Measurement of $W + b\bar{b}$ and a search for MSSM Higgs Bosons with the CMS Detector at the LHC

By

Isobel Rose Ojalvo

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Physics)

at the

University of Wisconsin – Madison

2014

Date of final oral examination: 2/11/2014

The dissertation is approved by the following members of the Final Oral Committee:

Wesley H. Smith, Bjorn Wiik Professor, Physics

Sridhara Dasu, Professor, Physics

Lisa Everett, Professor, Physics

Matthew Herndon, Professor, Physics

Terrence Millar, Professor, Mathematics

"The places you're going to are never on the map."

–Jim Henson

Abstract

This thesis describes a Standard Model (SM) cross section measurement of $\mathrm{W} + \mathrm{b\bar{b}}$ as well as a search for neutral Higgs bosons in the Minimal Supersymmetric Extension of the Standard Model (MSSM) decaying to tau pairs.

The measurement of $\mathrm{W} + \mathrm{b\bar{b}}$ was performed using proton-proton collisions at $\sqrt{s} = 7$ TeV in a data sample collected with the CMS experiment at the LHC corresponding to an integrated luminosity of 5.0 fb^{-1}. The $\mathrm{W} + \mathrm{b\bar{b}}$ events are selected in the $\mathrm{W} \to \mu\nu$ decay mode by requiring a muon with transverse momentum $p_{\mathrm{T}} > 25$ GeV and pseudorapidity $|\eta| < 2.1$, and exactly two b-tagged jets with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.4$. The measured $\mathrm{W} + \mathrm{b\bar{b}}$ production cross section in the fiducial region, calculated at the level of final-state particles, is 0.53 ± 0.05 (stat.) $\pm$ 0.09 (syst.) $\pm$ 0.06 (th.) $\pm$ 0.01 (lum.) pb, in agreement with the SM prediction. This measurement is a sensitive test of heavy quark production calculated with perturbative QCD. It also serves as an important benchmark in new physics searches which include a single isolated lepton and one or more b jets in the final state, as $\mathrm{W} + \mathrm{b\bar{b}}$ becomes an irreducible background.

Also presented is a search for the CP-even MSSM Higgs bosons, H and h, and the CP-odd MSSM pseudoscalar, A, in their decays to tau pairs. This search is performed using events recorded by the CMS experiment at the LHC in 2011 and 2012 at a center-of-mass energy of 7 TeV and 8 TeV respectively. The dataset corresponds to an integrated luminosity of 24.6 fb^{-1}, with 4.9 fb^{-1} at 7 TeV and 19.7 fb^{-1} at 8 TeV. To enhance the sensitivity to neutral MSSM Higgs bosons, the search includes the case where the Higgs boson is produced in association with a b-quark jet. No excess is observed in the tau-pair invariant-mass spectrum.

Acknowledgments

In the fall of 2008 I had been working as a Systems Engineer for the Space and Intelligence Systems division of Boeing for over a year when I decided I needed to make a career change. I had obtained an undergraduate degree in physics and mathematics from Rensselaer Polytechnic Institute in 2007, a university which was also my grandfather's alma mater. While at Boeing, I had the pleasure of working with Dr. Thomas Sutherland. Tom had five giant filing cabinets in his office which contained countless schematics, reports and papers concerning satellite design from his greater than 40 year tenure at Boeing. One of my favorite things about Tom was that no matter what problem we were trying to solve or what quantity we needed to estimate, he always insisted that we start from 'first principles'. Working with Tom was the first time I realized that one could break any real world physics problem, no matter how complex, down into manageable parts. When I told him I wanted to get my Ph.D. in physics he agreed to write a letter of recommendation for me.

As soon as I met Wesley during student prospective week, I knew I wanted to work in his group; anyone with that much energy and enthusiasm must be exciting to work for. I certainly was not wrong and was lucky to have the opportunity learn so much about, trigger, physics and technology from him. I really got to know Sridhara while working on calibration for the Regional Calorimeter Trigger and really liked his working style. I have learned very much about physics, analysis and computing from him and have been very lucky to have him as a mentor. I must really thank both of them for giving me a large amount of independence while doing analysis but still making sure I was pointed in the right direction. I have a huge amount of respect for both Wesley and Sridhara and look forward to working with them on even more complex projects as a post-doc.

Thank you to the fantastic professors and staff at the University of Wisconsin Madison who encouraged me to overcome and thrive in such a stressful environment. The graduate students are lucky to have Aimee and Renee looking out for them! Professor Lisa Everett taught an extremely useful and difficult Quantum Mechanics course. I learned from her how to solve problems under pressure and to continue working even if there seemed there was no possibility of success. I can not say enough good things about her and am so lucky to have found a physics role model.

Thank you to my graduate student colleagues who began with me in 2009, all 47 of you! If I mention one of them then I will feel bad for leaving others out, so I will be economical and say I enjoyed hanging out with the Dans. Kara and Xian are both my absolute favorites as well; I love that I could always got talk to Kara when I was grumpy and to Xian when I was happy.

My partner Aaron, if it weren't for you I probably would have keeled over due to malnourishment while I was taking courses, teaching and studying 16 hours a day. Then I moved to Switzerland and you put up with having a digital girlfriend for the last two years. You were always the first person to listen when anything went wrong.

Sharing an office with a variety of people including Will, Austin, Ian, Josh, Bethany, Michalis always kept the day to day life interesting. Thanks for not saying anything about my incredibly messy desk and moldy food in the fridge. Maria, now that I am also a post-doc too I would be a bit disappointed if we stopped arguing altogether. Thank you to my colleague and mentor Michalis, who I have learned so much from. Our physics arguments greatly helped me hone my skills for meetings, and presentations. Be sure to keep the POGs in order!!

Anyone who knows me well knows that, no matter how far away from them I'm currently living, I talk about my family constantly. My parents some how managed to raise six children and instill in them a life long dedication to learning. I will always

remember fondly our trips to the science museum and raising Monarch caterpillars when we were kids. Their 'grand experiment' has thus far been a huge success. Having three younger brothers, Daniel, Stephen and Jimmy, gave me an opportunity to test out my teaching and managerial skills at a very early age. This thesis is dedicated to my two sisters, Laureen and Katherine; having them as role models has always forced me to push myself to be a better version of myself. Katherine taught me how to stick up for myself on the bus when we were in middle school together and Laureen has always been extremely supportive. I could not ask for better big sisters.

After giving my advisors a first draft of my thesis I took my first real vacation in over two years. So, as I sit on a boat in the Adaman sea and watch the sun creep slowly above the horizon and think about the last 10 years that have culminated in achieving this important goal, I've come to realize that challenges and adversity are only opportunities in disguise.

Contents

List of Figures

List of Tables

Chapter 1

The Standard Model

1.1 Introduction

Scientific study of the building blocks of matter began in the late 19th century. In this section, a selected historical approach which highlights important advances in particle physics pertinent to the development of the standard model of physics is taken culminating in the recent discovery of a higgs boson.

