i}^{uncorr}$. It is important to note, that uncorrected reconstructed objects are used in the $\not\!\!E_T$ definition. The JEC and JER modify reconstructed energy-momentum vectors of jets. The propagation of these changes to $\not\!\!E_T$ is a two step procedure: 1) remove uncorrected jet energy-momentum 2) add corrected jet four-vector to the $\not\!\!E_T$: $\not\!\!E_T \rightarrow \not\!\!E_T + \mathbf{p}_{jet}^{uncorr} - \mathbf{p}_{jet}^{corr}$, where $\mathbf{p}_{jet}^{uncorr}$ and $\mathbf{p}_{jet}^{corr}$ are the jet energy-momentum vectors before and after the JEC and JER corrections, respectively.

4.3.5 b-tagging

We used the Combined Secondary Vertex b-tagging algorithm at tight operating point (64; 65) to identify jets with b-flavored content. The average b-tagging efficiency with this algorithm is $\sim 52\%$ for the $t\bar{t}$ sample, and a mistag rate is about one per mile.

In our analysis, we dynamically calculate weight for each event to correct for the discrepancy between the efficiency of b-tagging a jet in data and MC. The b-tag scale factor is measured to

be 0.90 ± 0.03 (47). For an event with N_T tagged and $N_{\bar{T}}$ untagged jets the correction weight is given by:

$$w = \prod_{i=i}^{N_T} SF_T(f_i, p_{T,i}) \times \prod_{j=1}^{N_{\bar{T}}} SF_{\bar{T}}(f_j, p_{T,j}) \tag{4.2}$$

where the first (second) product loops over all tagged (untagged) jets. The SF_T ($SF_{\bar{T}}$) is defined as follows:

$$SF_T(f, p_T) = \frac{\varepsilon_f^{data}(p_T)}{\varepsilon_f^{mc}(p_T)} \qquad SF_{\bar{T}}(f, p_T) = \frac{1 - \varepsilon_f^{data}(p_T)}{1 - \varepsilon_f^{mc}(p_T)} \tag{4.3}$$

where SF_T ($SF_{\bar{T}}$) corrects any discrepancy between data and MC efficiencies for a jet to pass (fail) the b-tagging requirement. In this analysis we use the values of SF_T provided by the b-tagging group for different jet flavors. In particular, we use SF_b as measured from the μ-jets sample and SF_l as measured from the dijet sample. Following the b-tag POG recommendations, we use $SF_c = SF_b$ with twice the quoted uncertainty.

The $SF_{\bar{T}}$ for heavy flavor jets, $f = \{b, c\}$, are derived from SF_T using

$$SF_{\bar{T}}(f = \{b, c\}, p_T) = \frac{1 - SF_T(f, p_T)\varepsilon_f^{mc}(p_T)}{1 - \varepsilon_f^{mc}(p_T)} \tag{4.4}$$

where $\varepsilon_f^{mc}(p_T)$ is the b-tagging MC-truth efficiency measured in the $t\bar{t}$ sample. In the case of light flavor, $SF_{\bar{T}}$ is derived using

$$SF_{\bar{T}}(l, p_T) = \frac{1 - \varepsilon_l^{data}(p_T)}{1 - \varepsilon_l^{data}(p_T)/SF_T(l, p_T)} \tag{4.5}$$

where $\varepsilon_l^{data}(p_T)$ is the efficiency of b-tagging a light jet measured from dijet samples.

Even though the number of events with tagged jets does not change, their contribution to the yields is modified by the event weight when allowing SF_b and SF_l to vary within their systematic uncertainties. The use of event weights thus allows for effective bin migration between the samples with 0 and ≥ 1 b-tagged jet.

CHAPTER 5

EVENT SELECTION AND QCD MODELING

We perform a model independent search for the production of a heavy resonance that decays to $t\bar{t}$ where one W boson from a top quark decays leptonically into an electron and a neutrino, and the other W boson decays hadronically. We expect a high-p_T electron, at least two high-p_T jets, and large missing transverse energy. For energetic $t\bar{t}$ events the decay products of the hadronically decaying top quark might not be well separated in ΔR in the detector and might be reconstructed as less than three jets. The electron from the semi-leptonically decaying top quark may not be well isolated. Therefore our event selection requires at least two jets in the final state and applies no isolation requirement on the electron.

5.1 Preselection

The goal of the preselection is to reduce the dataset in size which can be easily processed while still be inclusive enough to allow a sideband studies. We apply noise-reduction filters, HLT and event kinematic cuts. In addition, the preselection removes an overlap between the electron+jets and muon+jets channels by excluding the events with a muon. Compared to the muon+jets channel, the preselection in the electron+jets channel results in a sample that has a large QCD multijet contamination that is not properly described by the available QCD MC. To attain a proper data/MC comparison, we devise a series of topological cuts that remove most of the QCD multijet background.

Details of the preselection cuts we apply are listed below.

1. the event must pass High Level Trigger `HLT_Ele45_CaloIdVT_TrkIdT` (or `HLT_Ele65_CaloIdVT_TrkIdT` after run number 165970), [1]

2. at least one reconstructed Primary Vertex should be present in the event,

3. the event must have at least two cleaned jets with $p_T > 50$ GeV and $|\eta| < 2.4$, with a leading jet $p_T > 150$ GeV,

4. one Particle Flow electron with $p_T > 70$ GeV and $|\eta| < 2.5$; the electron must pass a Hyper Tight 1 electron identification with conversion rejection,

5. veto the event if a second lepton, either electron or muon, is present,

6. the electron must pass 2D cut, which is defined between the electron and the closest reconstructed jet with $p_T > 25$ GeV:

$$\Delta R(e, jet) > 0.5 \qquad \text{or} \qquad p_T^{rel}(e, jet) > 25 \text{ GeV} \ [2]$$

7. $H_T^{lep} = \not{E}_T + p_T^e > 150$ GeV,

---

[1] MC was simulated with HLT electron p_T threshold at 45 GeV. The trigger efficiency plateau is reached at 50 GeV. Because of the higher offline electron p_T cut, the trigger in MC is guaranteed to be in the fully efficient region.

[2] A transverse momentum p_T^{rel} of the electron relative to the direction of the closest reconstructed jet with $p_T > 25$ GeV is defined as the magnitude of the electron momentum orthogonal to the jet axis:

$$p_T^{rel} = |\mathbf{p}_e| \sin \alpha$$

where α is the angle between the jet and the electron momenta.

8. triangular cuts between the electron (e) or leading jet (j) and missing transverse energy:

$$-\frac{1.5}{75\text{GeV}}\not{E}_T + 1.5 < \Delta\phi\,\{(e \text{ or } j),\ \not{E}_T\} < \frac{1.5}{75\text{GeV}}\not{E}_T + 1.5$$

9. $\not{E}_T > 50$ GeV,

10. $p_T(t_{lep}) > 100$ GeV [1].

The good primary vertex requirement (step 2) guarantees the existence of a well reconstructed primary vertex. The primary vertex with the highest p_T sum of the attached tracks has to fulfill $|z| < 24$ cm, $\rho < 2$ cm, and $N_{dof} \geq 4$, where N_{dof} is the weighted number of tracks used to reconstruct the primary vertex.

The preselection steps 6 to 10 are exclusively designed to suppress the QCD multijet background. Figure 24 shows 2D cut. The largest background $W \to l\nu$ as well as signal samples have most of the events at large values of ΔR or p_T^{rel} whereas QCD multijet tends to accumulate at low region. The event selection cut 6 removes low region.

Figure 25 and Figure 26 show triangular cuts for the electron (leading jet) and missing transverse energy. The QCD multijet tends to have low values of missing transverse energy and is peaking in the opposite $\Delta\phi$ region compared to the signal Z' boson samples because of either lepton or $\not{E}_T$ misidentification. We remove corresponding regions with cuts 8 and 9.

Figure 27 shows the reconstructed p_T of the semileptonically decaying top quark. We attribute the gap at low values between data and MC to QCD multijet and remove corresponding events with cut 10.

---

[1] This variable is described in Chapter 6

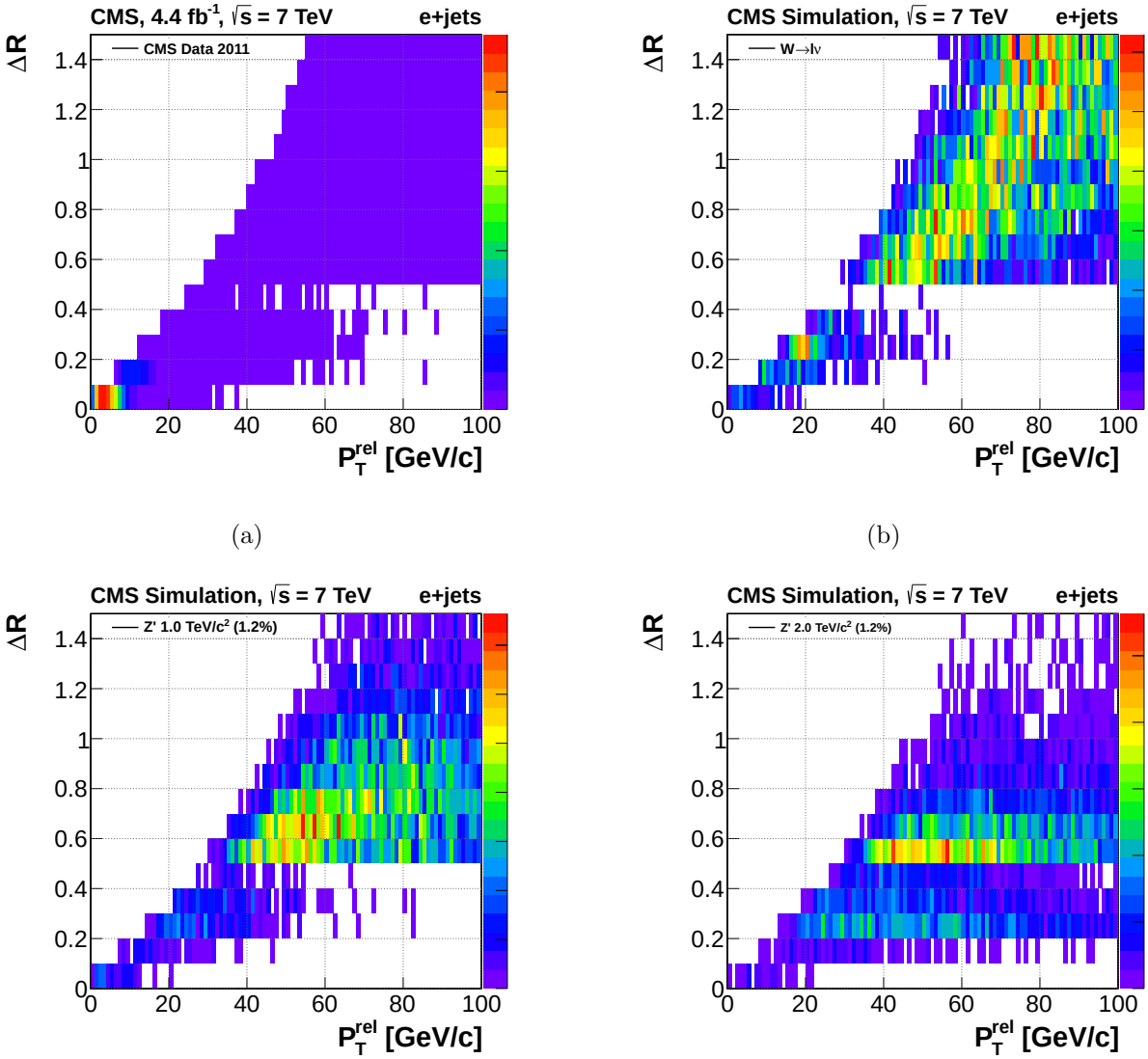


Figure 24: 2D cut in data (a), $W \to l\nu$ (b), and signal Z' boson samples with mass 1 TeV (c) and 2 TeV (d). The signal Z' boson samples are normalized to 1 pb.

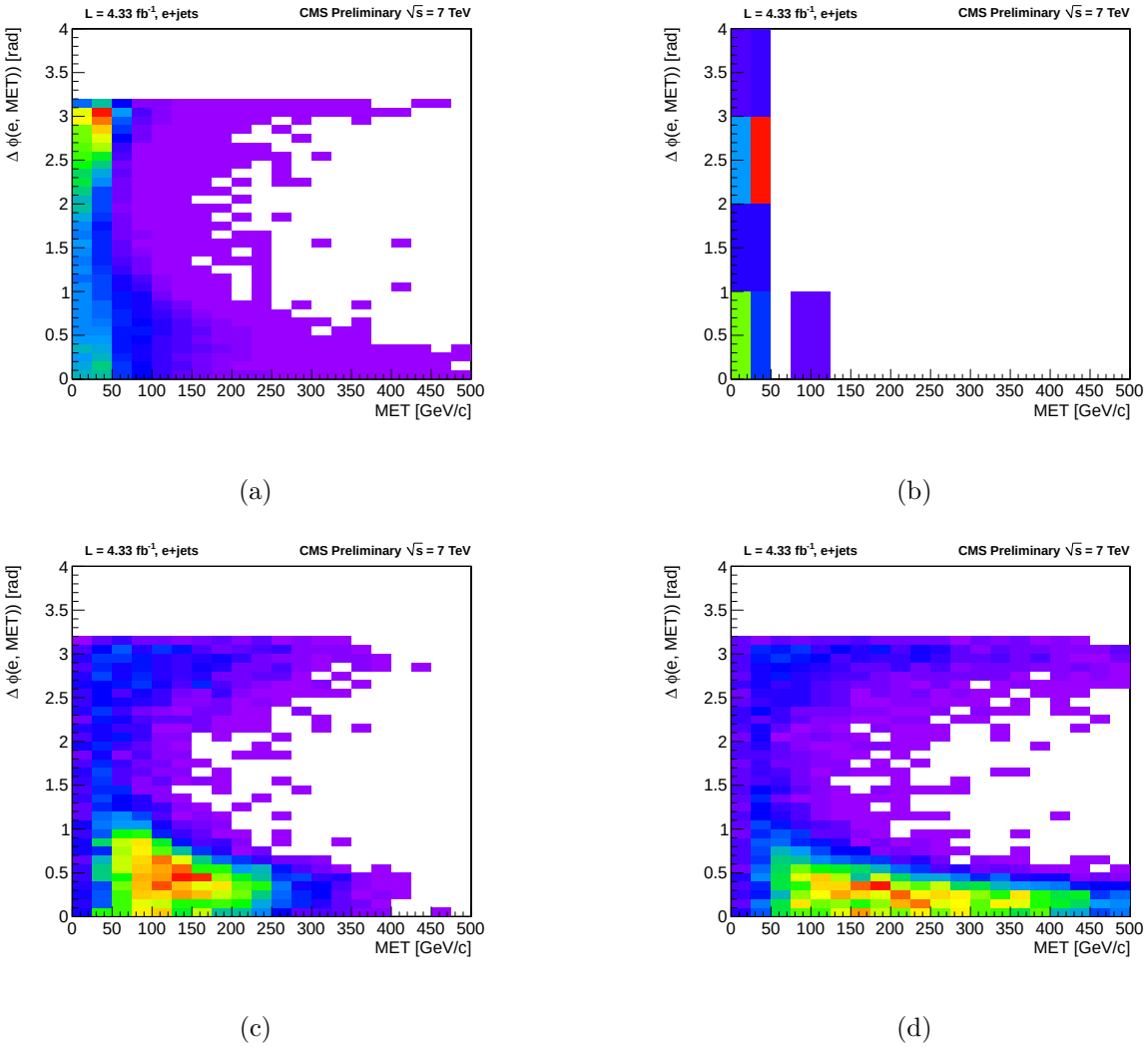


Figure 25: Triangular cut for the electron and missing transverse energy in data (a), QCD multijet (b), and signal Z' boson samples with mass 1 TeV (c) and 2 TeV (d). The signal Z' boson samples are normalized to 1 pb.

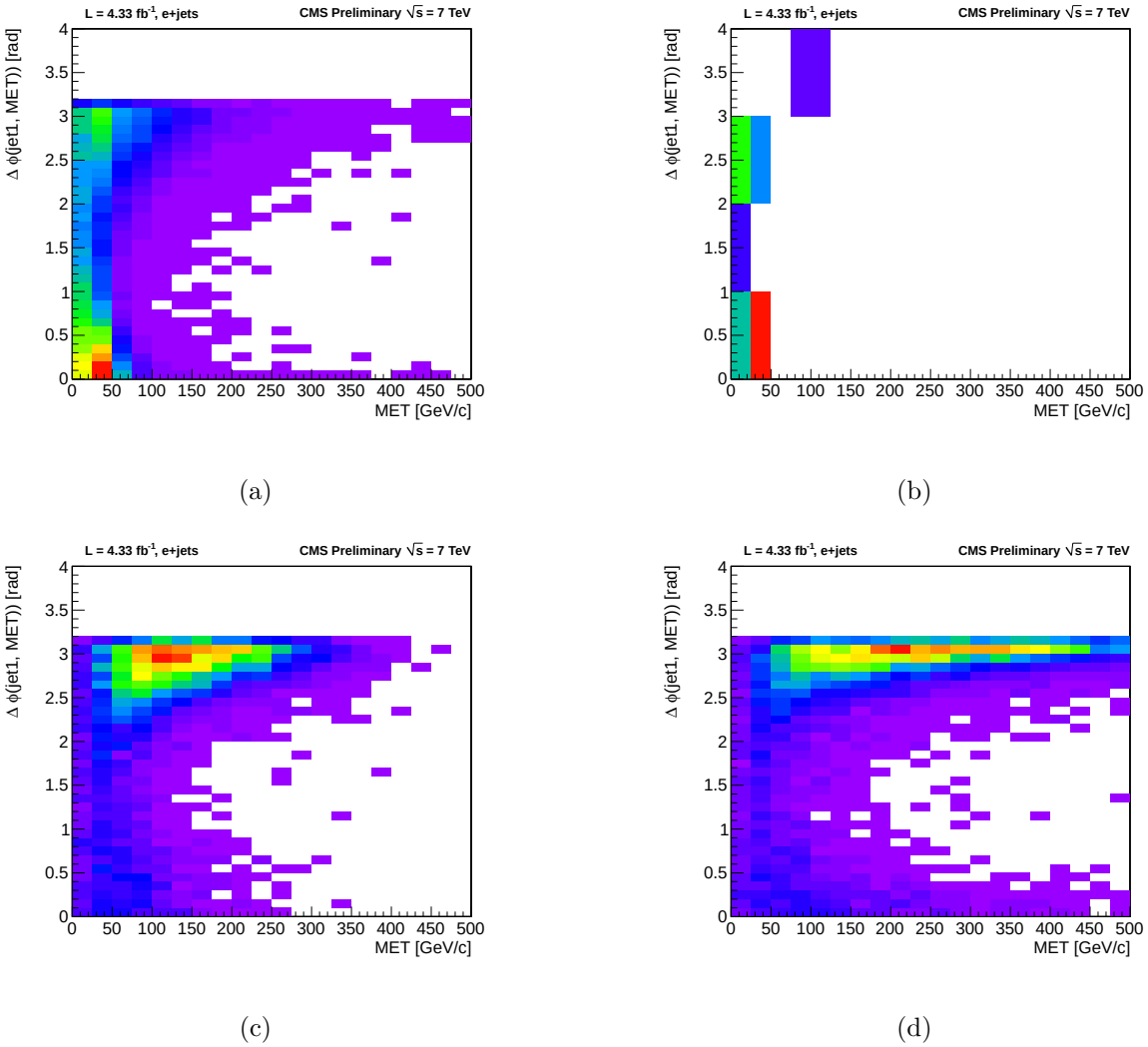


Figure 26: Triangular cut for the leading jet and missing transverse energy in data (a), QCD multijet (b), and signal Z' boson samples with mass 1 TeV (c) and 2 TeV (d). The signal Z' boson samples are normalized to 1 pb.

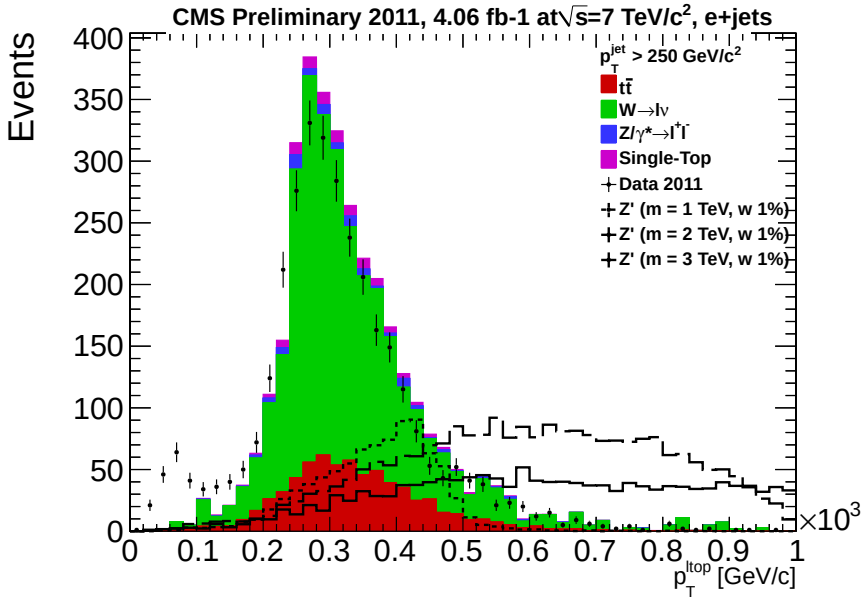


Figure 27: Data background comparison of the reconstructed leptonic top transverse momentum. The signal Z' boson samples are normalized to 1 pb.

5.2 High Level Trigger Efficiency

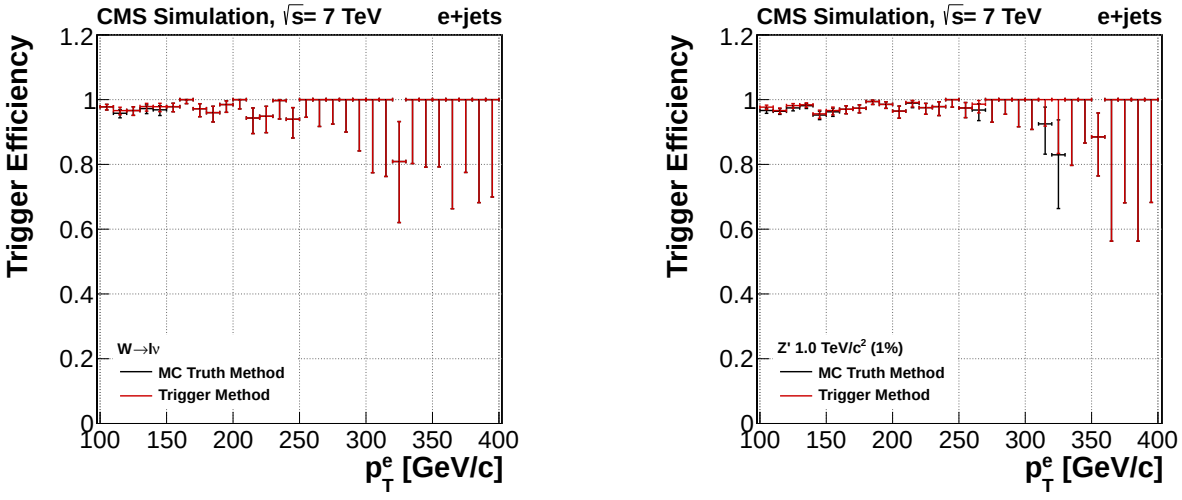
The first step of the preselection is the requirement that the candidate event passes the High Level Trigger.

Due to a highly boosted nature of the signal samples, we select a trigger without an isolation requirement on the electron, `HLT_ElePT_CaloIdVT_TrkIdT`, where `ElePT` is the candidate electron p_T threshold, and was equal to 45 GeV in the early data, but increased to 65 GeV starting from run 165970. The p_T threshold change was necessary for HLT to keep the output events rate below its maximum value due to the increase in instantaneous luminosity.

We studied the emulation of the `CaloIdVT_TrkIdT` term in the simulation and compared it to data by measuring the efficiency of the trigger for the events that pass the preselection and a reference trigger without the electron ID component $\varepsilon_{trig} = \frac{N_{trig}}{N_0}$, where N_0 is the number of preselected events that pass the reference trigger, and N_{trig} represents the number of preselected events that pass both: the reference and the trigger whose efficiency is being measured. It is assumed that the reference trigger is fully efficient for the signal events that passed the preselection. Finally, a scale factor can be derived to correct MC to data if efficiencies are different $SF_{trig} = \frac{\varepsilon_{data}}{\varepsilon_{MC}}$.

The 2011 Single Electron data stream contains only one HLT path that fulfills the reference trigger requirements: `HLT_Ele90_NoSpikeFilter`. It was only available in the first 200 pb^{-1} of data and then was removed from the trigger menu. Because of the higher threshold in the reference trigger, we measured the trigger efficiency with the full preselection and increased the electron transverse momentum cut $p_T > 100$ GeV.

We validated the efficiency measurement method using the W+jets, the largest SM background, and signal $M_{Z'} = 1$ TeV samples, and compared the measured efficiencies with the MC truth efficiency. The results of the test are shown in Figure 28 and agree very well, validating the proposed trigger efficiency measurement method in MC.



(a) (b)

Figure 28: A comparison between the MC truth efficiency and the measured efficiency for a W+jets (a) and signal $M_{Z'} = 1$ TeV samples (b) as a function of the electron p_T.

We used the proposed method and measured the trigger efficiency as a function of the electron transverse momentum p_T in the largest background $W \to l\nu$, and data. The measured

efficiencies are shown in Figure 29. Both distributions are fitted with a constant. The average trigger efficiency is 0.975 ± 0.004 for $W \to l\nu$ and 0.948 ± 0.015 for data. The ratio of measured efficiencies represents the data-MC scale factor and is equal to $SF_{trig} = 0.972 \pm 0.016$.

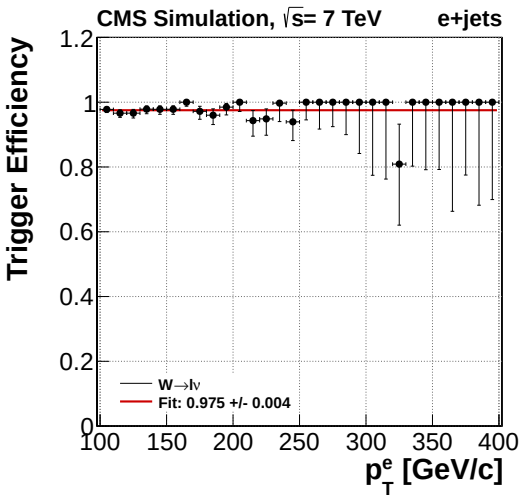
We repeated the same study with added isolation and impact parameter requirements to the offline electron identification in order to address a possible contamination of the data with QCD multijet events. The measured trigger efficiencies are shown in Figure 29. We obtain the correction scale factor $SF_{tirg}^{iso,ip} = 0.983 \pm 0.016$. The small increase in the value is consistent with our expectations that QCD multijet contamination is present in the non-isolated data sample, which would lower the measured trigger efficiency in data compared to that in MC.

Based on these studies, we assign a flat $SF_{trig} = 0.96 \pm 0.04$ to the trigger efficiency, which includes a conservative uncertainty to cover both the statistical uncertainties and the difference between SF_{trig} and $SF_{trig}^{iso,ip}$.

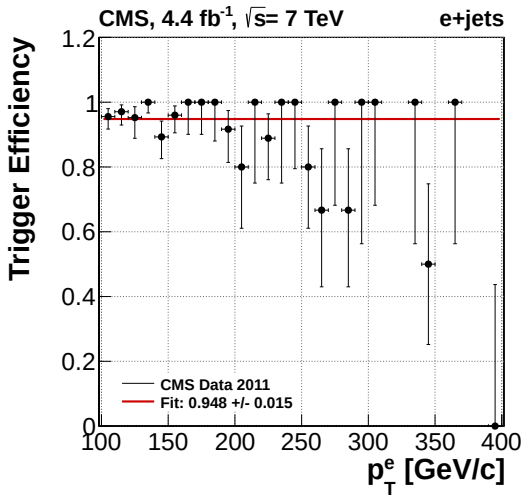
5.3 Electron Identification

We do not expect all electrons in the signal to be isolated and do not cut on the isolation. However, after selecting the events with high-energy electrons, which are candidates for the prompt electron from $t\bar{t}$ signal, some of these electrons will be isolated (and not clustered in jets), others will be non-isolated (and clustered to jets). To avoid a double-counting of the energy in the reconstruction, we perform a jet-electron cleaning. Technically, this is done by the following procedure:

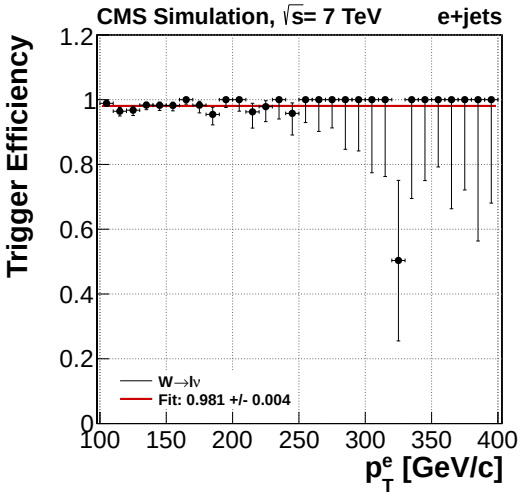
- we consider each electron which passes the ID cuts above and kinematic cuts as a candidate for a prompt electron,



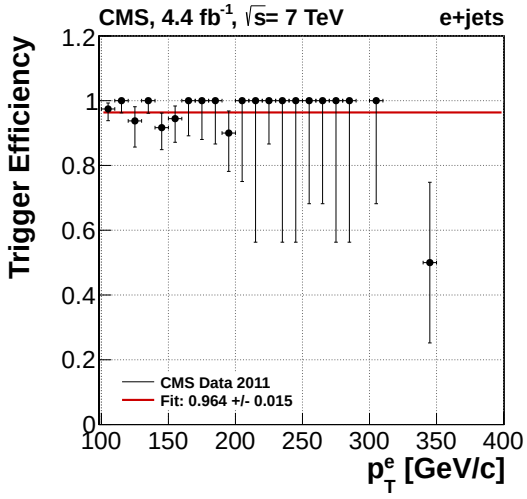
(a)



(b)



(c)



(d)

Figure 29: The trigger efficiencies measured in the $W \to l\nu$ (a) and 2011 data (b) with the full preselection and increased offline electron transverse momentum threshold $p_T > 100$ GeV. The trigger efficiency with the addition of isolation and impact parameter requirements in $W \to l\nu$ (c) and 2011 data (d).

- search the jet to which this electron has been clustered using the actual jet clustering information, not just a ΔR requirement,
- subtract the electron four momentum from the uncorrected jet energy-momentum,
- re-correct the jet based on the modified uncorrected 4-momentum.

The cleaned jets are then used in the analysis, including the selection of the events and the reconstruction of the invariant top pair mass. We perform the jet-electron cleaning rather than a jet re-clustering as this ensures that different analyses share the same jet definition.

The electron identification method efficiency is based on a set of cuts that are optimized to select electrons from Z or W decays and reject fakes from jets or conversions. These cuts are split into categories as follows:

- the identification cut matches a track with the electromagnetic calorimeter, searches for the energy in the hadronic calorimeter behind the electromagnetic calorimeter cluster, and analyzes the cluster shape,
- the impact parameter category cuts on the impact parameter with respect to the reconstructed vertex,
- the conversion cut rejects photon conversions,
- the isolation category calculates the isolation in the Tracker, electromagnetic, and hadronic calorimeters.

The algorithm generates four flags, one per category. The flags have nine severity levels and vary from Very Loose to Hyper Tight 4.

We measured the efficiency of the identification and conversion rejection cuts in the MC QCD multijet and Z' samples with $M_{Z'} = 1$ TeV and $M_{Z'} = 3$ TeV. The measurement requires each event to pass a relaxed preselection in order to have an adequate statistics in the QCD multijet sample. For this, we lowered the leading jet and electron transverse momentum thresholds to 50 GeV and 30 GeV respectively, and turned off the preselection steps 6 to 10. The results are shown in Figure 30.

Based on these plots, we select the Hyper Tight 1 operation point of the electron identification with conversion rejection for the analysis. This choice has a flat efficiency of $\sim 90\%$ for signal samples with masses between 1 TeV and 3 TeV, and a fake rate slightly below $\sim 30\%$ for the QCD multijet background. No requirement on the isolation is applied.

5.4 QCD Multijet Modeling

The preselection steps 6 to 10 are designed to reduce the QCD multijet background, but there is still relatively large fraction of these events that pass the cuts. The MC simulations of the QCD multijet cover the low region of $\hat{p}_T$ up to 170 GeV and do not cover the analysis phase-space in full.

In order to have an adequate description of the QCD multijet background, we model it from data in the control region. The events must pass the preselection but fail the electron identification (step 4) and triangular requirements (step 8). This modified preselection is called the inverted preselection. Figure 31 shows the $\not{E}_T$ and $M_{t\bar{t}}$ distributions for these events. The QCD multijet events dominate the resulting sample with a purity of $\sim 94\%$.

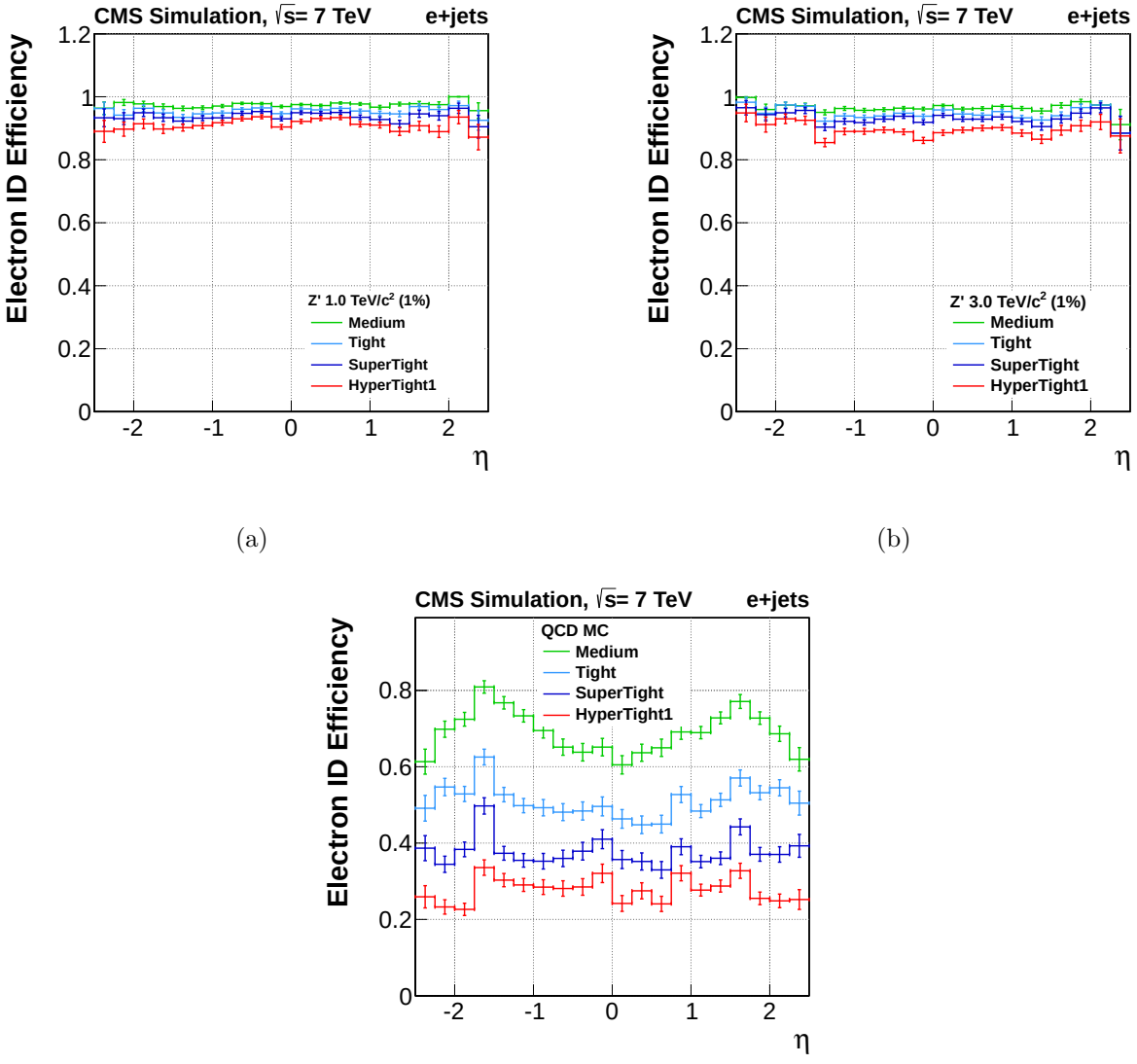
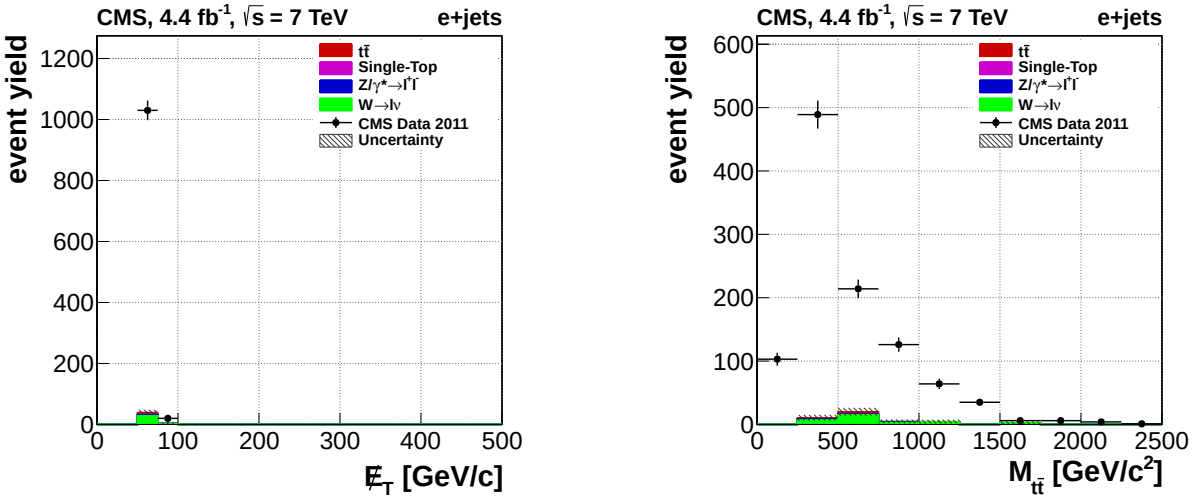


Figure 30: The measured efficiency of the electron identification and conversion rejection as a function of the electron η for different MC samples: $M_{Z'} = 1$ TeV (a), $M_{Z'} = 3$ TeV (b), MC QCD multijet (c).