1.2 Brief History of Particle Physics

The first leap forward in understanding of the particle physics field came in 1897 when J. J. Thomson fired what was then known as 'cathode rays' into a magnetic field[1]. He observed their circular orbit and by measuring the radius of the orbit, Thomson was able to make the first measurement of the electron mass to charge ratio. Thomson already knew these electrons were in some way associated to the atom and he hypothesized that the electrons were distributed evenly within the atom, much like plums in a pudding. In 1899 Rutherford tested this hypothesis by firing a beam of α particles at a thin gold sheet[2]. He observed that most of the α particles went straight through the sheet, but a few bounced off in various directions; this suggested that the α particles were made mostly of space with a very dense nucleus that would on occasion interact and cause the scattering. Rutherford named the nucleus of the lightest atom the 'proton'.

Even if at this time physicists had an idea of the nature of the most simple of atoms the relationship between the nucleus and the electrons was still not well understood. In 1914 Niels Bohr successfully developed a model of the Hydrogen atom by approximating it to a planet (the electron) circling the Sun (the nucleus). This rudimentary model that Bohr developed was very accurate in determining the Hydrogen spectrum, however, was ultimately insufficient in describing atomic physics.

In 1932, about thirty years after the discovery of the proton, Chadwick explained the difference between the mass and charge of larger atoms (for example He_4^2) with the discovery of the neutron.

Around the same time that Rutherford was measuring the mass of the electron, Planck was performing studies on black-body radiation. A black-body will absorb all radiation incident upon it and then re-radiate this energy, in classical physics this radiation can occur at any wavelength; this classical model lead to an 'ultra-violet catastrophe' whereby an ideal black body at thermal equilibrium will emit radiation with infinite power. Planck found that if electromagnetic radiation emission was quantized based upon the frequency, $E = h\nu$, then the ultraviolet catastrophe would be avoided. In 1905 Einstein took Planck's idea of the quantization of electromagnetic radiation a step further and claimed that radiation emission from an atom is intrinsically quantized, $E \leq h\nu - W$, where W is the work function which was found to be the amount of energy needed to transport an electron from its current bound energy level to an unbound free-state. This 'photoelectric effect' was long rejected by the physics community until decisively proven by Compton Scattering experiments which were performed by Compton in 1923. These results verified that light is both a particle and a wave.

The proposal of Dirac's equation in 1928 marks the ending of the era of non-relativistic quantum mechanics[3]. Dirac's famous equation combined high momentum physics described by Einstein's theory of relativity with the smallest of distances described by quantum mechanics. Although the positive energy solutions of Dirac's equation appeared to correctly describe electrons experimentally it had negative energy solutions, a feature which Dirac did not like at the time. The true triumph of the theory came in 1932 when Anderson discovered the positron. In the 1940's Feynman and Stuckelberg interpreted the positron and negative energy solutions as an elec-

tron moving backwards in time. This idea of antiparticles is of central importance to quantum field theory.

The first significant theory of the strong force was developed by Yukawa in 1934. Yukawa believed that the proton and the neutron were attracted to each other by some sort of field with a heavy mediator called the meson. By 1937 the pion (π) and the muon (μ) were identified in cosmic rays; while the π was the meson Yukawa had suggested it took much longer to understand the mystery of the μ as well as to develop an understanding of the the interactions of fundamental particles.

During the study of nuclear β decay in the 1930's a problem was observed. In β decays a nucleus transforms into a slightly lighter nucleus with the emission of an electron; if the neutron is at rest then the electron and proton must emerge back to back with equal and opposite momenta. However, when the energy of the electron was recorded over many experimental iterations it was shown to vary! A solution to this was proposed by Pauli that a neutral particle (the neutrino) carries off this missing energy; today this process is known to be $n \to p^+ + e^- + \bar{\nu}$.

By 1947 particle physics was, for the most part, well described and understood but late in 1947 while studying cosmic rays incident on a thick lead plate Butler and Rochester observed $K^0 \to \pi^+\pi^-$ and then in 1949 Brown observed $K^+ \to \pi^+\pi^+\pi^-$. This ushered in a slew of particle discovery in the 1950's. During this time many issues arose: Did the production of Λ via $p + pi^- \to K^0 + \Lambda$ take only 10^{-23} seconds but the decay $\Lambda \to p + \pi^+$ takes 10^{-10} seconds? Why was there no $p \to e^+ + \gamma$? These discoveries began the development of conservation laws such as baryon number and strangeness. Then, in 1961 Gell-Mann arranged baryons in geometric patterns according to their charge, Q, and strangeness, S, and successfully predicted the Ω^-. He called this the 'Eightfold Way' at the time to was considered an periodic table for particle physics and initiated many future advancements.

Despite all the best efforts, an individual quark has not been observed, which led to rampant skepticism about the quark model. This produced much skepticism about the quark model in the 1960's and 70's. To explain the absence of individual quarks the idea of 'quark confinement' was introduced, i.e. quarks are always confined within mesons or baryons. In the late 1960's physicists at the Stanford Linear Accelerator (SLAC) performed deep inelastic scattering experiments to study the sub-structure of the nucleon. By way of firing an electron at a nucleon and measuring the scattering of the outgoing electron the experimental results hinted that the nucleon was actually made up of many point-like constituents. These point-like constituents would eventually be called 'partons'. The discovery of these point-like constituents presented their own problem as per the Pauli exclusion principal two particles cannot exist in the exact same state. To solve this issue, W. Greenberg proposed that quarks come in 3 colors, red, green and blue, (which have anti-red, anti-green, and anti-blue partners) and that each quark bound state is actually colorless. The skepticism about the quark model remained widespread until the discovery of the J/ψ particle by separate groups at SLAC and MIT. The J/ψ is the $c\bar{c}$ meson; its ground and excited states were shown to be well described by Quantum Chromodynamics (QCD).

A very unanticipated third generation of lepton, the τ-lepton, was discovered at SLAC[4]. The τ has a heavy mass (1.8 GeV) and the lifetime of the τ is much shorter than that of the μ; the reconstruction of the τ is also more difficult due to the fact that it decays both to leptons ($\tau^- \rightarrow \mu^- \bar{\nu_\tau} \nu_\mu$) and to hadrons ($\tau^- \rightarrow \pi^+ \pi^- \pi^- \nu_\tau$). This discovery of a third lepton generation was unprecedented as there had previously been discovered only 2 generations of quarks and 2 generations of leptons.

In 1977 an Υ resonance was observed via its decay to $\mu^+ \mu^-$ at approximately 9.5 GeV. It was later shown that this peak at 9.5 GeV was actually three $b\bar{b}$ resonances, with the ground state at 9.4GeV. With a mass of 4.3GeV, the b-quark is much more

massive than the c-quark (1.2 GeV); the b-quark's long lifetime and high mass means it is of great importance in the discovery of new physics.

By 1979 the production of a 2 jet event, $q\bar{q}$ had already been observed in electron-positron collisions at PETRA via the process $e^+e^- \rightarrow q\bar{q}$. However, the observation of a 3-Jet event at PETRA was extremely exciting as it was the first evidence of a high energy quark emitting a gluon via Bremsstrahlung, $e^+e^- \rightarrow q\bar{q}g$[5].

Fermi theorized that the weak interaction in β decays occur at a single point. However, it was quickly seen that at high energies this theory must fail and there must exist an intermediate vector boson to mediate the incredibly weak interaction (10^{13} times weaker than the strong interaction). Electroweak theory was proposed by Glashow, Weinberg and Salam and it predicted the mass of the $W^{\pm}$ and Z bosons. In the 1970s CERN began construction of a proton-proton collider, the Super Proton Synchrotron (SPS), that would operate with a center of mass energy of 540 GeV. The UA1 detector was constructed on this ring by a large collaboration at CERN led by Carlo Rubbia. In January of 1983 Rubbia announced the discovery of the W, and 5 months later the Z was reported as well.