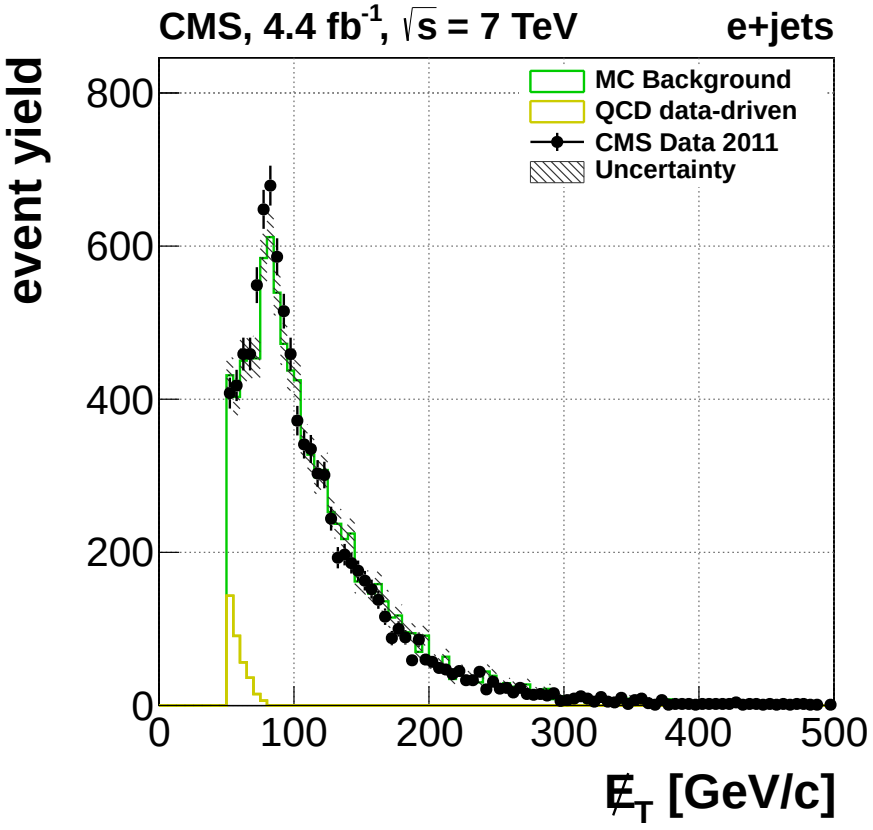
(a) (b)

Figure 31: Data background comparison of the missing transverse energy (a) and $M_{t\bar{t}}$ (b) in the QCD multijet dominated region: the events must pass the preselection but fail the electron identification and triangular requirements. The empty region between the data and SM background prediction is attributed to the QCD multijet.

We extract the fraction of QCD multijet events by fitting the $\not\!\!E_T$ distribution in the data after the full preselection with a QCD multijet template derived with the inverted preselection and a separate template given by the sum of all other SM sources after the full preselection, as shown in Figure 32. The fit is implemented using the TFractionFitter (66), which takes into account the effect of MC statistics. The fraction of QCD multijet background is estimated to be $\sim 5\%$ from the fit. We use this value to normalize the QCD multijet template from data, and use the normalized templates as a model for the QCD multijet, assuming a normalization uncertainty of 50%.

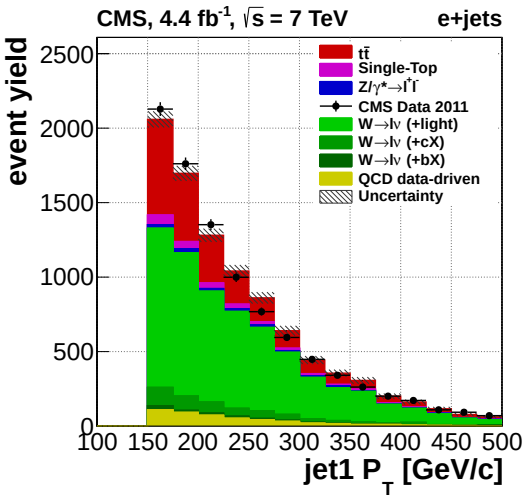
5.5 Control Plots

After the preselection, we perform a comparison between the data and SM prediction of different object kinematic distributions in the event. All distributions have the pileup, jet energy scale and resolution corrections applied according to the description in Section 4.3. The QCD multijet background is modeled from and scaled to data. The background uncertainty band on the plots includes 2.2% and 5% errors on the luminosity and trigger, respectively. The results are shown in Figure 33 to Figure 35. A good agreement between the data and background is observed in all cases.

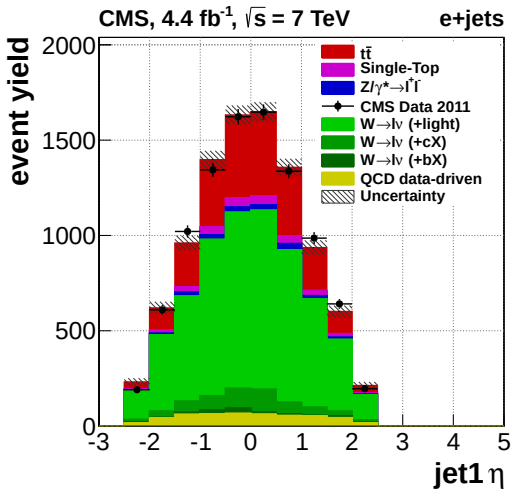


(a)

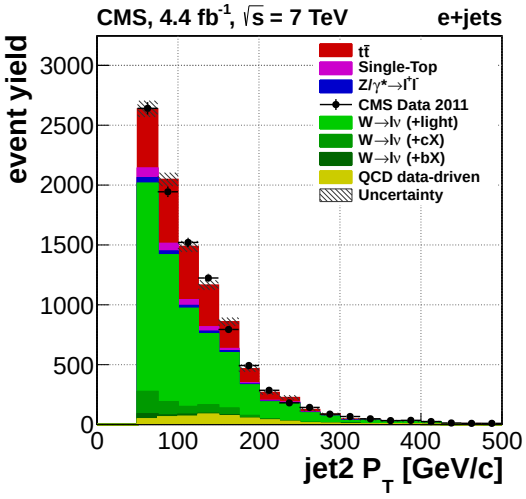
Figure 32: The missing transverse energy distribution in data fitted to the sum of MC background and data-driven QCD, which is extracted in the control region, where the events must pass the preselection but fail the electron identification or triangular requirements.



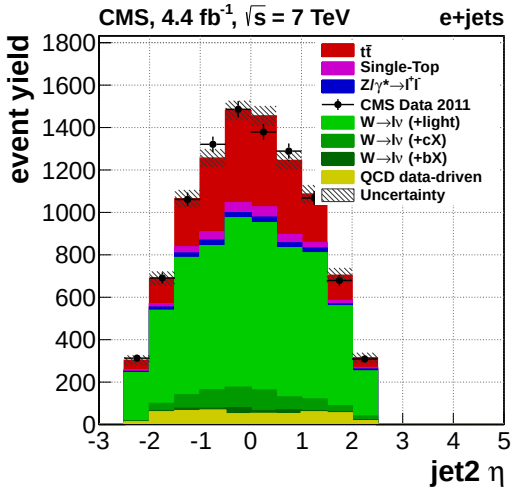
(a)



(b)

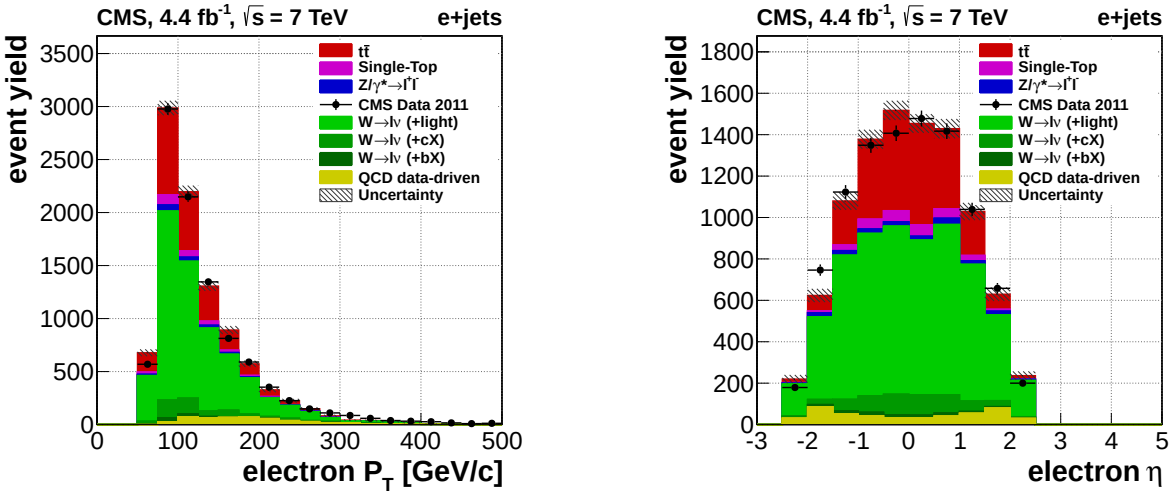


(c)



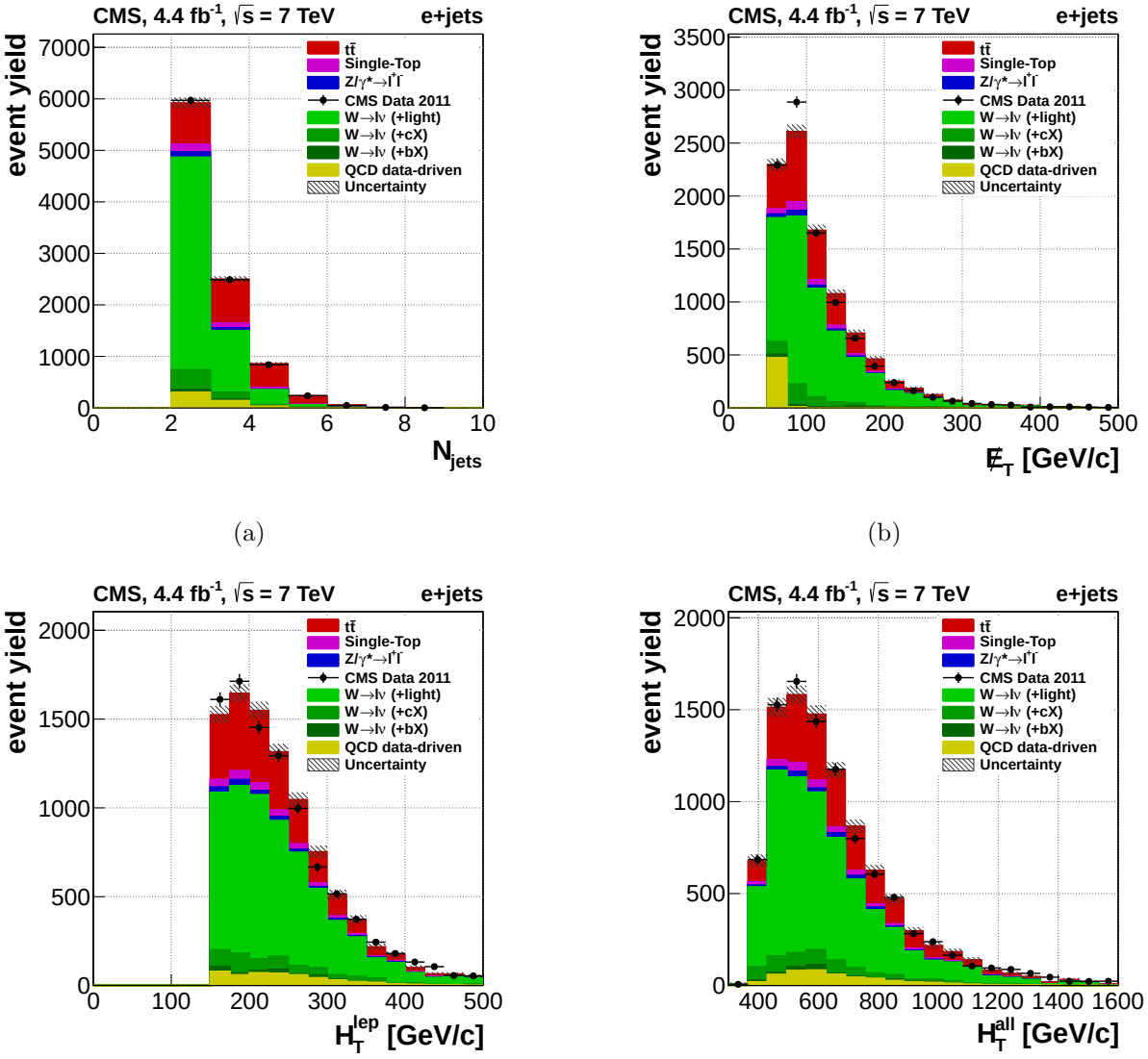
(d)

Figure 33: Data background comparison after the preselection for the leading jet p_T (a) and η (b), the second leading jet p_T (c) and η (d). The QCD multijet and SM background are scaled to data with the $\not\!\!E_T$ fit results. The error uncertainty band includes 2.2% and 5% errors on the luminosity and trigger.



(a) (b)

Figure 34: Data background comparison after the preselection for the electron p_T (a) and η (b). The QCD multijet and SM background are scaled to data with the $\not\!\!E_T$ fit results. The error uncertainty band includes 2.2% and 5% errors on the luminosity and trigger.



(a) (b)

(c) (d)

Figure 35: Data background comparison after the preselection of a number of jets (a), the $\not\!\!E_T$ (b), H_T^{lep} (c), and H_T^{all} (d). The QCD multijet and SM background are scaled to data with the $\not\!\!E_T$ fit results. The error uncertainty band includes 2.2% and 5% errors on the luminosity and trigger.

CHAPTER 6

EVENT RECONSTRUCTION AND FINAL SELECTION

6.1 Event Reconstruction

We reconstruct the four-vectors of the top and the antitop quark candidates for each event using a χ^2 approach. The objective is to assign each object in the event - charged lepton, $\not\!\!E_T$, or jets - to either the leptonic or the hadronic side of the $t\bar{t}$ pair decay. We call each possible objects assignemt a hypothesis. Each event will have my different hypotheses and ambiguities in the assignments are resolved using the criteria specified below.

The resonance reconstruction is a two steps process: first, we reconstruct the neutrino longitudinal momentum component, p_z, then assign the objects to the leptonically or hadronically decaying tops.

6.1.1 Neutrino Reconstruction

In the current analysis, we interpret the $\not\!\!E_T$ as the transverse component of the momentum of the neutrino, which is produced from the W boson semi-leptonic decay, $W \to l + \nu$. Assuming the W boson was on-shell, one can derive a quadratic equation for the z component of the momentum of the neutrino from the energy-momentum balance (see Appendix A for a full derivation). The solutions are given by:

$$p_{z,\nu} = \frac{\alpha\, p_{z,l}}{p_{T,l}^2} \pm \sqrt{\frac{\alpha^2\, p_{z,l}^2}{p_{T,l}^4} - \frac{E_l^2\, p_{T,\nu}^2 - \alpha^2}{p_{T,l}^2}} \tag{6.1}$$

where $\alpha = \frac{m_W^2}{2} + \mathbf{p}_{T,l} \cdot \mathbf{p}_{T,\nu}$; $\mathbf{p}_{T,l}$ and $\mathbf{p}_{T,\nu}$ are the transverse momenta of the lepton and neutrino, respectively; m_W is the W boson mass; p_l and p_ν are the magnitude of the electron and neutrino momenta, respectively; E_l and $p_{z,l}$ are the energy and longitudinal momentum of the lepton. Equation 6.1 has either 0, 1, or 2 solutions. In the absence of a real solution, we take the real part of the complex solution. If there are two solutions then both are considered in the resonance reconstruction.

6.1.2 Jets Assignment

We construct a list of all possible jet assignments, called hypotheses. Each jet is assigned to either the semileptonically decaying top quark, the hadronically decaying top quark, or neither of the two. A total number of hypotheses per event is equal to $N_{\nu-sol} \cdot 3^{N_{jets}}$, where $N_{\nu-sol}$ is the number of considered neutrino p_z-solutions in the event, and is always either 1 or 2. The N_{jets} is the number of selected jets in the event.

For each hypothesis, the top quark 4-vectors are given by the sum of the 4-vectors of all assigned decay products (jets, lepton, neutrino). The electron and neutrino are always added to the semi-leptonically decaying top quark. We also require each top quark to have at least one jet assigned. If more than one jet is assigned to the semi-leptonically decaying top quark, only the leading jet is used in the $M_{t\bar{t}}$ calculation.