This summary of observations and tests of the standard model brings us to July 2nd 2012 when physicists separately from the CMS and ATLAS experiments at CERN both independently reported observation of a Higgs-like boson at approximately 125 GeV. The discovery of this particle confirmed that spontaneous symmetry breaking via the Higgs Mechanism gives the vector bosons their masses. In the following sections the Standard Model (SM) and an extension of this model, the Minimally Supersymmetric Standard Model (MSSM), is described in more detail. The discovery of a SM Higgs-like boson that couples to bosons and fermions is a major milestone in the story of particle physics. It remains to be seen where the story will go; a search for Super Symmetry is presented with no strong evidence yet to support this model.

However, many important questions about the universe remain unanswered and so the saga continues.

1.3 The Standard Model of Particle Physics

Following the advances of the 20th century it has been experimentally determined that the fundamental particles in the SM consists of 3 generations of quarks and leptons where each generation has two particles. Quarks are spin one-half particles

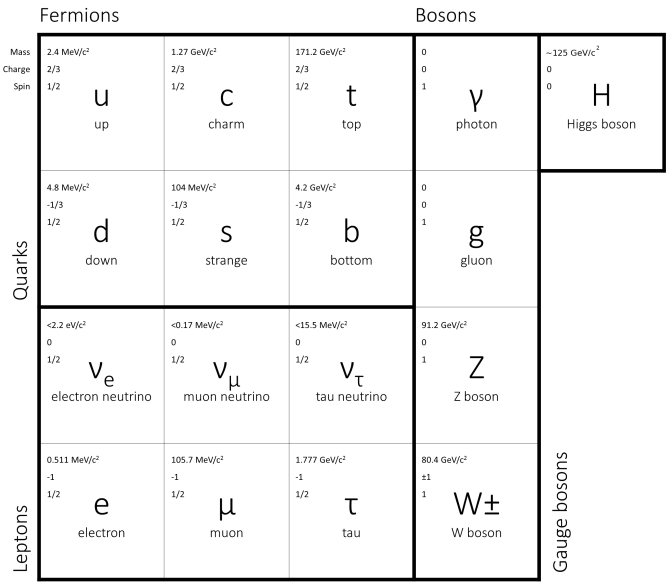


Figure 1.1: Standard Model Particles

that are triplets of the color group. There are 6 flavors of quarks, these are grouped into three generations: t (top) and b (bottom), c (charm) and s (strange), u (up) and d (down). The (u,c,t) quarks have a charge of $\frac{2}{3}$ while the (d,s,b) have a charge of $-\frac{1}{3}$. The fermions (e,μ,τ) all have a charge of -1 while their neutrino partners are neutral. Recent experiments in neutrino oscillation have hinted that neutrinos have a very small mass. Each fermion has an anti-particle partner with the same mass but opposite charge.

Each fundamental force is associated with spin 1 mediator particles. The weak interactions are mediated by the $W^{\pm}$ and Z; the electromagnetic interactions are mediated by the photon and the strong interactions are mediated by the 8 colored

gluons. The Higgs Mechanism is responsible for giving mass to the $W^{\pm}$, Z and the fermions. The SM of particle physics describes very successfully the electroweak and strong interaction of elementary particles over a wide range of energies.

The Standard Model of particle physics (SM) follows a $SU(3) \times SU(2)_L \times U(1)$ symmetry. The $SU(2)_L \times U(1)$ describes electroweak interactions. The $SU(3)$ group describes color and the interactions between gluons and quarks. In the unbroken $SU(2)_L \times U(1)$ symmetry gauge bosons and fermions are massless. However, Spontaneous Symmetry Breaking which is the result of couplings to a Higgs field with a non-zero expectation value results in the $W^{\pm}$ and Z bosons acquiring mass. The quarks and leptons can also acquire mass through couplings to the Higgs field. The study of SM W boson production in association with b-quarks is described in this thesis.

Recent discovery of a particle that closely resembles the SM Higgs boson completes the Standard Model. The discovery of this SM Higgs-like boson was announced jointly by the CMS and ATLAS collaborations in July of 2012 [6, 7]. As of the writing of this thesis the Higgs boson has been observed at CMS via its decay to ZZ^*, $\gamma\gamma$ and $\tau\tau$. The mass of this Higgs boson, as measured by CMS, is 125.6 $\pm$ 0.4(stat.) $\pm$ 0.2(syst.). This thesis also details a search for beyond the SM Higgs bosons. At the LHC, searches for beyond the standard model physics are of central importance.

Chapter 2

$\mathrm{W} + \mathrm{b\bar{b}}$ Production

2.1 Production in Proton-Proton Collisions

In a hadron collider, such as the LHC, W bosons are primarily produced through the annihilation of an up-type quark (Q=$\pm\frac{2}{3}$) and a down-type anti-quark (Q=$\pm\frac{1}{3}$). W coupling is favored within quark generations, $u\bar{d} \to W^+$. Although not 0, cross-generation, $s\bar{u} \to W^-$, transition is suppressed. This is paramaterized by introducing

$$\begin{pmatrix} u \\ d' \end{pmatrix}, \begin{pmatrix} c \\ s' \end{pmatrix}, \begin{pmatrix} t \\ b' \end{pmatrix} \tag{2.1}$$

where (d', s', b') are mixed states of (d, s, b); namely,

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.2}$$

The Cabbibo-Kobayashi-Maskawa matrix $\mathbf{V}$ is written in terms of three generalized Cabbibo angles and a complex phase. The standard model provides no way to calculate the weak mixing angles, instead they are taken from data. With this effect in mind we can write the matrix element for W production through quark fusion.

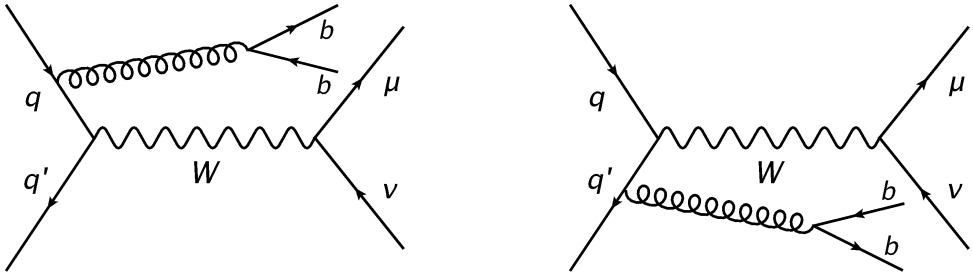


Figure 2.1: Leading Order W + $b\bar{b}$ diagrams

Neglecting the mass of the quarks, spins and the W polarization, the matrix element is,

$$\sum |M|^2 = |V_{qq'}|^2 \frac{8G_F}{\sqrt{2}} M_W^4 \tag{2.3}$$

The above equation is summed over all quarks, qq'. The cross section for W production is,

$$\sigma = 2 \times 2\pi \frac{G_F}{\sqrt{2}} M_W^2 |V_{qq'}|^2 \delta(s - M_W^2) \tag{2.4}$$

Where s is the center of mass energy squared of the two partons, $s = (p_1 + p_2)^2$. Importantly, the term $\delta(s - M_W^2)$ leads us to the conclusion that incoming quarks are required to have a center of mass energy equal to the mass of the W. In 2011 at the LHC protons were collided with an energy of $\sqrt{S} = 7$ TeV, and 8 TeV in 2012. For a W boson to be created two interacting partons must have suitable momentum fractions x, y such that $(x + y)S = s$. At a proton-proton collider the production cross section is:

$$\sigma = 2 \times \frac{2\pi G_F}{3\sqrt{2}} \int dxdy \sum V_{q,\bar{q}'} xy S f_q(x) f_{\bar{q}'}(y) \tag{2.5}$$

where $f_q(x)$ and $f_{\bar{q}'}(y)$ are parton distribution functions which represent the probability densities to find a parton carrying a momentum faction x. The branching ratio of the W to $\mu\ \nu$, as predicted by theory, is approximately 0.11.