Of these $N_{\nu-sol} \cdot 3^{N_{jets}}$ hypotheses, we select one hypothesis per event. The selection criteria is based on the idea that the reconstructed top quark masses should be close to the true top quark mass. We build:

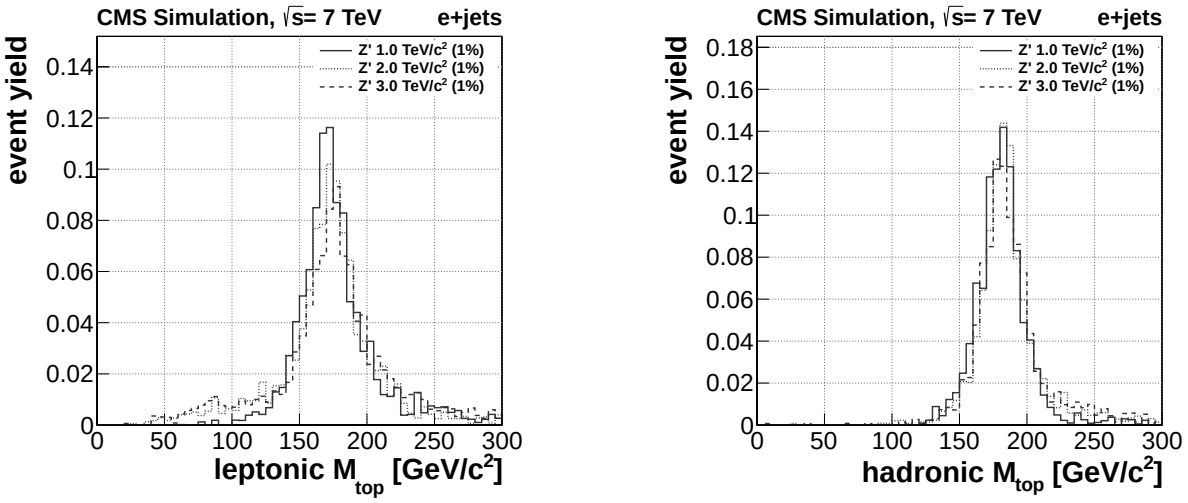
$$\chi^2 = \left[\frac{M_{lep} - \bar{M}_{lep}}{\sigma_{M_{lep}}}\right]^2 + \left[\frac{M_{had} - \bar{M}_{had}}{\sigma_{M_{had}}}\right]^2 \tag{6.2}$$

where M_{lep} and M_{had} are the reconstructed invariant masses of the semileptonically and hadronically decaying top quarks for each hypothesis. The parameters $\bar{M}_{lep}$, $\bar{M}_{had}$, $\sigma_{M_{lep}}$, and $\sigma_{M_{had}}$ are chosen such that resulting χ^2 is small for the correct hypothesis.

We extract the values of these parameters from the MC Z' events in which all four partons of the final state top quark decay products are matched to the reconstructed jets used in the hypothesis, using $\Delta R(parton, jet) < 0.3$ as a matching criterion. One reconstructed jet can be matched to more than one parton if the hadronic top quark has large p_T and the partons have a small opening angle. In this case, it can even happen that all the decay products are reconstructed as only one jet and all three partons of the hadronic top quark are matched to the same jet. In those events, where all four partons can be matched (we call these the matchable events), there is always an obvious correct hypothesis which has the jet assignment according to the matching.

Figure 36 shows the distribution of $\mathrm{M_{lep}}$ and $\mathrm{M_{had}}$ for the correct hypotheses in the matchable events. We fit this mass distribution with a Gaussian function to find the values of $\bar{M}_{lep}$, $\bar{M}_{had}$, $\sigma_{M_{lep}}$, and $\sigma_{M_{had}}$. The mean and width of the Gaussian fit are found to be independent

of the Z' boson mass to a very good approximation. We find the values $\bar{M}_{lep} = 172 \pm 18$ GeV, $\bar{M}_{had} = 179 \pm 14$ GeV.

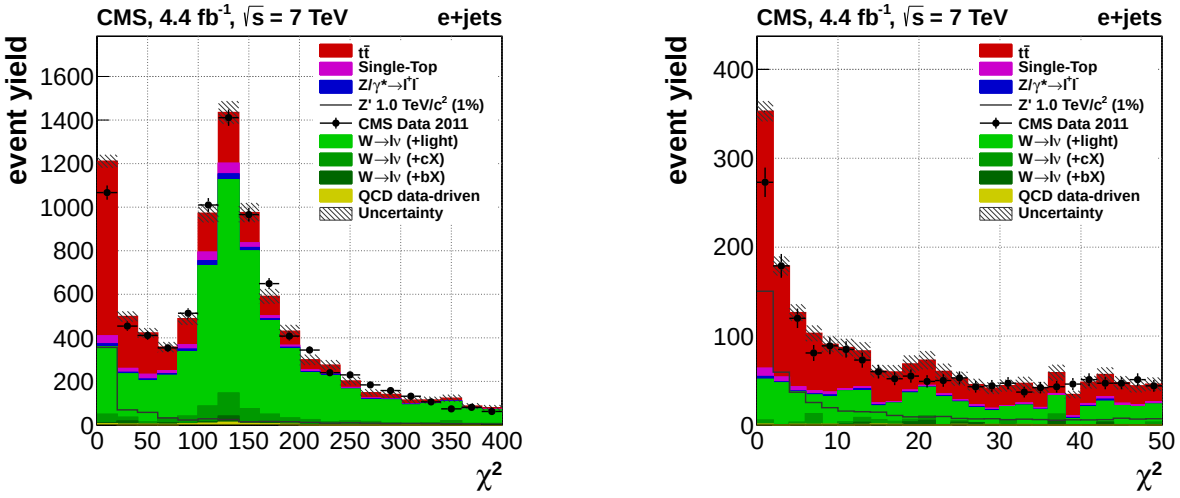


(a) (b)

Figure 36: The reconstructed invariant top quark mass on the leptonic (a) and hadronic (b) sides for the correct hypothesis in the matchable events. The samples are normalized to the unit area.

By construction, the χ^2 is expected to be small for the correct jet-parton assignment. Therefore, we select the hypothesis with lowest value of the χ^2 for each event. The comparison between

the data and simulated events for the χ^2 variable in the wide and narrow ranges is shown in Figure 38. The signal Z' boson sample is normalized to 5 pb in these plots.



(a) (b)

Figure 37: Data background comparison of the χ^2 in a wide range (a) and at low values (b). The signal Z' boson sample is normalized to 5 pb.

The signal and irreducible $t\bar{t}$ background tend to have low values of the χ^2 because of the correct leptonic and hadronic top quark reconstructions as expected. Using the matchable events, we can evaluate the fraction of events in which the selected hypothesis coincides with the correct one. Depending on the sample, this fraction is between 65% and 70%. Other

backgrounds have large values of the χ^2. The $W \to l\nu$ background peak at 140 and long tail in the irreducible $t\bar{t}$ background are attributed to misreconstruction. This is a strong indication that an additional cut on the value of the χ^2 can be applied to reduce this background.

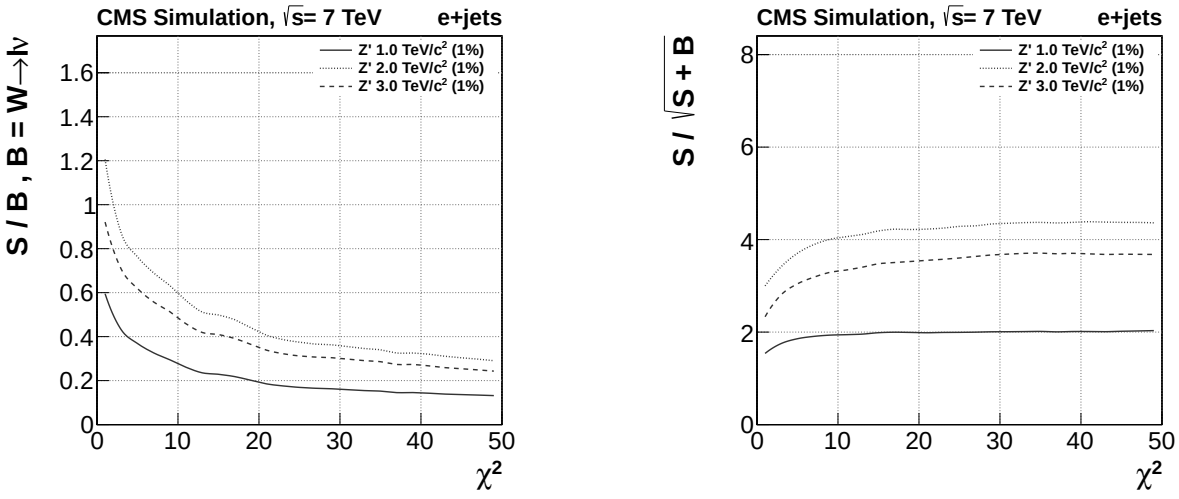
Figure 38a shows the signal over W+jets background ratio for three different signal samples with $M_{Z'} = 1,\ 2,$ and 3 TeV. The signal yield is obtained assuming a cross section of 1 pb. As can be seen, the rejection of W+jets relative to the signal events increases monotonically, e.g. the harder the cut, the cleaner the sample would be.

We also need to ensure that there is a proper MC statistics left after the cut for the statistical analysis. Figure 38b shows $S/\sqrt{S+B}$, where B accounts for all background samples. It can be seen from this figure, that the performance of the cut tends to drop for $\chi^2 \lesssim 10$. We therefore choose a value of $\chi^2 < 8$ which is a good compromise between the sample purity and available MC statistics.

6.2 Final Selection

In order to increase the sensitivity of the search, we extend the preselection discussed in Chapter 5, with cut on $\chi^2 < 8$. The χ^2 quantity is an indication of how $t\bar{t}$-like an event is: $t\bar{t}$ events are expected to have low values of χ^2 whereas non-$t\bar{t}$ background events have large χ^2 values. We estimated that less than 1% of the QCD multijet background is left and 50% of the signal is kept after the final selection.

In addition, we split events into two sub-channels depending on the number of found b-tagged jets: 0-btag and $\geq$ 1-btag. The 0-btag channel holds events with zero b-tagged jets, and $\geq 1-$ btag channel contains events with at least one b-tagged jet.



(a) (b)

Figure 38: Signal over $W \to l\nu$ background (a) and signal over square root of sum of the signal and background for all background samples (b). The signal Z' boson samples are normalized to 1 pb.

The number of events left after the final selection is summarized in Table V. The $t\bar{t}$ entry includes contributions from dileptonic $t\bar{t}$ ($t\bar{t} \rightarrow (bl\nu)(\bar{b}l'\nu)$, where $l : e, \mu, \tau$) and hadronic ($t\bar{t} \rightarrow (q\bar{q}'b)(\bar{b}q''\bar{q}''')$) decay channels. The Z' boson samples are normalized to 1 pb cross section. The total yields of the simulated samples are normalized to data using the scale factors derived in the maximum likelihood fit of a $M_{t\bar{t}}$ as discussed in Chapter 8.

We compared data and background after the event selection and resonance reconstruction with control plots. The results are shown in Figure 39 to Figure 43. All the distributions are normalized to data using the scales obtained from the maximum likelihood procedure described in Chapter 8. The MC uncertainty band includes all the systematic errors considered in the analysis as listed in Chapter 7. A good agreement between the data and background is observed.

The jets multiplicity plot has a large fraction of the signal Z' boson in the two jets bin with very little background. The fraction of signal in that bin increases with larger resonance mass considered. As can be seen, there is more 3 TeV signal events than 2 TeV. This gives large signal over backround ratio in this jets multiplicity bin to our analysis.

Sample	$\mathbf{N_{b-tag} = 0}$	$\mathbf{N_{b-tag} \geq 1}$
Z', $M = 1$ TeV	17.1	36.5
Z', $M = 1.5$ TeV	44.7	55.4
Z', $M = 2$ TeV	62.1	52.8
Z', $M = 3$ TeV	57.2	36.9
Single Top	9.3	14.6
$W \to l\nu$	89.4	5.5
$Z/\gamma^* \to l^+l^-$	5.6	0.0
$t\bar{t}$	171.6	335.8
Total Background	276 ± 58	356 ± 50
Data 2011	277	354

TABLE V: Number of expected and observed events in 4.4 fb^{-1} of the data. The Z' boson samples are normalized to 1 pb cross section. The $W \to l\nu$ sample represents merged light and heavy flavor content. The total yield of the simulated samples are scaled with scale factors obtained by maximizing the binned likelihood function used in the statistical evaluation discussed in Chapter 8. The total background yield errors are estimated by accounting for those uncertainties originated from the luminosity, lepton identification, jet energy scale and resolution, b-tagging, and pileup corrections.

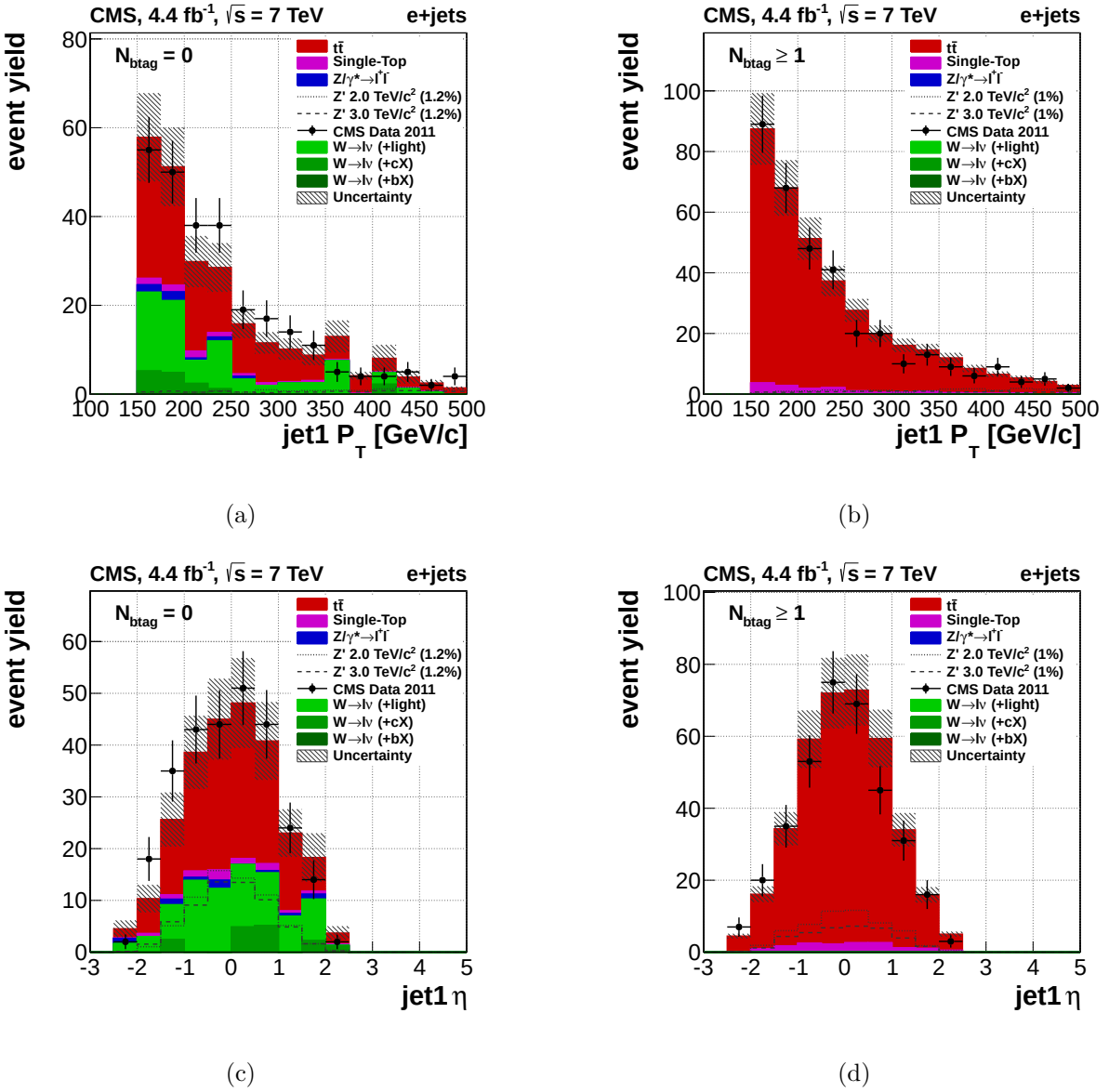


Figure 39: Data background comparison of the leading jet p_T in the 0-btag (a) and $\geq$ 1-btag (b) channels, the leading jet η in the 0-btag (c) and $\geq$ 1-btag (d) channels. All distributions are normalized to the data using the scales obtained from the maximum likelihood procedure. The signal Z' boson samples are normalized to 1 pb. The MC uncertainty band includes all the systematic errors considered in the analysis.

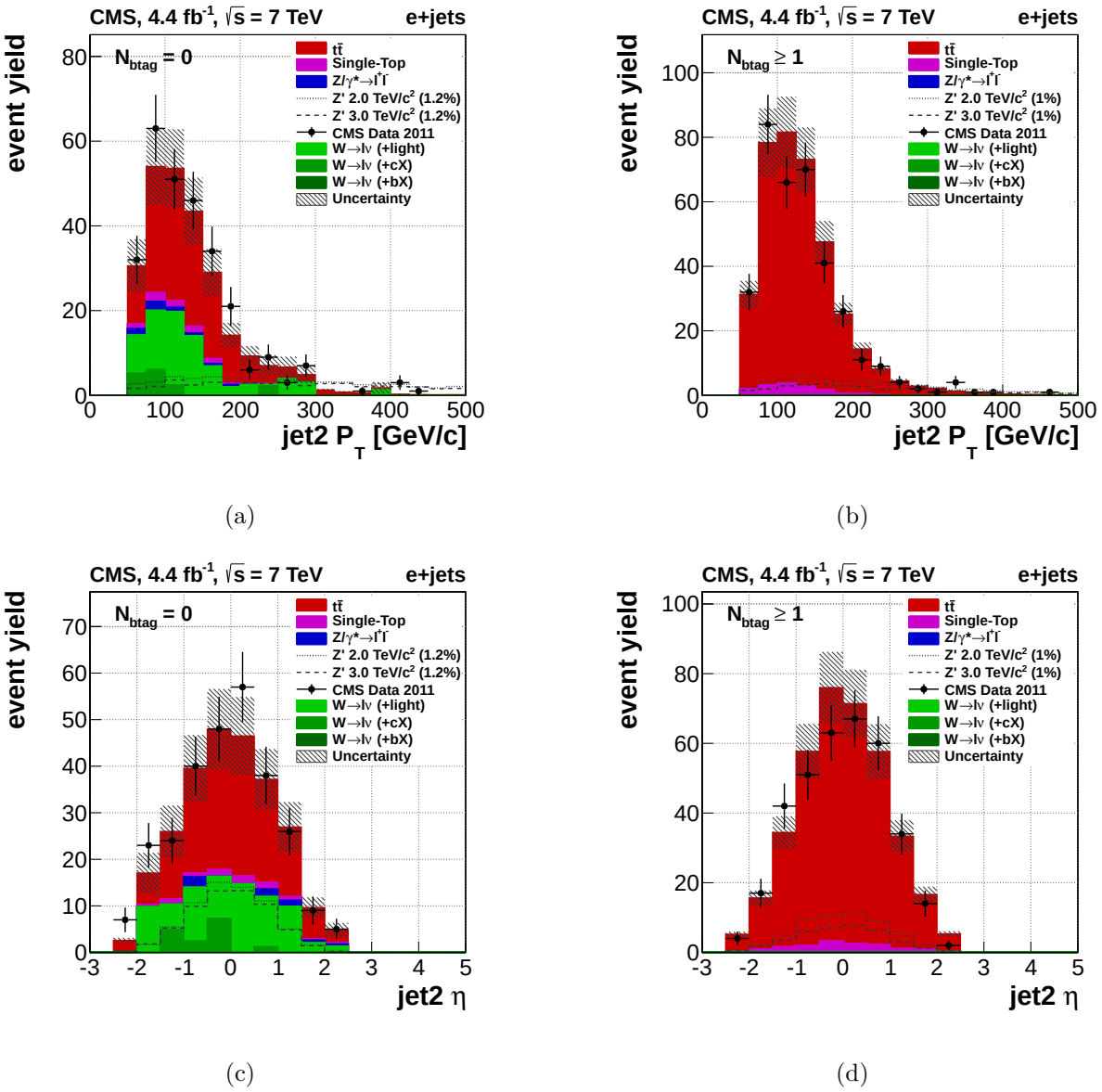


Figure 40: Data background comparison of the second leading jet p_T in the 0-btag (a) and $\geq$ 1-btag (b) channels, the second leading jet η in the 0-btag (c) and $\geq$ 1-btag (d) channels. All distributions are normalized to the data using the scales obtained from the maximum likelihood procedure. The signal Z' boson samples are normalized to 1 pb. The MC uncertainty band includes all the systematic errors considered in the analysis.

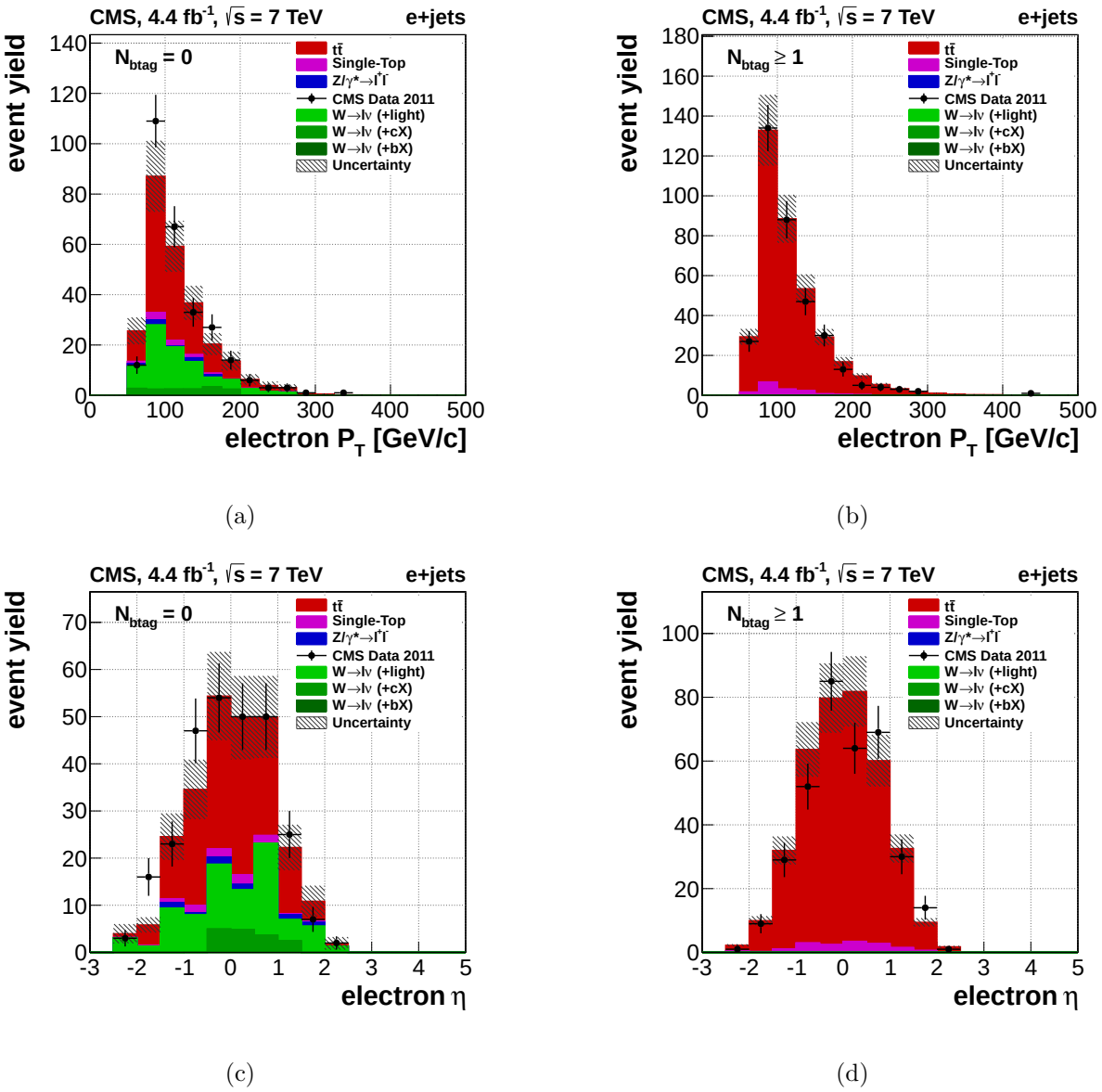


Figure 41: Data background comparison of the electron p_T in the 0-btag (a) and $\geq$ 1-btag (b) channels, the electron η in the 0-btag (c) and $\geq$ 1-btag (d) channels. All distributions are normalized to the data using the scales obtained from the maximum likelihood procedure. The MC uncertainty band includes all the systematic errors considered in the analysis.

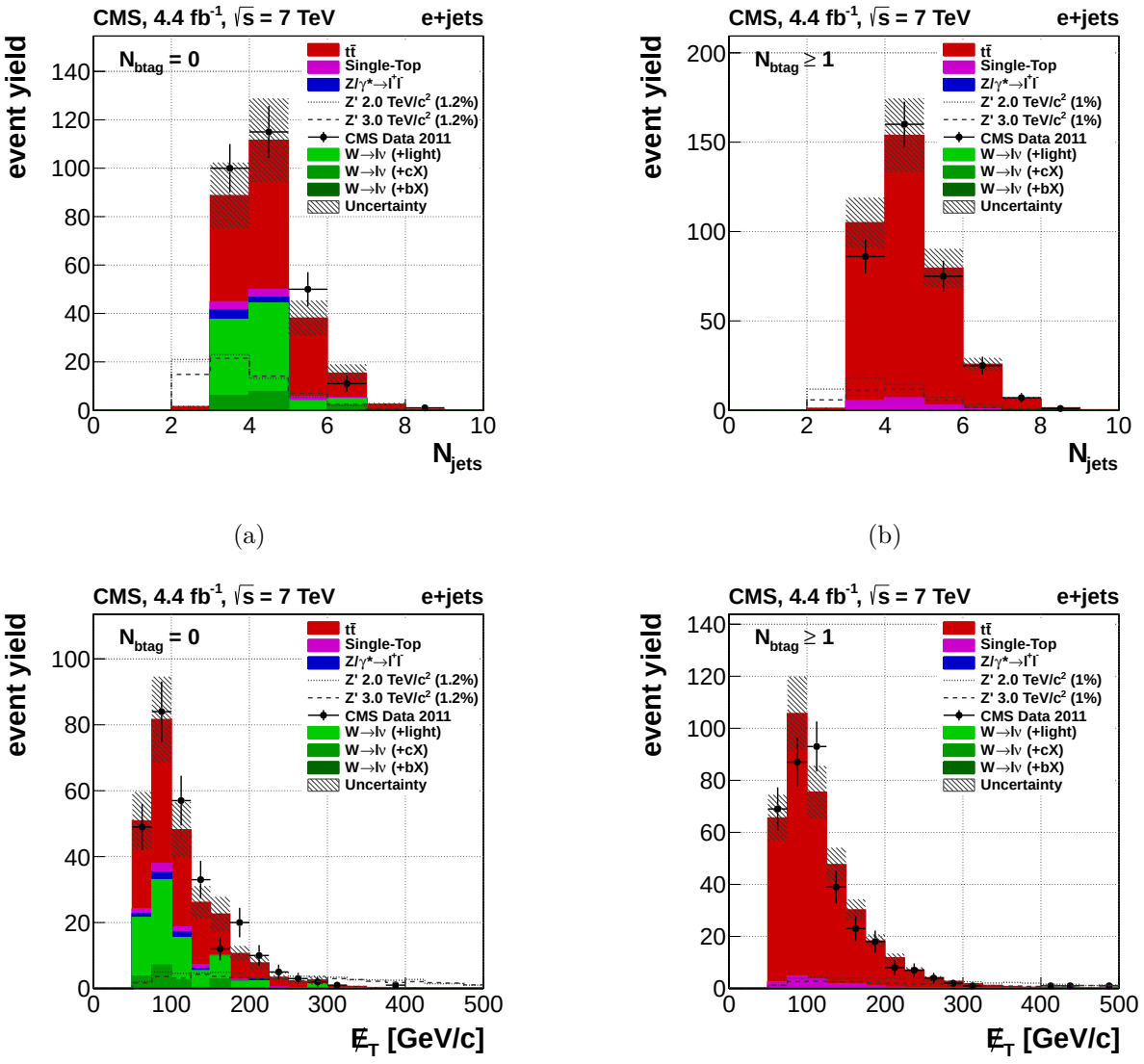


Figure 42: Data background comparison of the number of jets in the 0-btag (a) and $\geq$ 1-btag (b) channels, the $\not\!\!E_T$ in the 0-btag (c) and $\geq$ 1-btag (d) channels. All distributions are normalized to the data using the scales obtained from the maximum likelihood procedure. The signal Z' boson samples are normalized to 1 pb. The MC uncertainty band includes all the systematic errors considered in the analysis.

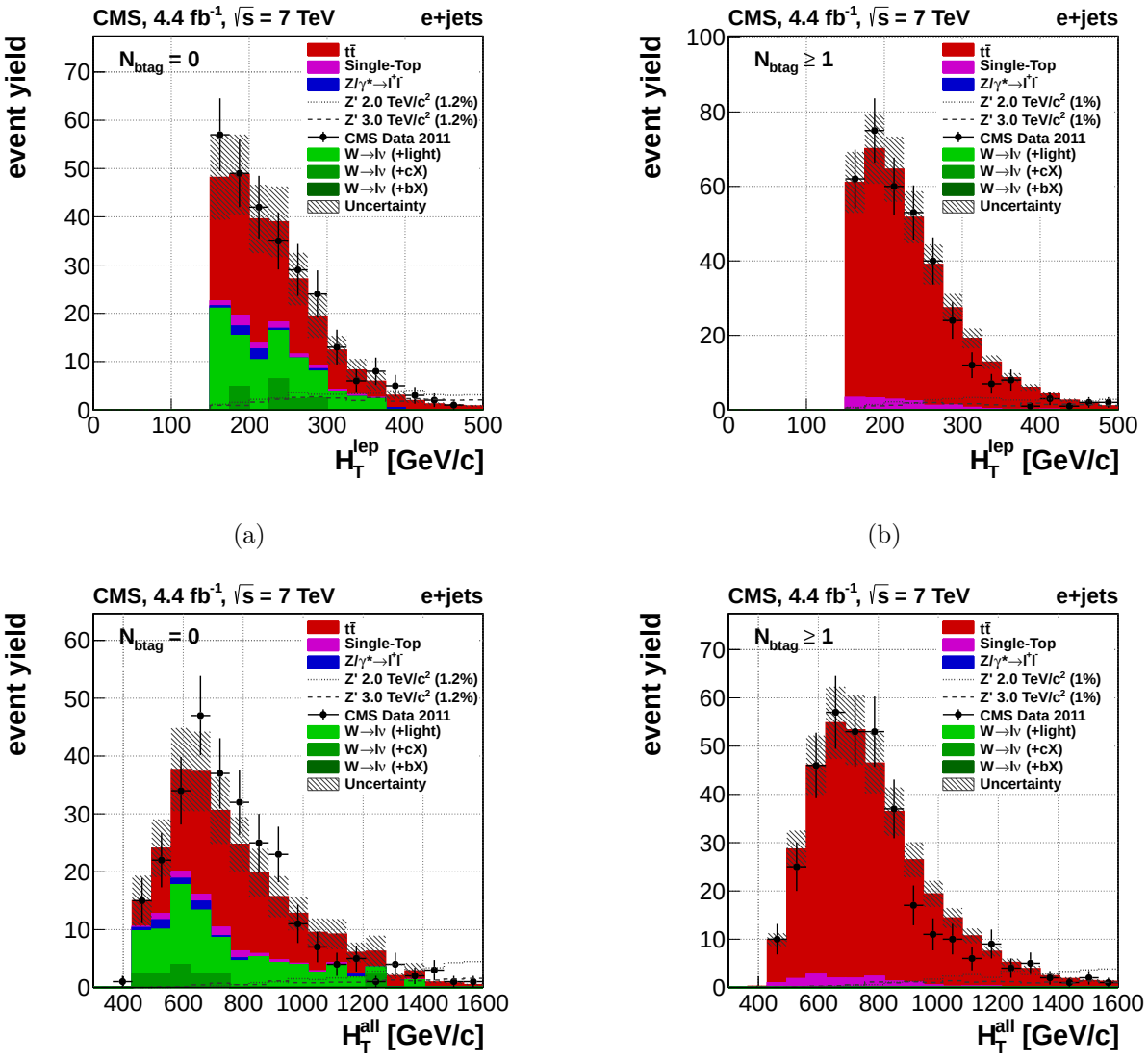


Figure 43: Data background comparison of the H_T^{lep} in the 0-btag (a) and $\geq$ 1-btag (b) channels, the H_T^{all} in the 0-btag (c) and $\geq$ 1-btag (d) channels. All distributions are normalized to the data using the scales obtained from the maximum likelihood procedure. The signal Z' boson samples are normalized to 1 pb. The MC uncertainty band includes all the systematic errors considered in the analysis.

CHAPTER 7

SYSTEMATIC UNCERTAINTIES

This analysis depends on both the normalization and shape of the reconstructed $M_{t\bar{t}}$ distribution. We therefore investigated sources of systematic uncertainties that affect only the normalization or the normalization and shape of the $M_{t\bar{t}}$ distributions. The list of all considered uncertainties in the analysis is given in Table VI.

7.1 Normalization Uncertainties

To investigate the systematic uncertainties due to change in the normalization we vary the cross sections for $t\bar{t}$ by 15% (67), single top by 50%, W_{light}+jets by 50%, and $W_{\text{heavyflavor}}$+jets and Z+jets by 100%. In addition, we vary the luminosity by 2.2% and assume a conservative 5% uncertainty on the lepton trigger and electron identification uncertainty.

7.2 Shape Uncertainties

We consider the following sources of the systematic uncertainties which affect the shape of the reconstructed resonance mass templates in the statistical evaluation:

- Uncertainties in the Jet Energy Scale are of the order of a few percent and vary as a function of the jet p_T and η. We vary the jet energy scale by $\pm\sigma$. The variation in the jet energies are propagated to the $\not\!\!E_T$.

Source	**Uncertainty**	**Type**
Cross Sections		RATE
$t\bar{t}$	15%	
Single-Top	50%	
$W \to l\nu$ (+light)	50%	
$W \to l\nu$ (+bX, +cX)	100%	
$Z/\gamma^* \to l^+l^-$	100%	
Other		RATE
Luminosity	2.2%	
Trigger and Electron ID	5%	
Scale for $W \to l\nu$	$2Q^2$ and $0.5Q^2$	
Matching for $W \to l\nu$	2 and 0.5 x default	
		RATE + SHAPE
Jet Energy Scale	$\pm\sigma(p_T, \eta)$	
Jet Energy Resolution	$\pm\sigma(\eta)$	
b-tagging	$\pm\sigma(p_T, \eta)$	
Mistag Rate	$\pm\sigma(p_T, \eta)$	
Pileup	$\pm\sigma$	
Scale for $t\bar{t}$	$2Q^2$ and $0.5Q^2$	
Matching for $t\bar{t}$	2 and 0.5 x default	

TABLE VI: Systematic uncertainties.

- Uncertainties in the Jet Energy Resolution are 10% worse in data than in Monte-Carlo. To account for this, we vary the jet energy resolution, depending on the jet η, by $\pm\sigma$. The variation in the jet energy resolution is propagated to the $\not\!\!E_T$.
- The uncertainty of the b-tagging scale factor is a function of the jet p_T and is derived from the analysis of the μ+jets sample. For the jets with $p_T > 670$ GeV, the value of SF_b for 670 GeV is used and the quoted uncertainty is doubled.
- The uncertainty for the mistag rate is a function of the jet p_T. For the jets with $p_T > 670$ GeV we use the SF_l value at 670 GeV twice the quoted uncertainty.
- We varied the measured minimum bias cross section with $\pm\sigma$ of its uncertainty (73.5 ± 5.5 mb) to produce different expected pileup distributions.
- We studied the effect due to the $Q^2 = M^2(t) + \sum p_{T,jet}^2$ scale uncertainty. We use samples with two different Q^2 scales,scale up and scale down, for the W+jets and $t\bar{t}$ samples. These samples were re-generated with MadGraph by scaling the event-by-event Q^2 value by 0.5 (2) for the scale down (scale up) sample. We change the scale coherently in both samples. In the case of W+jets, there is not enough statistics to properly qualify the effect of these uncertainties on the shape of the $M_{t\bar{t}}$ distribution. Therefore we assigned only a normalization uncertainty based on the relative yield changes between the nominal and systematically shifted samples before the χ^2 cut and b-tagging categorization.
- We studied the effect due to the uncertainty in the extra hard parton radiation by varying the jet matching threshold for the W+jets and $t\bar{t}$ samples with a factor 0.5 and 2 from its default value. The default values are 30 GeV for $t\bar{t}$ and 15 GeV for W+jets. We change

the matching threshold coherently in both samples. In the case of W+jets, there is not enough statistics to properly quantify the effect of these uncertainties on the shape of the $M_{t\bar{t}}$ distribution. Therefore we assigned only a normalization uncertainty based on the relative yield changes between the nominal and systematically shifted samples before the χ^2 cut and b-tagging categorization.