As shown in figure 2.1, the W + $b\bar{b}$ process at leading order (LO) in perturbative QCD calculations requires that one of the initial partons radiates a gluon and the

gluon then splits into a $b\bar{b}$ pair. Next to leading order (NLO) calculations, with more complicated diagrams, include perturbative corrections of $O(\alpha_S^2)$. When perturbative corrections are taken into account each additional vertex involving the gluon adds a factor of order α_s. A full calculation of the theoretical $W + b\bar{b}$ cross section at NLO can be found here [8, 9]. Validation of theoretical calculations in perturbative QCD is important.

$W + b\bar{b}$ is an irreducible background to many searches including: hW where a W radiates a higgs boson and the higgs bosons then decays to a b quark pair. The beyond the SM higgs boson search presented in this thesis also suffers from $W + b\bar{b}$ background. A measurement of $W + b\bar{b}$ via the W decay to $\mu + \nu$ is presented as a central focus of this thesis.

2.2 Jet Hadronization and b-quarks

During a high energy interaction involving quarks in the final state (for instance $g \to q\bar{q}$) the final state quarks fly off as individual particles for a very brief moment. After they reach a distance of $10^{-15}m$ the strong interaction is so great that new quarks and anti-quarks are produced. These new quarks and anti-quarks then combine in many different combinations to form baryons and mesons. This process is known as hadronization and is illustrated in figure 2.2. Although many QCD next to leading order calculations have been performed, hadronization can not yet be directly calculated. Instead monte carlo event generators simulate this hadronic production using parton showering with various fragmentation models.

If we consider a high energy parton, b, with with energy E_b which produces a hadron, B, with energy E_B then the hadron's energy fraction is $z = E_B/E_b$. The

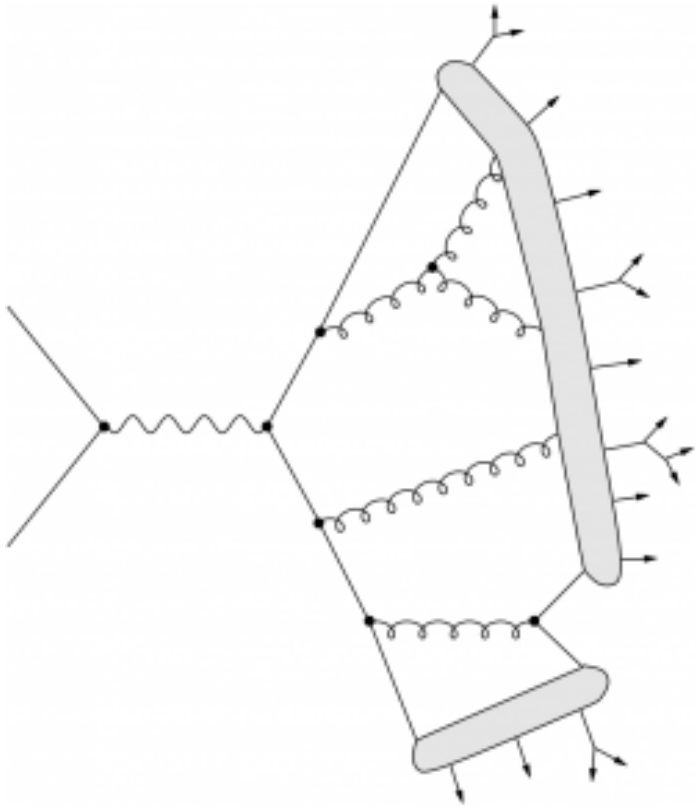


Figure 2.2: Jet Hadronization

probability of finding B in the range z to dz for a heavy b quark is,

$$\int D_b^B(z)dz = \int \frac{Constant}{z(1-\frac{1}{z}-\frac{\epsilon_b}{1-z})}dz \tag{2.6}$$

Here $\epsilon_b = \langle m_b^2 + p_{T,b}^2 \rangle / \langle m_B^2 + p_{T,B}^2 \rangle$, and is expected to be proportional to m_b^{-2}. In the case where the fragmenting parton is a b quark, it loses only a very small amount of its energy to materialize a number of light quark pairs and a majority of the energy is carried off by the B hadron. For comparison, the average b-meson lifetime is $1.55 \pm 0.06ps$, this corresponds to a proper lifetime of $c\tau = 463\mu m$. For a b-meson with $20GeV$ momentum this would correspond to 1.9 mm in the lab frame. Therefore a small V_{cb} allows for detection of a separated vertex. This long lifetime is extremely helpful in distinguishing between jets which originate from b quarks and jets that originate from lighter quarks or gluons that do not produce heavy quarks.

Chapter 3

Supersymmetry and the MSSM

The standard model performs well in describing experimental observations at energies around the electroweak scale of $O(246\text{GeV})$. The recent discovery of a standard model-like higgs boson (m_h =125 GeV) brings with it more confidence in this theory. However, one must explain why the mass of the Higgs boson is as low as 125 GeV whereas the planck scale is $O(1.22 \times 10^{19}\text{GeV})$ [10].

3.1 Motivations for SUSY

A fermion loop correction to m_h, seen in figure 3.1a, manifests as a negative-sign correction to the higgs mass, namely,

$$\Delta m_H^2 = -\frac{|\lambda_f|^2}{8\pi^2}\Lambda_{UV}^2 \tag{3.1}$$

where Λ_{UV} is the ultraviolet high-mass cutoff, therefore, the correction to the higgs mass is large. It turns out that the correction due to a scalar loop, as seen in figure 3.1b, has an opposite sign to that of the fermion loop. It is quite unnatural that the difference of two large numbers has to be tuned carefully to obtain the 125 GeV observed Higgs boson mass. Supersymmetry supposes that there exists a scalar

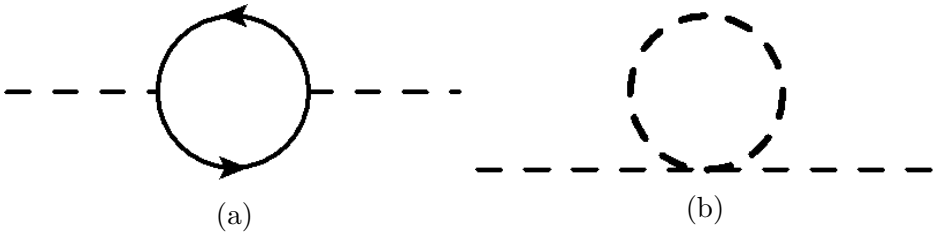


Figure 3.1: One loop quantum corrections for a fermion (3.1b) and a scalar (3.1b).

particle for every fermion and vice versa. Therefore, this theory of symmetry between bosons and fermions can naturally lead to a low mass higgs.