Figure 44 shows variation after a maximum likelihood fit for nuisance parameters associated to systematic uncertainties that affect the sample normalization and shape in units of standard deviation. This fit is used to normalize all the control plots after the full selection. The variation is shown for the combination with boosted muon+jets channel. The picked values of nuisance parameters are within one standard deviation.

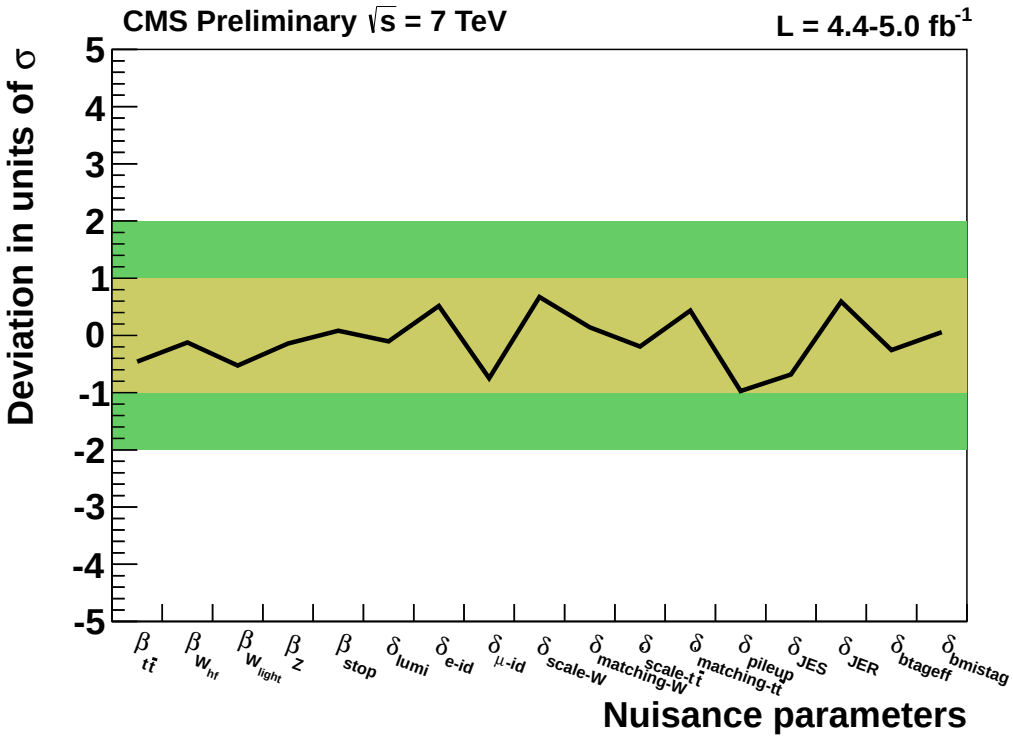


Figure 44: Variation after maximum likelihood fit for nuisance parameters associated to those systematic uncertainties that affect both sample normalization and shape in units of standard deviation.

CHAPTER 8

STATISTICAL ANALYSIS

We use a CLs statistical method to extract 95% Confidence Level upper limits on the production cross section of heavy resonances that decay to $t\bar{t}$. For the CLs method we use a profile likelihood ratio as a test statistic. We use the $t\bar{t}$ invariant mass spectrum in the 0-btag and $\geq$ 1-btag channels for the template and THETA framework (68) for the implementation of this method.

8.1 Likelihood

The likelihood function is a product of the Poisson probability for each bin in the template:

$$\mathcal{L}(\beta_k|\text{data}) = \prod_{i=1}^{N_{bins}} \frac{\mu_i^{n_i} \cdot e^{-\mu_i}}{n_i!} \tag{8.1}$$

where n_i is the number of observed events in the bin i, e.g. data, and μ_i is the Poisson mean for the same bin given by:

$$\mu_i = \sum_k \beta_k \cdot T_{k,i} \tag{8.2}$$

where k denotes all of the considered processes, and $T_{k,i}$ is the content of the bin i of the template for the process k. These templates are scaled such that the bin entries are predicted Poisson mean according to the assumed cross section and luminosity. The model parameters β_k are the Poisson mean for the process k normalized to the prediction from the simulation. For

the signal samples, we scale the templates to the arbitrary cross section of 1 pb which allows the interpretation of the $\beta_{Z'}$ as the signal production cross section times branching ratio to $t\bar{t}$ in pb. The MC background templates are scaled according to theory cross section and luminosity.

The predicted yield for the bin i, μ_i, can be modified by systematic uncertainties, described in Chapter 7. For each systematic uncertainty u, a nuisance parameter δ_u is introduced with a Gaussian prior around 0 with width 1, which modifies μ_i using a template interpolation technique interpolating the nominal template, not affected by the uncertainty, and the shifted templates, which are obtained by applying a $\pm 1\sigma$ systematic shift to the MC, and re-rederiving the templates. The interpolation is defined such that a value of $\delta_u = 0$ reproduces the nominal template, and values of $\delta_u = \pm 1$ lead to the shifted templates. The normalization uncertainties correspond to β_k parameters, which have log-normal priors with widths given in Table VI. In the toy experiments performed within the CLs method all nuisance parameters are drawn from their respective priors. We use a flat prior for the signal samples.

8.2 CLs method

To set the exclusion limits on the BSM signal hypothesis, we define a test statistic $q_{\beta_{Z'}} = q_\alpha$, which depends on the hypothesized signal rate $\alpha = \beta_{Z'}$ (69). The test statistic is:

$$q_\alpha = -2 \ln \frac{\mathcal{L}(\hat{\hat{\delta}}_\alpha, \alpha | \text{data})}{\mathcal{L}(\hat{\delta}, \hat{\alpha} | \text{data})} \tag{8.3}$$

where $\hat{\hat{\delta}}_\alpha$, $\hat{\alpha}$, and $\hat{\delta}$ are the values of the parameters δ and α that maximize the likelihoods in the numerator and denominator, and the subscript α in $\hat{\hat{\delta}}_\alpha$, indicates that the maximization of

the likelihood in the numerator is done under the hypothesis of a signal of strength α. One-sided limits on the BSM resonance production rate are obtained by constraining $\hat{\alpha} \leq \alpha$ in the denominator.

We use a modified frequentist construction CL_s (70; 71) to calculate the 95% C.L. upper limits on the BSM resonance to $t\bar{t}$ cross section times branching ratio. We define two tail probabilities associated with the observed data: a probability to obtain a value of the test statistic q_α larger than the observed value q_α^{obs} for the signal + background:

$$CL_{s+b} = P(q_\alpha \geq q_\alpha^{\text{obs}} | \delta, \alpha) \tag{8.4}$$

For the background-only hypotheses:

$$CL_b = P(q_\alpha \geq q_\alpha^{\text{obs}} | \delta, \alpha = 0) \tag{8.5}$$

The ratio of the two gives:

$$CL_s = \frac{CL_{s+b}}{CL_b} \tag{8.6}$$

We adjust α until we reach $CL_s = 0.05$ to quote the upper limit on α at the 95% Confidence Level.

The expected upper limits are calculated using background-only pseudo-experiments ($\alpha = 0$) and calculating the upper limit for each pseudo-experiment. The expected limit is given by the

median of the distribution of upper limits, and the central 68% and 95% give the ±1 and ±2 standard deviation exclusions.

8.3 Background Statistics

Figure 45 shows the $M_{t\bar{t}}$ distribution in logarithmic scale for events with zero and at least one b-tagged jet. The main background is $t\bar{t}$ and it has a steeply falling spectrum in both channels. Starting from the invariant masses around 2 TeV, the number of available simulated events is not enough to have a good estimation of the uncertainty in the background expectation. Also, the background and data content of the first bins in the distribution are zero due to the selection thresholds.

To ensure a proper background modeling in the entire $M_{t\bar{t}}$ range, we merge bins in the $M_{t\bar{t}}$ distribution by requiring a minimum number of background events per bin. The bins are chosen such that the statistical uncertainty on the number of expected background events due to the limited number of simulated events is not worse than 30% in all channels. Figure 45 shows a data background comparison of $M_{t\bar{t}}$ after the re-binning. We use the re-binned histograms to compute the likelihood and extract the limits.

The uncertainty due to a finite size of the simulated samples is taken into account by a method inspired by Barlow and Beeston (72), but is modified to use only one additional nuisance parameter per bin, sometimes called Barlow-Beeston lite (73).

In this method, the yield μ_i in each bin is modified by adding a nuisance parameter ν_i, which has a Gaussian prior around zero with a width corresponding to the uncertainty estimated from

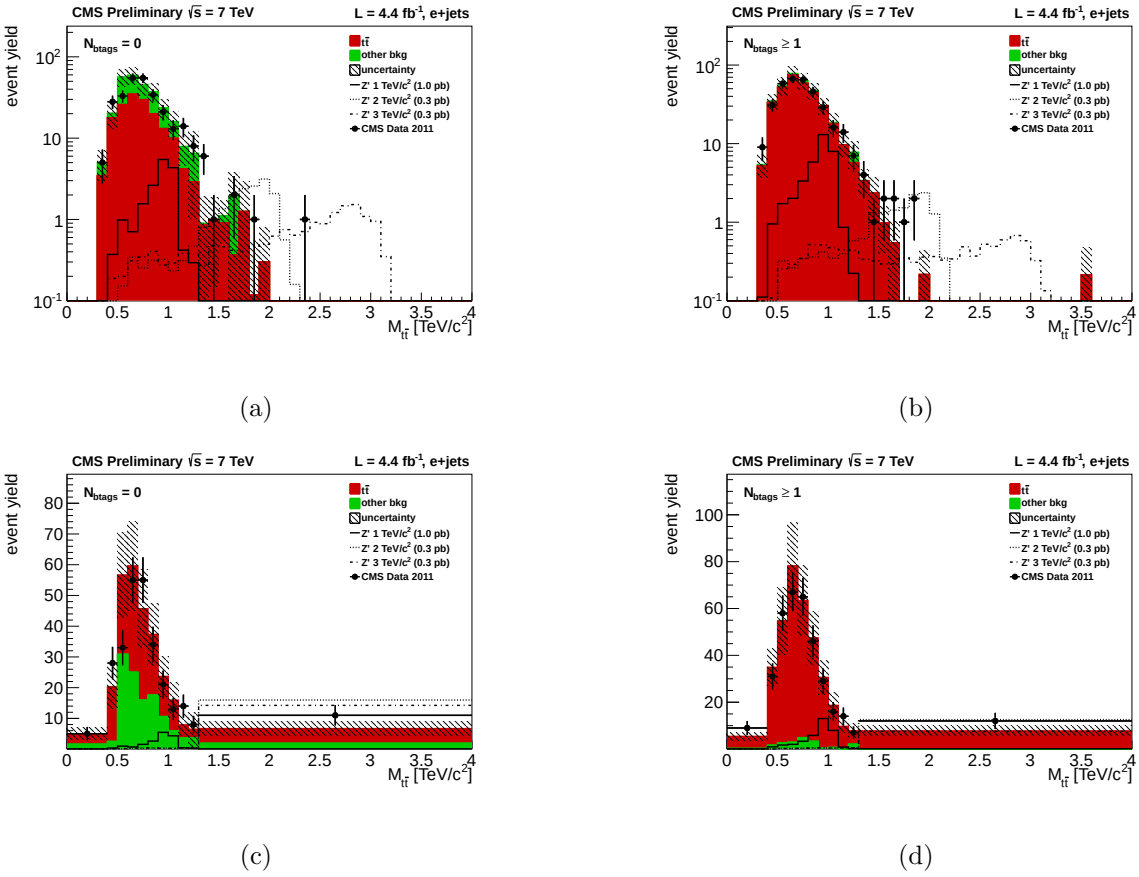


Figure 45: Data background comparison of $M_{t\bar{t}}$ before re-bin in the 0-btag (a) and $\geq$ 1-btag (b) channels, and after the re-bin in the 0-btag (c) and $\geq$ 1-btag (d) channels. The background yields are obtained from a maximum likelihood fit. The hatched bands in the plots correspond to yield changes when applying the following systematics: JES, JER, b-tagging uncertainties and pielup. The signal Z' boson samples are normalized to 1 pb for 1 TeV mass, and 0.3 pb for 2 and 3 TeV masses.

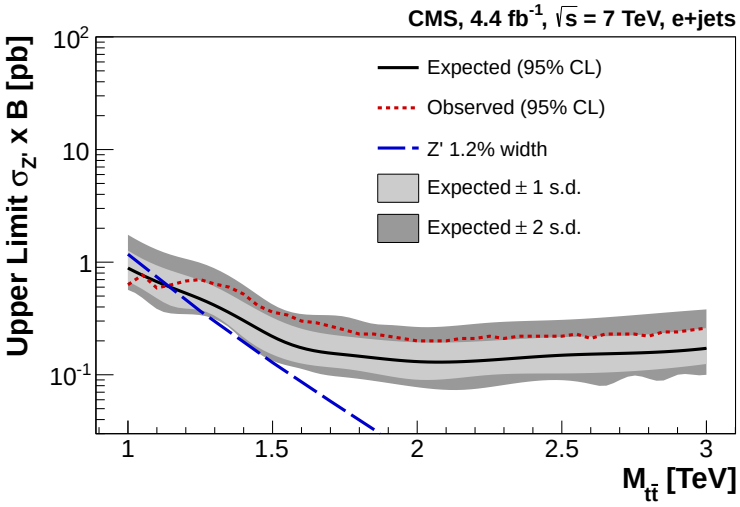
the number of MC events in this bin. This introduces one additional nuisance parameter per bin.

The numerical maximization of the likelihood function performed in the test statistics evaluation becomes unstable with too many nuisance parameters. However, the test statistic evaluation actually requires a profile likelihood function in which the newly introduced nuisance parameters ν_i have been maximized out:

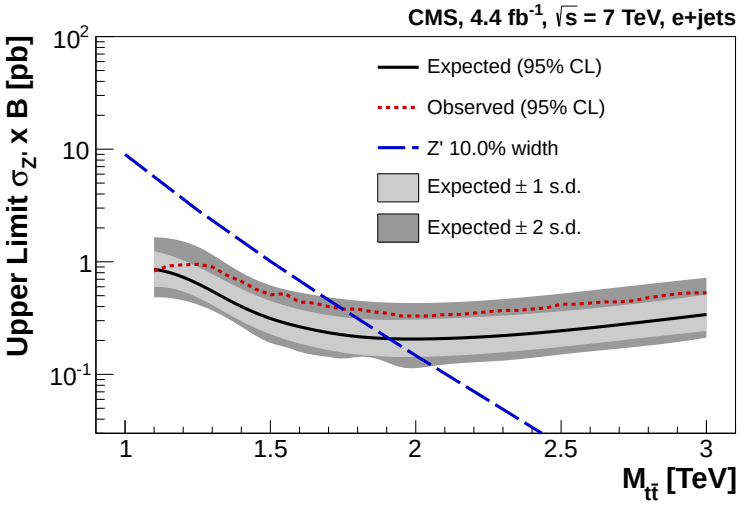
$$\mathcal{L}_p(\beta_k, \delta_u|\text{data}) = \max_{\nu_i} \mathcal{L}(\beta_k, \delta_u, \nu_i|\text{data}) \tag{8.7}$$

The maximization with respect to ν_i can be performed analytically. For the calculation of the test statistic, $\mathcal{L}_p$ is used which accounts for the uncertainty for limited simulated sample size. The extracted 95% C.L. upper limits on the product of the BSM resonance production cross section and the branching fraction with $\Gamma_{Z'}/M_{Z'} = 1.2\%$ and 10% as a function of the invariant mass is shown in Figure 46.

After analyzing 4.4 fb^{-1} of CMS data from 2011 in the electron+jets channel, we found no evidence of such massive BSM resonances and exclude at the 95% C.L. topcolor Z' bosons with the width $\Gamma_{Z'}/M_{Z'} = 1.2\%$ (10%) for masses between 1.0 and 1.15 TeV (1.0 and 1.75 TeV). A subpicobarn limit is set for both resonances at masses larger than 2 TeV.



(a)



(b)

Figure 46: The 95% C.L. upper limits on the product of the production cross section $\sigma_{Z'}$ and the branching fraction B of hypothesized resonances that decay into $t\bar{t}$ as a function of the invariant mass of the resonance. The Z' production with $\Gamma_{Z'}/M_{Z'} = 1.2\%$ (a) and 10% (b) compared to theoretical predictions based on (1). The ± 1 and ± 2 s.d. excursions from the expected limits are also shown.

CHAPTER 9

COMBINATION AND RESULTS

We combine the electron+jets results from the previous chapter with the corresponding results in the muon+jets channel (74). The muon+jets channel has a selection similar to the electron+jets with minor modifications in the event selection. The resonance reconstruction uses the same χ^2 technique as in the electron+jets channel. Figure 47 shows a data background comparison of the reconstructed mass spectrum in the muon+jets channel. The distribution is broader than in the electron+jets channel due to different kinematic selection. To combine the electron and muon channels, we used the reconstructed resonance mass distributions from both channels as input for the statistical evaluation and limit setting procedure with the same systematic uncertainties as described in Chapter 7.

We also combined the electron and muon boosted results with a complementary analysis optimized for resonances with masses smaller than 1 TeV (26; 28). This threshold analysis uses the lepton+jets final state but requires isolated leptons, electron or muon, and at least 3 jets. Figure 48 shows the reconstructed $M_{t\bar{t}}$ distributions for the threshold analysis. A good agreement is observed between the data and SM predictions.

Both analyses, threshold and boosted, share the same events in the selection and resonance reconstruction. Therefore the reconstructed resonance mass distributions are correlated between the searches and can not be used as independent inputs by `THETA` for the statistical evaluation. We extracted the limits independently for each analysis and define a transition be-

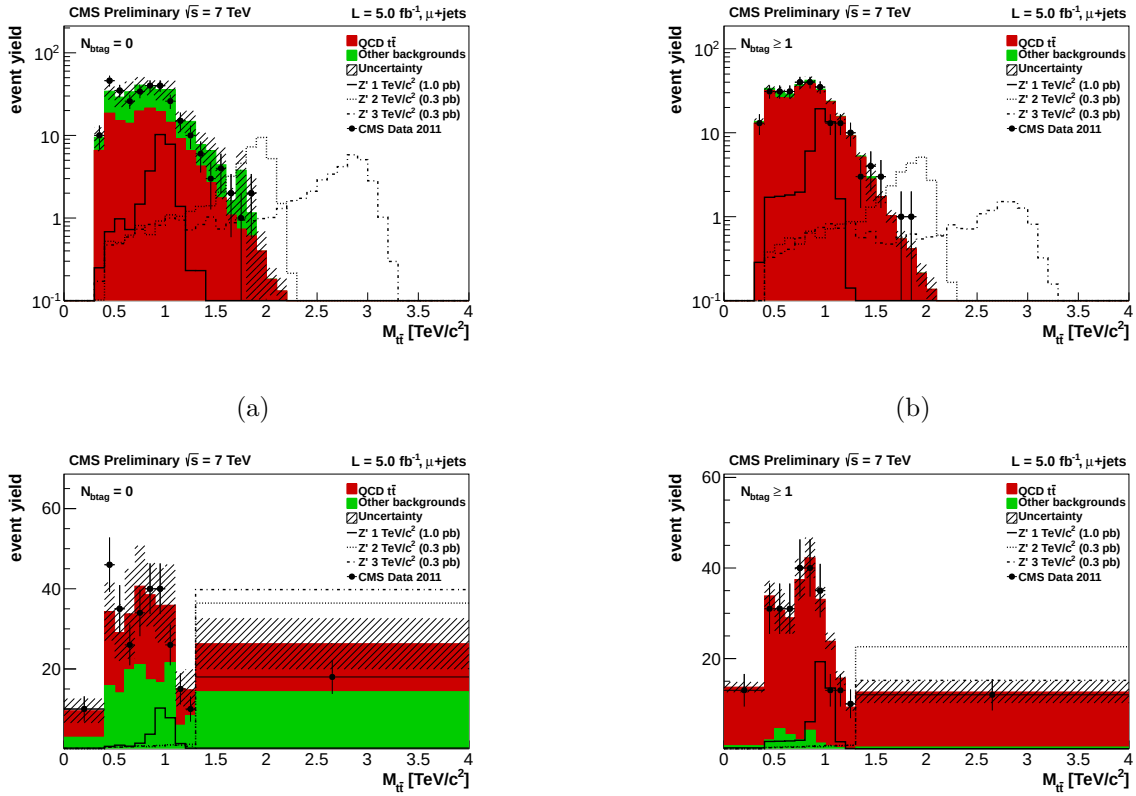


Figure 47: Data background comparison of the $t\bar{t}$ reconstructed invariant mass in the muon+jets channel before the re-bin in the 0 b-tagged (a) and ≥ 1 b-tagged (b) samples, and after the re-bin in the 0 b-tagged (c) and ≥ 1 b-tagged (d) samples. The background yields are obtained from the maximum likelihood fit as described in the text. A cross section of 1.0 pb (0.3 pb) is used for the Normalization of the $M_{Z'} = 1$ TeV ($M_{Z'} = 2$ TeV and $M_{Z'} = 3$ TeV) samples. The hatched bands in the plots correspond to yield changes when applying the following systematics: JEC, JER, b-tagging uncertainties and pileup.

tween the threshold and the boosted analyses, chosen based on the sensitivity of the expected limit.

Figure 49 and Figure 50 show the expected and observed 95% C.L. upper limits for the product of the production cross section and branching fraction of the hypothesized resonances that decay into $t\bar{t}$ as a function of the invariant mass of the resonance. The blue dashed lines indicate the values predicted by various models for the new physics processes. No evidence of such massive resonances is observed and the 95% C.L. upper limits on the product of the production cross section and the branching fraction are set.

The expected mass exclusion region for the topcolor Z' with $\Gamma_{Z'}/M_{Z'} = 1.2\%$ is $M_{Z'} < 1.53$ TeV, the observed exclusion is $M_{Z'} < 1.49$ TeV. For wide resonances with $\Gamma_{Z'}/M_{Z'} = 10\%$, the exclusion region is $M_{Z'} < 2.04$ TeV for both the expected and observed limits. For the Kaluza-Klein excitation of a gluon the exclusion mass region is $M_{g_{KK}} < 1.82$ TeV for both the expected and observed limits.

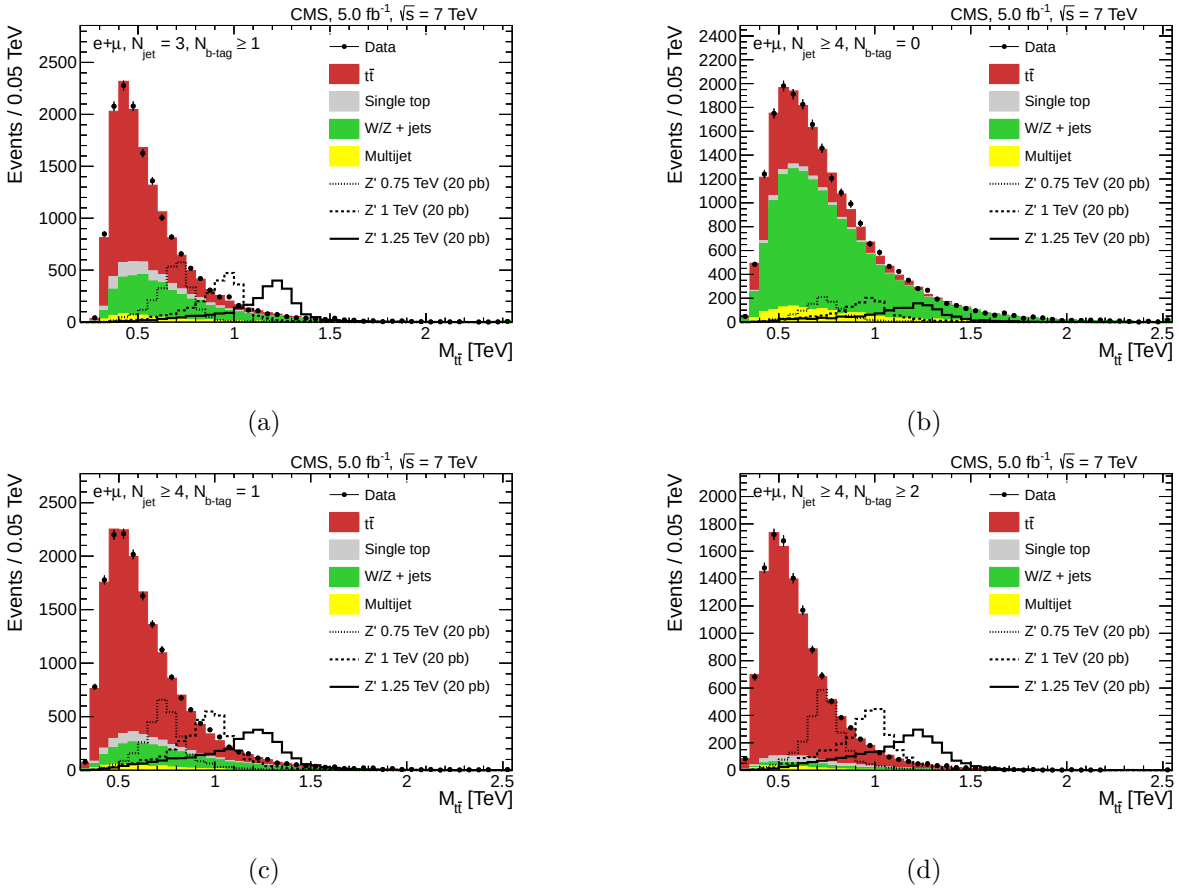
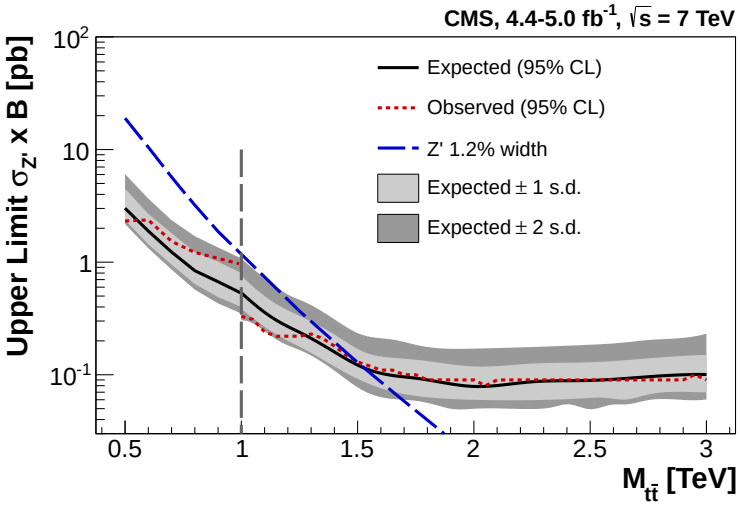
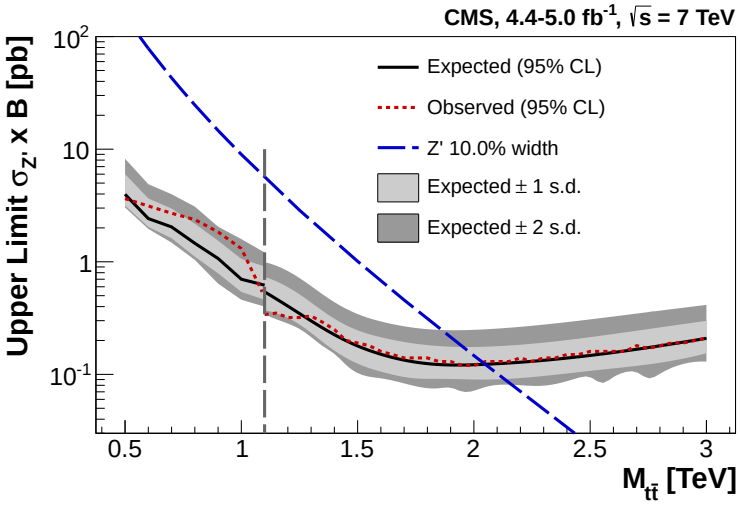


Figure 48: A comparison of the reconstructed $M_{t\bar{t}}$ in the data and SM predictions for the threshold analysis in different jet and b-tags multiplicity channels: (a) $N_{jets} = 3$, $N_{b-tag} \geq 1$, (b) $N_{jets} \geq 4$, $N_{b-tag} = 0$, (c) $N_{jets} \geq 4$, $N_{b-tag} = 1$, (d) $N_{jets} \geq 4$, $N_{b-tag} \geq 2$. The expected signal contributions for narrow-width topcolor Z' models at different masses are also shown. For clarity, a cross section times branching fraction of 20 pb is used for the normalization of the Z' samples.

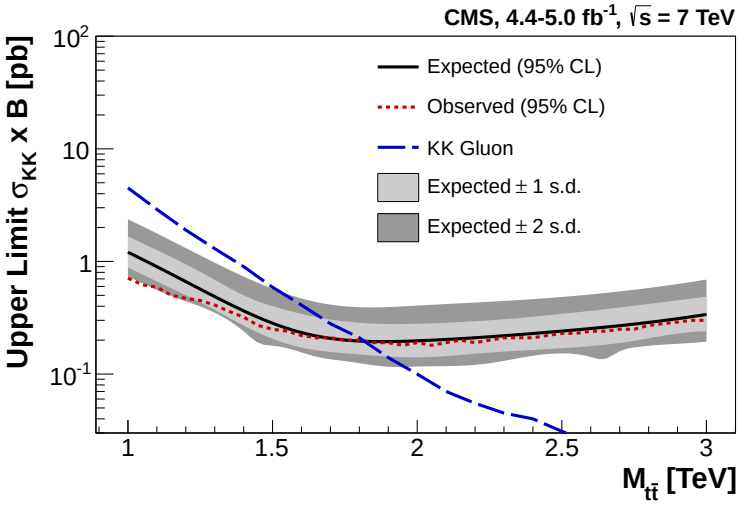


(a)



(b)

Figure 49: The 95% C.L. upper limits on the product of the production cross section $\sigma_{Z'}$ and the branching fraction B of hypothesized resonances that decay into $t\bar{t}$ as a function of the invariant mass of the resonance. The Z' production with with $\Gamma_{Z'}/M_{Z'} = 1.2\%$ (a) and 10% (b) compared to theoretical predictions based on (1). The ± 1 and ± 2 s.d. excursions from the expected limits are also shown. The vertical dashed line indicates the transition between the threshold and the boosted analyses, chosen based on the sensitivity of the expected limit.



(a)

Figure 50: The 95% C.L. upper limits on the product of the production cross section σ_{KK} and the branching fraction B of Kaluza-Klein excitation of gluon production from (2) compared to the theoretical prediction of that model. The ± 1 and ± 2 s.d. excursions from the expected limits are also shown.

CHAPTER 10

SUMMARY

Results from a model-independent search for the production of heavy resonances decaying to $t\bar{t}$ are presented. The analyzed data sample corresponds to the integrated luminosity of 4.4 fb^{-1} recorded in 2011 by the CMS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC and uses events in the electron plus jets final state. No evidence of a massive resonances is found and the limits on the production cross section times branching fraction are set: topcolor Z' bosons are excluded at the 95% C.L. with the width $\Gamma_{Z'}/M_{Z'} = 1.2\%$ (10%) for masses between 1.0 and 1.15 TeV (1.0 and 1.75 TeV).

We combined the results with the muon+jets boosted channel and lepton+jets threshold analysis. The combination didn't find evidence of BSM resonances. Topcolor Z' bosons with the mass width $\Gamma_{Z'}/M_{Z'} = 1.2\%$ (10%) are excluded at 95% C.L. for masses below 1.49 (2.04) TeV. An upper limit of 0.3 (1.3) pb is set on the production cross section times branching fraction for the resonance mass of 1 TeV.

The Kaluza-Klein excitations of a gluon with masses below 1.82 TeV at 95% C.L. in the Randall-Sundrum model are excluded in the combined lepton+jets search. An upper limit of 0.7 pb is set on the production cross section times branching fraction for the resonance mass of 1 TeV.

Figure 51 compares the 95% C.L. expected upper limits on the product of the production cross section $\sigma_{Z'}$ and the branching fraction B of hypothesized resonances that decay into $t\bar{t}$

as a function of the invariant mass of the resonance between different BSM searches. The results for the Z' production with $\Gamma_{Z'}/M_{Z'} = 1.2\%$ are shown. The searches cover all $t\bar{t}$ system final states: dilepton analysis (28), lepton+jets threshold (26) and boosted searches (74), and all-hadronic analysis (27). As can be seen from the plot, current lepton+jets boosted analysis has outstanding sensitivity and best result to date in the central region of the hypothesized resonance masses $1 - 2$ TeV. With the increase of the LHC integrated luminosity and energy from 5 fb^{-1} at $\sqrt{s} = 7$ TeV in 2011 to 20 fb^{-1} at $\sqrt{s} = 8$ TeV in 2012 the limits on the cross section times branching fraction can be further improved because of the increase of the statistics in the large invariant mass region. We expect to reach subpicobarn region for the masses above 1 TeV. Furthermore, the use of new techniques designed exclusively for the boosted top analyses such as top-tagging can improve the resonance reconstruction and the limits.

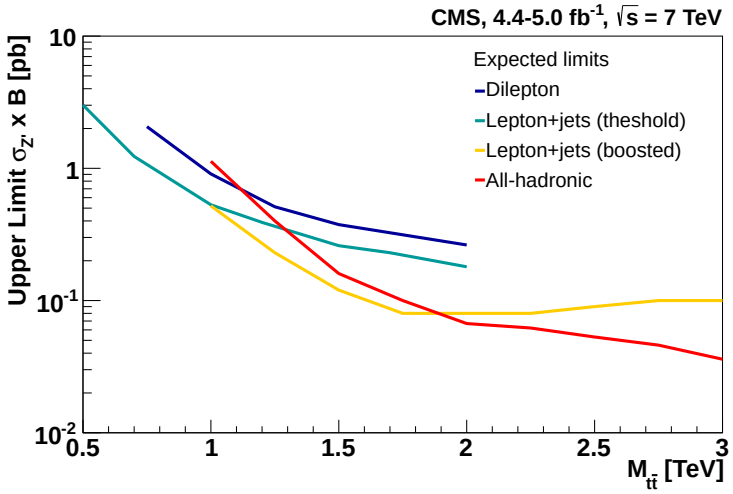


Figure 51: A comparison of the 95% C.L. expected upper limits on the product of the production cross section $\sigma_{Z'}$ and the branching fraction B of hypothesized resonances that decay into $t\bar{t}$ as a function of the invariant mass of the resonance between different BSM searches. The results for the Z' production with $\Gamma_{Z'}/M_{Z'} = 1.2\%$ are shown.

APPENDICES

Appendix A

NEUTRINO RECONSTRUCTION

The neutrino is produced from the semi-leptonic decay of the W boson in current analsysis $W \to l + \nu$. Given the lorentz energy-momentum vectors of each object and assuming the W boson decay on shell, one can write:

$$\mathbf{P}_W = \mathbf{P}_l + \mathbf{P}_\nu \tag{A.1}$$

The equation can be simplified if both sides are squared.

$$\mathbf{P}_W^2 = \mathbf{P}_l^2 + \mathbf{P}_\nu^2 + 2\,\mathbf{P}_l \cdot \mathbf{P}_\nu \tag{A.2a}$$

$$= m_l^2 + m_\nu^2 + 2\,(E_l\,E_\nu - \mathbf{p}_l \cdot \mathbf{p}_\nu) \tag{A.2b}$$

The masses of the decay products are neglected compared to high values of momenta and are dropped from the derivation. Thus, the equation is simplified to:

$$\frac{m_W^2}{2} = E_l \sqrt{p_{T,\nu}^2 + p_{z,\nu}^2} - \mathbf{p}_{T,l} \cdot \mathbf{p}_{T,\nu} - p_{z,l}\,p_{z,\nu} \tag{A.3}$$

where $\mathbf{p}_T$ is the transverse momentum:

$$E^2 = p_T^2 + p_z^2 \tag{A.4}$$

Appendix A (Continued)

Here the mass is dropped. Let:

$$\alpha = \frac{m_W^2}{2} + \mathbf{p}_{T,l} \cdot \mathbf{p}_{T,\nu} \tag{A.5}$$

Re-arrange the terms in the Equation A.3:

$$\alpha + p_{z,l}\, p_{z,\nu} = E_l \sqrt{p_{T,\nu}^2 + p_{z,\nu}^2} \tag{A.6}$$

Square both sides of the equation.

$$\alpha^2 + p_{z,l}^2\, p_{z,\nu}^2 + 2\alpha\, p_{z,l}\, p_{z,\nu} = E_l^2\, (p_{T,\nu}^2 + p_{z,\nu}^2) \tag{A.7}$$

Let:

$$a = p_{z,l}^2 - E_l^2 = -p_{T,l}^2 \tag{A.8a}$$

$$b = \alpha\, p_{z,l} \tag{A.8b}$$

$$c = \alpha^2 - E_l^2\, p_{T,\nu}^2 \tag{A.8c}$$

A quadratic equation for the netrino longitudinal momentum is:

$$a\, p_{z,\nu}^2 + 2\, b\, p_{z,\nu} + c = 0 \tag{A.9}$$

Appendix A (Continued)

The solution is given by:

$$p_{z,\nu} = \frac{\alpha\, p_{z,l}}{p_{T,l}^2} \pm \sqrt{\frac{\alpha^2\, p_{z,l}^2}{p_{T,l}^4} - \frac{E_l^2\, p_{T,\nu}^2 - \alpha^2}{p_{T,l}^2}} \tag{A.10}$$

CITED LITERATURE

1. R. M. Harris and S. Jain, "Cross Sections for Leptophobic Topcolor Z-prime Decaying to Top-Antitop," *Eur.Phys.J.* **C72** (2012) 2072, arXiv:1112.4928 [hep-ph].