To define SUSY a set of generators which transforms a fermionic state into a bosonic state and vice versa is needed. The simplest operator which can perform these operations is a 2 component Weyl spinor Q such that,

$$Q|\text{Boson}\rangle = |\text{Fermion}\rangle \qquad Q|\text{Fermion}\rangle = |\text{Boson}\rangle. \tag{3.2}$$

The generators Q and $Q^\dagger$ must obey the anticommutation and commutation algebra,

$$\{Q, Q^\dagger\} = P^\mu \tag{3.3}$$

$$\{Q, Q\} = \{Q^\dagger, Q^\dagger\} = 0 \tag{3.4}$$

$$[P^\mu, Q] = [P^\mu, Q^\dagger] = 0 \tag{3.5}$$

where P^μ is the generator of space-time translations.

In the supersymmetric model the SM and SUSY particles are arranged into supermultiplets. These supermultiplets must contain both the SM fermion and the boson superpartner. In the simplest approach, a spin $\frac{1}{2}$ weyl fermion (for example, e) must have a spin-0 superpartner ($\tilde{e}$), likewise, a supermultiplet with a spin-1 gauge boson (W) would have a spin $\frac{1}{2}$ weyl fermion superpartner ($\tilde{W}$). The naming convention for super particles calls for the names of the fermionic standard model particles to be prepended by an 's', for example, 'selectron' is short of 'scalar electron'; the gauge

bosons are appended with an '-ino', for example $\tilde{W}$ is known as a 'wino'. These particles are displayed in figure 3.2. If SUSY remained unbroken then m_e must be equal to $m_{\tilde{e}}$. Since sparticles like $\tilde{e}$ have not yet been discovered this implies that SUSY must be a broken symmetry.

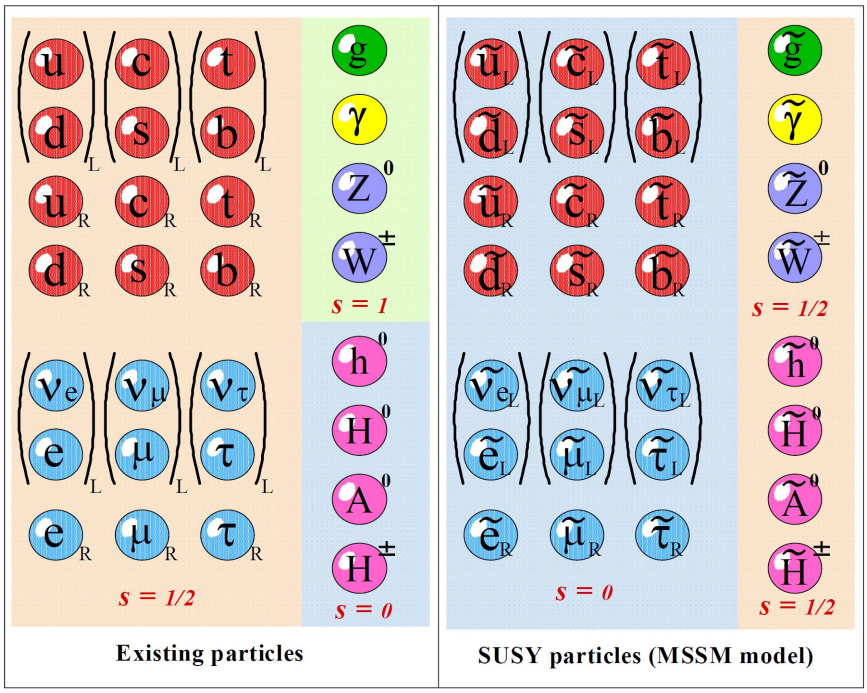


Figure 3.2: Particle Spectrum in the MSSM

3.2 Higgs Sector of the SM

Electroweak interactions happen on a very short range (10^{-14} m). To confine weak interactions to this short range the $W^{\pm}$ and Z bosons must be massive. However, the lagrangians governing fundamental interactions of particles are written in terms of gauge symmetries and in an unbroken gauge theory the gauge bosons must be massless. If one were to simply break the symmetry by adding in a mass term then the theory would no longer be renormalizeable. Spontaneous Symmetry Breaking (SSB) breaks the symmetry but still requires that the Lagrangian is invariant under

gauge transformations. This can done by introducing a complex $SU(2)_L$ doublet scalar higgs field,

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{3.6}$$

This two component complex scalar field has four degrees of freedom, three of them give mass to the $W^{\pm}$ and Z, the fourth appears as a massive physical particle, the Standard Model Higgs boson. The Lagrangian describing interactions with this field is

$$\mathscr{L} = (D_\mu \Phi^\dagger)(D^\mu \Phi) + V(\Phi) \tag{3.7}$$

where D_μ is a covariant derivative and the scalar potential has the form,

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \frac{\lambda}{4}(\Phi^\dagger \Phi)^2. \tag{3.8}$$

If $\lambda > 0$ and $\mu^2 > 0$ then the shape of the potential is symmetric, however, for $\mu^2 < 0$ the minimum is a circle of radius v where $v^2 = -\mu^2/\lambda$. Figure 3.3 illustrates the shape of this potential. This is interpreted as a non-vanishing expectation value of the higgs field in the vacuum state. The mass of the higgs boson is $m_h = \sqrt{2\lambda v^2}$. SSB generates mass for the both the gauge bosons and the fermions. Neglecting radiative corrections these are,

$$M_W = \frac{1}{2} g v, M_Z = \frac{\sqrt{g^2 + g'^2}}{2}, m_f = \frac{G_f v}{\sqrt{2}}. \tag{3.9}$$

The coupling between the higgs field and the massive gauge bosons is proportional to their mass squared while the coupling to the massive fermions is proportional to m_f/v.

3.3 Higgs Sector of the MSSM

The most simple supersymmetric extension of the standard model is the MSSM. In the MSSM there are two higgs doublets with an $SU(2)_L$ symmetry,

$$\Phi_1 = \begin{pmatrix} \phi_1^{0*} \\ -\phi_1^- \end{pmatrix} \qquad \Phi_2 = \begin{pmatrix} \phi_2^+ \\ -\phi_2^0 \end{pmatrix}, \tag{3.10}$$

The Φ_1 has a hypercharge of -1 and gives mass to each of the down-type quarks and charged leptons whereas Φ_2 gives masses to the up-type quarks. The extra doublet is needed to cancel out the corresponding supersymmetric higgs fermion contributions. H_1^0 and H_2^0 acquire vacuum expectation values v_1 and v_2 where

$$v = \sqrt{2}(v_1^2 + v_2^2)^{\frac{1}{2}}. \tag{3.11}$$

The ratios of v_1 and v_2 is written as,

$$\tan(\beta) = \frac{v_2}{v_1} \tag{3.12}$$

Consequently the MSSM model has a total of five higgs bosons: two neutral CP-even higgs, h and H, a CP-odd pseudoscalar, A, and two charged higgs, $H^{\pm}$. By translating the fields in (3.10) to their minima and diagonalizing the matrix the following higgs