2. K. Agashe *et al.*, "LHC Signals from Warped Extra Dimensions," *Phys. Rev. D* **77** (2008) 015003, arXiv:hep-ph/0612015 [hep-ph].

3. S. L. Glashow, "Partial-symmetries of weak interactions," *Nuclear Physics* **22** no. 4, (1961) 579 – 588. http://www.sciencedirect.com/science/article/pii/0029558261904692.

4. **CDF** Collaboration, "Observation of Top Quark Production in $p\bar{p}$ Collisions with the Collider Detector at Fermilab," *Phys. Rev. Lett.* **74** (Apr, 1995) 2626–2631. http://link.aps.org/doi/10.1103/PhysRevLett.74.2626.

5. "Review of Particle Physics, 2012-2013. Review of Particle Properties," *Phys. Rev. D* **86** no. 1, (2012) 010001. http://cds.cern.ch/record/1481544.

6. J. M. Campbell, J. Huston, and W. Stirling, "Hard Interactions of Quarks and Gluons: a Primer for LHC Physics," *Rept.Prog.Phys.* **70** (2007) 89, arXiv:hep-ph/0611148 [hep-ph].

7. D. E. Morrissey, T. Plehn, and T. M. Tait, "Physics Searches at the LHC," *Phys.Rept.* **515** (2012) 1–113, arXiv:0912.3259 [hep-ph].

8. J. L. Rosner, "Prominent Decay Modes of a Leptophobic Z-prime," *Phys. Lett.* **B387** (1996) 113–117, arXiv:hep-ph/9607207.

9. K. R. Lynch, E. H. Simmons, M. Narain, and S. Mrenna, "Finding Z-prime Bosons Coupled Preferentially to the Third Family at LEP and the Tevatron," *Phys.Rev.* **D63** (2001) 035006, arXiv:hep-ph/0007286 [hep-ph].

10. M. S. Carena, A. Daleo, B. A. Dobrescu, and T. M. Tait, "Z-prime Gauge Bosons at the Tevatron," *Phys.Rev.* **D70** (2004) 093009, arXiv:hep-ph/0408098 [hep-ph].

11. C. T. Hill and S. J. Parke, "Top Production: Sensitivity to New Physics," *Phys.Rev.* **D49** (1994) 4454–4462, arXiv:hep-ph/9312324 [hep-ph].

12. P. H. Frampton and S. L. Glashow, "Chiral Color: an Alternative to the Standard Model," *Physics Letters B* **190** no. 1-2, (1987) 157–161. http://www.sciencedirect.com/science/article/B6TVN-472JMJG-P5/2/90a0c6335472f2ad2b10a609415c85b5.

13. R. M. Godbole and D. Choudhury, "Nonstandard, Strongly Interacting Spin One t anti-t Resonances," arXiv:0810.3635 [hep-ph].

14. D. Dicus, A. Stange, and S. Willenbrock, "Higgs Decay to Top Quarks at Hadron Colliders," *Phys.Lett.* **B333** (1994) 126–131, arXiv:hep-ph/9404359 [hep-ph].

15. L. Randall and R. Sundrum, "A Large Mass Hierarchy from a Small Extra Dimension," *Phys.Rev.Lett.* **83** (1999) 3370–3373, arXiv:hep-ph/9905221 [hep-ph].

16. N. Arkani-Hamed, S. Dimopoulos, and G. Dvali, "The Hierarchy Problem and New Dimensions at a Millimeter," *Phys.Lett.* **B429** (1998) 263–272, arXiv:hep-ph/9803315 [hep-ph].

17. C. T. Hill, "Topcolor: Top Quark Condensation in a Gauge Extension of the Standard Model," *Phys. Lett. B* **266** (1991) 419. Updates in http://arxiv.org/abs/hep-ph/9911288.

18. **CDF** Collaboration, T. Aaltonen *et al.*, "Limits on the Production of Narrow $t\bar{t}$ resonances in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," *Phys.Rev.* **D77** (2008) 051102, arXiv:0710.5335 [hep-ex].

19. **CDF** Collaboration, T. Aaltonen *et al.*, "Search for Resonant $t\bar{t}$ Production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," *Phys.Rev.Lett.* **100** (2008) 231801, arXiv:0709.0705 [hep-ex].

20. **CDF** Collaboration, T. Aaltonen *et al.*, "A Search for Resonant Production of $t\bar{t}$ pairs in 4.8 fb^{-1} of Integrated Luminosity of $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV," *Phys.Rev.* **D84** (2011) 072004, arXiv:1107.5063 [hep-ex].

21. **D0** Collaboration, V. M. Abazov *et al.*, "Search for a Narrow $t\bar{t}$ Resonance in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV," *Phys.Rev.* **D85** (2012) 051101, arXiv:1111.1271 [hep-ex].

22. **CDF** Collaboration, T. Aaltonen *et al.*, "Search for Resonant Production of $t\bar{t}$ Decaying to Jets in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV," *Phys.Rev.* **D84** (2011) 072003, arXiv:1108.4755 [hep-ex].

23. **D0** Collaboration, "Search for Resonances in the Lepton Plus Jets Final State in Collisions at $\sqrt{s} = 1.96$ TeV," *Physics Letters B* **668** no. 2, (2008) 98 – 104. http://www.sciencedirect.com/science/article/pii/S0370269308010149.

24. **D0** Collaboration, "Search for $t\bar{t}$ Resonances in the Lepton+Jets Final State in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV," *D0 Note* **5882-CONF** (2009) . http://www-d0.fnal.gov/Run2Physics/WWW/results/prelim/TOP/T83/.

25. **CMS** Collaboration, "Search for Heavy Narrow Resonances Decaying to $t\bar{t}$ in the Muon+Jets Channel," *CMS Physics Analysis Summary* **CMS-PAS-EXO-11-055** (2011) . http://cds.cern.ch/record/1376673.

26. **CMS** Collaboration, "Top Pair Invariant Mass Distribution in Lepton+Jets in 2011 Data," *CMS Physics Analysis Summary* **CMS-PAS-TOP-11-009** (2011) . http://cds.cern.ch/record/1429634.

27. **CMS** Collaboration, "Search for BSM $t\bar{t}$ Production in the Boosted All-Hadronic Final State," *CMS Physics Analysis Summary* **CMS-PAS-EXO-11-006** (2011) . http://cds.cern.ch/record/1370237.

28. **CMS** Collaboration, "Top Pair Invariant Mass Distribution in Dileptons," *CMS Physics Analysis Summary* **CMS-PAS-TOP-11-010** (2011) . http://cds.cern.ch/record/1438739.

29. O. S. Brning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report*. CERN, Geneva, 2004. http://cds.cern.ch/record/782076.

30. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical Design Report CMS. CERN, Geneva, 2006. http://cds.cern.ch/record/922757.

31. *CMS Physics: Technical Design Report Volume 2: Physics Performance*, vol. 34. 2006. http://cds.cern.ch/record/942733.

32. *ATLAS detector and physics performance: Technical Design Report, 1*. Technical Design Report ATLAS. CERN, Geneva, 1999. http://cds.cern.ch/record/391176.

33. *ATLAS detector and physics performance: Technical Design Report, 2*. Technical Design Report ATLAS. CERN, Geneva, 1999. http://cds.cern.ch/record/391177.

34. Cortese, P and Fabjan, Christian Wolfgang and Riccati, Lodovico and Safark, Karel and de Groot, Hans, *ALICE physics performance: Technical Design Report*. Technical Design Report ALICE. CERN, Geneva, 2005. http://cds.cern.ch/record/879894.

35. *LHCb reoptimized detector design and performance: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2003. http://cds.cern.ch/record/630827.

36. **CMS** Collaboration, "The CMS Experiment at the CERN LHC," *Journal of Instrumentation* **3** (Aug., 2008) .

37. **CMS** Collaboration, "CMS Tracking Performance Results from Early LHC Operation," arXiv:1007.1988 [hep-ex].

38. **CMS** Collaboration, "Performance of the CMS Hadron Calorimeter with Cosmic Ray Muons and LHC Beam Data," *Journal of Instrumentation* **5** (Mar., 2010) 3012, arXiv:0911.4991 [physics.ins-det].

39. **CMS** Collaboration, "Performance of CMS Muon Reconstruction in pp Collision Events at $\sqrt{s} = 7$ TeV," *Journal of Instrumentation* **7** (Oct., 2012) 2P, arXiv:1206.4071 [physics.ins-det].

40. **CMS** Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET," *CMS Physics Analysis Summary* **CMS-PAS-PFT-09-001** (2009) . http://cds.cern.ch/record/1194487.

41. **CMS** Collaboration, "Track Reconstruction in the CMS Tracker (in preparation)," *CMS Physics Analysis Summary* **CMS-PAS-TRK-09-001** (2009) .

42. **CMS** Collaboration, "Tracking and Primary Vertex Results in First 7 TeV Collisions," *CMS Physics Analysis Summary* **CMS-PAS-TRK-10-005** (2010) . http://cds.cern.ch/record/1279383.

43. M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k(t) Jet Clustering Algorithm," *JHEP* **0804** (2008) 063, arXiv:0802.1189 [hep-ph].

44. **DO** Collaboration, Abazov *et al.*, "Subjet Multiplicity of Gluon and Quark Jets Reconstructed with the k_T Algorithm in $p\bar{p}$ Collisions," *Phys. Rev. D* **65** (Feb, 2002) 052008. http://link.aps.org/doi/10.1103/PhysRevD.65.052008.

45. S. Bentvelsen and I. Meyer, "The Cambridge Jet Algorithm: Features and Applications," *Eur.Phys.J.* **C4** (1998) 623–629, arXiv:hep-ph/9803322 [hep-ph].

46. T. Speer, K. Prokofiev, R. Frhwirth, W. Waltenberger, and P. Vanlaer, "Vertex Fitting in the CMS Tracker," Tech. Rep. CMS-NOTE-2006-032, CERN, Geneva, Feb, 2006. http://cds.cern.ch/record/927395.

47. **CMS** Collaboration, S. Chatrchyan *et al.*, "Identification of b-quark jets with the CMS experiment," arXiv:1211.4462 [hep-ex].

48. P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, *et al.*, "Implications of CTEQ Global Analysis for Collider Observables," *Phys.Rev.* **D78** (2008) 013004, arXiv:0802.0007 [hep-ph].

49. J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, "MadGraph 5 : Going Beyond," *JHEP* **1106** (2011) 128, arXiv:1106.0522 [hep-ph].

50. S. Alioli, P. Nason, C. Oleari, and E. Re, "A General Framework for Implementing NLO Calculations in Shower Monte Carlo Programs: the POWHEG BOX," *JHEP* **1006** (2010) 043, arXiv:1002.2581 [hep-ph].

51. T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual," *JHEP* **0605** (2006) 026, arXiv:hep-ph/0603175 [hep-ph].

52. M. Asai, "Geant4-a Simulation Toolkit," *Trans.Amer.Nucl.Soc.* **95** (2006) 757.

53. J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Dubois, *et al.*, "Geant4 Developments and Applications," *IEEE Trans.Nucl.Sci.* **53** (2006) 270.

54. M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, "Matching Matrix Elements and Shower Evolution for Top-Quark Production in Hadronic Collisions," *JHEP* **0701** (2007) 013, arXiv:hep-ph/0611129 [hep-ph].

55. J. M. Campbell and R. Ellis, "MCFM for the Tevatron and the LHC," *Nucl.Phys.Proc.Suppl.* **205-206** (2010) 10–15, arXiv:1007.3492 [hep-ph].

56. R. Kleiss and W. J. Stirling, "Top Quark Production at Hadron Colliders: Some Useful Formulae," *Z. Phys. C* **40** (1988) 419.

57. K. Melnikov and F. Petriello, "Electroweak Gauge Boson Production at Hadron Colliders Through $O(\alpha_s^2)$," *Phys. Rev. D* **74** (2006) 114017, arXiv:hep-ph/0609070.

58. N. Kidonakis, "Next-to-Next-to-Leading-Order Collinear and Soft Gluon Corrections for t-Channel Single Top Quark Production," *Phys. Rev. D* **83** (2011) 091503, arXiv:1103.2792 [hep-ph].

59. N. Kidonakis, "NNLL resummation for s-channel single top quark production," *Phys. Rev. D* **81** (2010) 054028, arXiv:1001.5034 [hep-ph].

60. N. Kidonakis, "Two-Loop Soft Anomalous Dimensions for Single Top Quark Associated Production with a W- or H-," *Phys. Rev. D* **82** (2010) 054018, arXiv:1005.4451 [hep-ph].

61. C. T. Hill, "Topcolor Assisted Technicolor," *Phys. Lett. B* **345** (1995) 483, arXiv:hep-ph/9411426 [hep-ph].

62. **CMS** Collaboration, A. J. Zsigmond, "Inelastic Proton-Proton Cross Section Measurements in CMS at $\sqrt{s} = 7$ TeV," arXiv:1205.3142 [hep-ex].

63. **CMS** Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS," *Journal of Instrumentation* **6** (Nov., 2011) 11002, arXiv:1107.4277 [physics.ins-det].

64. **CMS** Collaboration, "b-Jet Identification in the CMS Experiment," *CMS Physics Analysis Summary* **CMS-PAS-BTV-11-004** (2011) . http://cds.cern.ch/record/1427247.

65. **CMS** Collaboration, "Measurement of b-Tagging Efficiency Using $t\bar{t}$ Events," *CMS Physics Analysis Summary* **CMS-PAS-BTV-11-003** (2011) . http://cds.cern.ch/record/1421611.

66. R. J. Barlow and C. Beeston, "Fitting Using Finite Monte Carlo Samples," *Comput.Phys.Commun.* **77** (1993) 219–228.

67. **CMS** Collaboration, "Measurement of the Top Quark Pair Differential Cross Sections at $\sqrt{s} = 7$ TeV," *CMS Physics Analysis Summary* **CMS-PAS-TOP-11-013** (2012) . http://cds.cern.ch/record/1422425.

68. T. Muller, J. Ott, and J. Wagner-Kuhr, "Theta - a Framework for Template-Based Modeling and Inference," *CMS Internal Note* **IN-2010/017** (2010) .

69. **ATLAS and CMS** Collaboration, "Procedure for the LHC Higgs Boson Search Combination in Summer 2011," ATL-PHYS-PUB-2011-011, CMS NOTE-2011/005, 2011. https://cds.cern.ch/record/1379837.

70. A. L. Read, "Presentation of Search Results: The CLs Technique," *J. Phys.* **G28** (2002) 2693.

71. T. Junk, "Confidence Level Computation for Combining Searches with Small Statistics," *Nucl. Instrum. Meth.* **A434** (1999) 435, arXiv:hep-ex/9902006 [hep-ex].

72. R. J. Barlow and C. Beeston, "Fitting Using Finite Monte Carlo Samples," *Comput.Phys.Commun.* **77** (1993) 219–228.

73. J. S. Conway, "Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra," *ArXiv e-prints* (Mar., 2011) , arXiv:1103.0354 [physics.data-an].

74. **CMS** Collaboration, "Search for High-Mass Resonances Decaying to Top Quark Pairs in the Lepton+Jets Channel," *CMS Physics Analysis Summary* **CMS-PAS-EXO-11-093** (2012) . http://cds.cern.ch/record/1454695.

VITA

Name	Samvel Khalatian
Education	Ph.D., Physics University of Illinois at Chicago, Chicago, Illinois 2013
	M.Sc., Physics National Research Nuclear University "MEPhI", Moscow, Russia 2004
Summer Schools	CTEQ Summer School on QCD Analysis and Phenomenology University of Madison, Wisconsin 2011
	The 5^{th} CERN-Fermilab Hadron Collider Physics Summer School Fermilab, Batavia, Illinois 2010
	10^{th} ICFA School on Instrumentation in Elementary Particle Physics Itacuruca, Rio de Janeiro, Brazil 2003
Awards	Kouvel Award for Graduate Research University of Illinois at Chicago, Chicago, Illinois 2011
Teaching	Teaching Assistant University of Illinois at Chciago, Chicago, Illinois 2007

Experience

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

Research Assistant (CMS Experiment)
University of Illinois at Chicago, Chicago, Illinois
2006 - 2007

Research Assistant (D0 Experiment)
Boston University, Boston, Massachusetts
2005 - 2006

Intern (CDF Experiment)
Fermi National Accelerator Laboratory, Illinois
2004

Presentations

Search for High Mass Resonances Decaying to TTbar in the Lepton+Jets Channel
User's Meeting Poster Session, FermiLab, Illinois
2012

Search for TTbar Resonances and Implications for New Physics Models
24^{th} Rencontres de Blois: Particle Physics and Cosmology, Blois, France
2012

Search for High-Mass Resonances Decaying to TTbar in the Lepton+Jets Channel with the CMS Detector
LHC Students Poster Session, CERN, Switzerland
2012

non-ROOT I/O
Future Computing in Particle Physics (ATLAS Workshop), Edinburgh e-Science Institute, Scotland
2011

Measurement of the b-Tagging Efficiency Using Events with Jets Containing Muons
FNAL Prospective Students Meeting, FermiLab, Illinois
2011

Measurement of the b-Tagging Efficiency Using Events with Jets Containing Muons
APS April Meeting, Anahaim, California
2011

non-ROOT I/O: Performance Boost
CMS Computing Week, CERN, Switzerland
2011

Efficiency Measurement with System8
CMS Week, CERN, Switzerland
2010

Measuring the b-Tagging Efficiency with the System8 Method
UIC High-Energy Physics Meeting, University of Illinois at Chicago, Illinois
2010

Introduction to System8
UIC b-Tagging Group Meeting, University of Illinois at Chicago, Illinois
2010

CMS Tracker Upgrade
IEEE Nuclear Science Symposium and Medical Imaging Conference, Orlando, Florida
2009

Conference Proceedings

S.Khalatian, *"Search for tt Resonances and Implications for New Physics Models"*
24^{th} Rencontres de Blois: Particle Physics and Cosmology

Publications

CMS Collaboration, *"Search for resonant $t\bar{t}$ production in lepton+jets events in pp collisions at $\sqrt{s} = 7$ TeV"*
Journal of High Energy Physics, http://arxiv.org/abs/1209.4397

CMS Collaboration, *"Search for High-Mass Resonances Decaying to Top Quark Pairs in the Lepton+Jets Channel"*
CMS-PAS-EXO-11-093, http://cds.cern.ch/record/1454695/

CMS Collaboration, *"Search for High-Mass Resonances Decaying to TTbar in the Electron+Jets Channel"*
CMS-PAS-EXO-11-092, http://cds.cern.ch/record/1423037