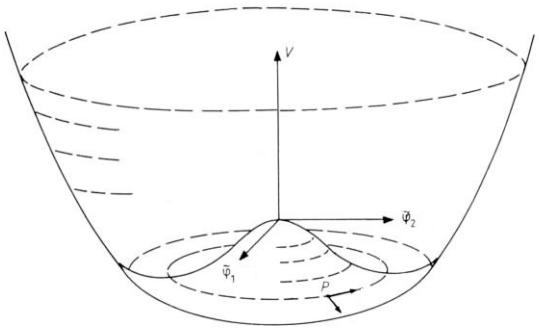


Figure 3.3: $V(\phi)$ potential with $\lambda > 0$ and $\mu^2 < 0$

boson states are obtained,

$$\begin{pmatrix} H_1^0 \\ H_2^0 \end{pmatrix} = \sqrt{2} \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} Re\phi_1^{0*} - v_1 \\ Re\phi_2^0 - v_2 \end{pmatrix} \tag{3.13}$$

$$H_3^0 = \sqrt{2}(\sin\beta Im\phi_1^{0*} + \cos\beta Im\phi_2^0) \tag{3.14}$$

$$H^- = (H^+)^* = -\phi_1^- \sin\beta + \phi_2^- \cos\beta \tag{3.15}$$

In the MSSM at tree level the number of independent parameters can be reduced to 3: m_W, $\tan\beta$ and the mass of the pseudoscalar, m_A. The neutral boson masses are,

$$m_h^2 = \frac{1}{2}\left(m_A^2 + M_Z^2 - \left[(m_A^2 + M_Z^2)^2 - 4M_Z^2 m_A^2 \cos^2(2\beta)\right]^{\frac{1}{2}}\right) \tag{3.16}$$

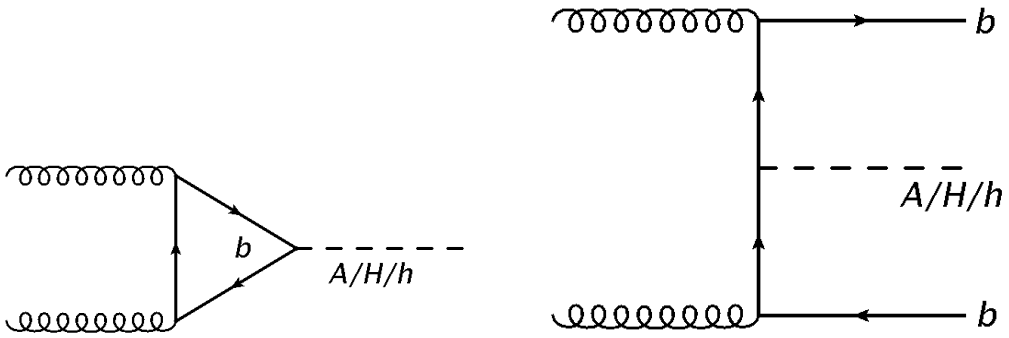
$$m_H^2 = \frac{1}{2}\left(m_A^2 + M_Z^2 + \left[(m_A^2 + M_Z^2)^2 - 4M_Z^2 m_A^2 \cos^2(2\beta)\right]^{\frac{1}{2}}\right). \tag{3.17}$$

While the masses of the two charged bosons are,

$$m_{H^\pm}^2 = m_A^2 + M_W^2. \tag{3.18}$$

In order to make reliable phenomenological predictions, loop corrections must be included which depend on the particle masses and free parameters of the SUSY model. Due to the large number of free parameters, searches for MSSM Higgs bosons are expressed in terms of benchmark scenarios where the lowest-order parameters tanβ and M_A are varied, while fixing the other parameters that enter through radiative corrections to benchmark values. The $m_h^{\max}$ scenario [11, 12] yields expected limits in the tanβ and M_A plane. The following parameters are fixed in the $m_h^{\max}$ scenario:

$$M_{SUSY} = 1\text{TeV}, \mu = -200\text{GeV},$$
$$m_{\tilde{g}} = 0.8M_{SUSY}, M_A \leq 1000\text{GeV},$$
$$X_t = 2M_{SUSY}, A_b = A_t$$



(a) (b)

Figure 3.4: MSSM Higgs Production diagrams

where M_{SUSY} is the third generation squark mass parameter, μ is the higgsino mass parameter, $m_{\tilde{g}}$ is the gluino mass, $X_t = A_t - \mu/\tan\beta$, A_b denotes the higgs-sbottom coupling and A_t is the trilinear higgs-stop coupling. The gaugino mass parameter, M_1, is fixed via the GUT relation,

$$M_1 = \frac{5\sin^2\theta_w}{3\cos^2\theta_w} M_2 \tag{3.19}$$

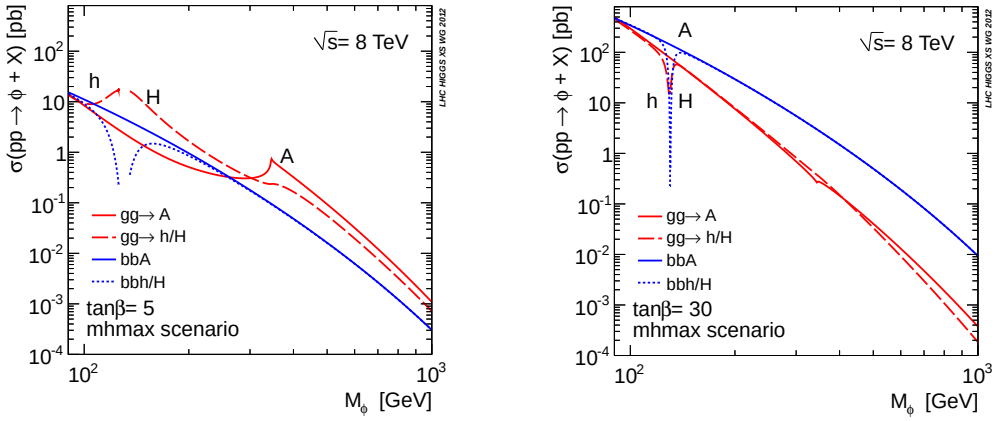
where M_2 is the SU(2)-gaugino mass parameter. Results in this thesis are interpreted both in the context of the MSSM $m_h^{\max}$ scenario and also in a model independent way, in terms of upper limits on $\sigma\cdot\mathrm{BR}(A/H/h \to \tau\tau)$ for gluon-fusion and b-associated neutral higgs boson production.

3.4 MSSM Higgs Production

The two higgs doublet model of the MSSM is responsible for interesting phenomenological effects which do not occur in the SM. The dominant production mechanisms for the higgs particles are gluon fusion and associated production of b quarks. The neutral MSSM Higgs boson production cross section for small and moderate values of $\tan\beta$ is high for gluon fusion ($gg \to A/H/h$) shown in figure (3.4a).

At large values of tanβ the b-associated production is the dominant contribution due to the enhanced bottom Yukawa coupling; therefore, associated production of b quarks with $A/H/h$ becomes a dominant signature. Production of $gg \to A/H/h + bb$ is shown in figure (3.4b). Identification of a b jet in the final state also serves to reduce further unwanted backgrounds such as $Z \to \tau\tau$.

Figure (3.5a) shows cross sections for various MSSM higgs boson production mechanisms at a value of $\tan\beta = 5$ and figure (3.5b) shows them at $\tan\beta = 30$. Performing an analysis with and without b jet associated production is useful for probing a larger region of MSSM phase space. Futhermore, due to recently improved τ identification techniques developed at CMS [13] and enhanced couplings to τ and b-quarks the search for $A/H/h \to \tau\tau$ at the LHC is of particular interest and will continue to be an important mode for MSSM discovery into the 2015 run.



(a) $\tan\beta = 5$ (b) $\tan\beta$=30

Figure 3.5: The cross sections for various MSSM higgs bosons and production mechanisms at 8 TeV [14]

Chapter 4

Large Hadron Collider

The Large Hadron Collider (LHC) is a two ring superconducting hadron accelerator and collider. It is located on the border of Switzerland and France to the northwest of the metropolitan area of Geneva. The LHC is installed in a 26.7 km long tunnel which was originally constructed from 1984 to 1989 for the Large Electron Positron (LEP) collider. While a hadron hadron collider does not have the same limitations due to synchrotron radiation as an electron positron collider, the financial benefits of building the LHC in an existing collider tunnel provided a strong motivation. The LHC is designed to collide protons at a center of mass energy of 14 TeV. The overall purpose of constructing a hadron collider with such a high center of mass energy is to explore the physics beyond the standard model.

4.1 Layout

A schematic layout of the LHC is shown in figure 4.1; the LHC follows the LEP tunnel geometry. The tunnel is 2.7 m in diameter and houses a twin-bore magnet which provides both rings in the same structure. As can be seen in figure 4.1, the

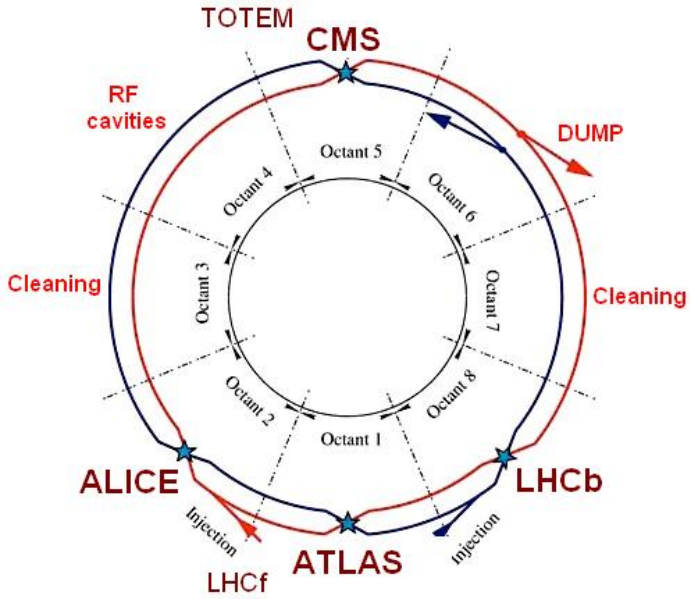


Figure 4.1: LHC Experimental and Utility Insertion Layout

LHC can be schematically divided into 8 octants. At the center of each octant is a straight section and between each of the 8 straight section there are 8 arcs. Each straight section is 528 m long and can serve as an experimental point, where a beam crossing occurs, or as a utility insertion point. The LHC is host to five experiments: ALICE is a dedicated ion experiment, LHCB is designed to study B-physics, TOTEM detects protons from elastic scattering at small angles, and two detectors designed to study a wide range of physics processes, CMS and ATLAS. CMS and ATLAS are both high luminosity experiments; the ATLAS experiment is located at Point 1 and, on the opposite side of the ring, CMS is located at Point 5. Located at points 3 and 7 are collimation systems for beam cleaning, the beam dump is at point 6 and point 4 houses an RF system for acceleration.

Each of the arcs that stretch between the straight sections are made of 23 arc cells. Each arc cell is 106.9 m in length and is comprised of two half cells each of which are 53.45 m long. A schematic of the interior of the arcs showing the dipole magnets, cryogenic system and overall layout can be see in figure 4.2.

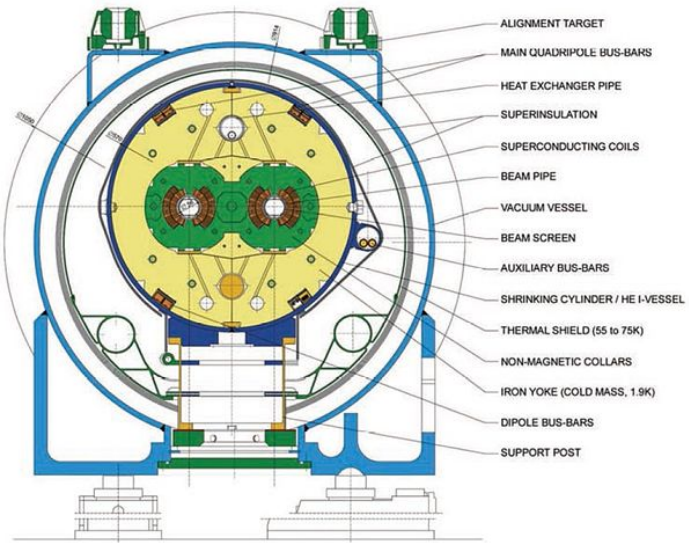


Figure 4.2: Cross section of cryodipole

4.2 Performance Goals and Constraints

The LHC was designed with the purpose of exposing the physics that lies beyond the standard model. To achieve this, proton collisions must have high energies, to explore regions of phase space that had remained out of reach at previous particle colliders, and high intensities, to search for rare processes. Considering a process with an event cross section of σ_{event} the number of events, N, delivered per second as a function of machine luminosity, L, is,

$$N = L\sigma_{event}. \tag{4.1}$$

The machine luminosity itself depends on optimization of the beam parameters,

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi\epsilon_n \beta*} F \tag{4.2}$$

Here, N_b is the number of particles per bunch, n_b is the number of bunches per beam, f_{rev} is the revolution frequency, γ_r is the relativistic factor, and F is the geometric luminosity reduction factor due to the crossing angle at the collision point. The normalized transverse beam emittance, ϵ_n, is a measure of the beams cross sectional

size and $\epsilon_n\pi$ is the total cross sectional area occupied by beam particles. F is the geometric luminosity reduction factor due to the crossing angle at the interaction point (IP),

$$F = \left(1 + \left(\frac{\theta_c \sigma_z}{2\sigma*}\right)^2\right)^{-1/2}. \tag{4.3}$$

In (4.3), θ_c is the full crossing angle at the IP, σ_z is the RMS bunch length, and $\sigma*$ is the transverse RMS beam size at the IP. A number of factors constrain the luminosity that the LHC is capable of delivering: During nominal operation a decrease in bunch intensity occurs due to collisions in the IPs. Initially, the decay time of the bunch intensity is,

$$\tau_{nuclear} = \frac{N_{tot,0}}{L\sigma_{tot}k}, \tag{4.4}$$

where $N_{tot,0}$ is the initial beam intensity, L is the initial luminosity and σ_{tot} is the total cross section and k the number of IPs. The luminosity as a function of time is then,

$$L(t) = \frac{L_0}{(1 + t/\tau_{nuclear})^2}, \tag{4.5}$$

here, L_0 is the initial luminosity. When accounting for effects due to a gradual increase in ϵ_n this results in a net luminosity lifetime of

$$\tau_L = 14.9h. \tag{4.6}$$

This further implies that it takes only ≈10h for the beam to reach $1/e$ of its initial intensity.

The maximum total integrated luminosity is then, per fill is,

$$L_{int} = L_0\tau_L\left[1 - e^{-T_{run}/\tau_L}\right] \tag{4.7}$$

Another factor which effects the integrated luminosity that can be delivered per year is turnaround time between fills. After a fill has been ended it is expected the

turnaround time needed between the end of the fill and the start of a new fill to be ≈7 hours. Therefore, if the machine is operated for 200 days per year and an optimum run time of 12 hours this leads to a theoretical total maximal delivered integrated luminosity per year of 80 fb^{-1} to 120 fb^{-1}. The operating conditions for 2011 and 2012 are outlined in section 4.3.

Refering to equation (4.2) N_b/ϵ_n is limited by the non-linear beam-beam interaction that occurs when beams collide. This beam-beam interaction is parameterized by a linear tune shift, ξ,

$$\xi = \frac{N_b r_p}{4\pi\epsilon_n} \tag{4.8}$$

where r_p is the classical radius of the proton and ϵ_n is the beam emittance. Data from previous hadron colliders show that when summed over all IPs ξ should not exceed 0.015. The transverse beam emittance ϵ_n is limited by the mechanical aperture of the LHC arcs which is given by the beam screen dimensions. The beam screen has a height of 2×17.3 mm and a total width of 2×22 mm. In terms of the RMS beam size the beam requires a minimum aperture of 10 σ; this corresponds to a nominal beam size of 1.2 mm. Finally, if this is combined with a peak β function of 180 m in the LHC arcs, the maximum acceptable transverse beam emittance is $\epsilon_n = 3.75\mu m$ and a maximum bunch intensity of $N_b = 1.15 \times 10^{11}$.

A total beam current of 0.584 A corresponds to a stored energy of approximately 362 MJ; the LHC magnet system has a stored electromagnetic energy of more than 1 GJ. At the end of a fill or in case of a malfunction the stored energy in the beam needs to be dumped. Therefore, the beam and magnet dumping systems lead to additional constraints on the maximum allowed energies and intensities.

Power loss due to synchrotron radiation is given by larmor's formula,

$$P = \frac{\mu_0 q^2 a^2 \gamma^4}{6\pi c} \tag{4.9}$$

Where μ_0 is vacuum permeability, q is the charge of the proton, a is the centripetal acceleration, and $\gamma = (1 - (v/c)^2)^{-1/2}$. This puts further requirements on the LHC design as the heat generated in this process and from luminosity-induced losses and interaction with a resistive wall must be absorbed by the cryogenic system.

In order to collide two counter-rotating proton beams separate magnetic fields with opposite dipoles are needed. To maintain good operating conditions the magnets must be identical in field strength. Due to the space constraint in the LEP/LHC tunnel twin bore magnets are used that consist of two sets of coils and beam channels within the same structure and sharing the same cooling. The superconducting technology used in the twin bore magnets is at the leading edge. The magnetic field strength in a solenoid is given by

$$B = \mu n I \tag{4.10}$$

where n is the number of turns per unit length and I is the current in the winding. To achieve the strong magnetic field needed for a charged particle collider a very high current density is needed. The high current is achieved using superconducting metal that has an extremely low resistance. The current density is then increased by using multiple windings of wires. A cable configuration known as Rutherford cables is able to minimize the self coupling of the wires while providing a higher density than simply winding or braiding the cables [15]. At the Tevatron, HERA and RHIC experiments, Niobium Titanium (NbTi) was used in the Rutherford cables and cooled to ≈4K. The LHC dipole magnets use NbTi in Rutherford cables as well but are cooled to a temperature of less than 2K. This decrease in temperature by a factor of two allows the dipole magnets to be operated at a much higher magnetic field, up to 8T. The cooling of the magnetic field is done by a cryogenic system that makes use of liquid He. With the lowering of temperature comes a reduced heat capacity of the cables themselves. Therefore, the energy deposition threshold that can raise the

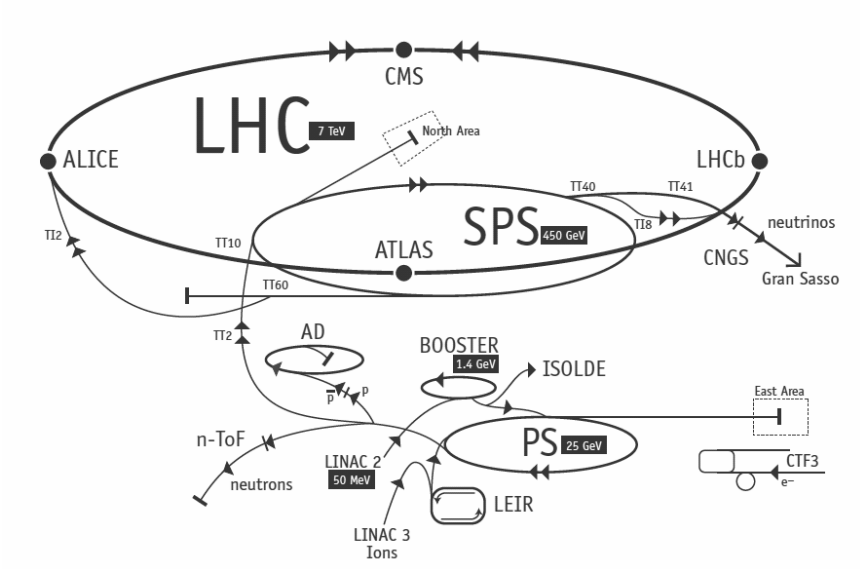


Figure 4.3: LHC and Experiments Layout

temperature of the windings above their critical temperature and cause the coil to enter its resistive state (known as a 'quench') is greatly reduced. To mitigate these issues a tighter control of current movement and heat dissipation in the cables is required.

4.3 Operation

From their creation in Linear Accelerator 2 until they reached a final energy per proton of 3.5 TeV (4 TeV) in 2011 (2012), the proton beams transverse a great distance of accelerator complex as shown in figure 4.3. Initially, hydrogen atoms are stripped of their electrons by passing them through an electric field. These protons then undergo their first round of acceleration in Linear Accelerator 2 (LINAC 2). In LINAC2 the protons are accelerated in conductors which are alternately charged positive or negative by radiofrequency (RF) cavities; this brings the protons up to 50 MeV and forms the initial bunches. After this the protons are fed into the Proton

Synchrotron Booster (PSB) which is made up of four superimposed synchrotron rings; here the energy is increased to 1.4 GeV by RF cavities. In the PSB the beams start to get squeezed and each bunch is split into 3. Next, they are injected into the Proton Synchrotron (PS) where they are accelerated to 26 GeV; each bunch is split into 4 here and, using an 80 MHz RF system, bunches are shortened so they can fit into the 200 MHz brackets of the next RF system. Before reaching the LHC, the final stage of acceleration, takes place in the Super Proton Synchrotron where they are brought up to 450 GeV. Finally, the beam is injected into the LHC where protons are ramped up until they reached a final energy of 3.5 TeV in the 2011 run. In the 2012 run the final energy was 4 TeV. The operating conditions of the 2011 and 2012 runs are further detailed in the next section.

Operating Conditions in 2011 and 2012

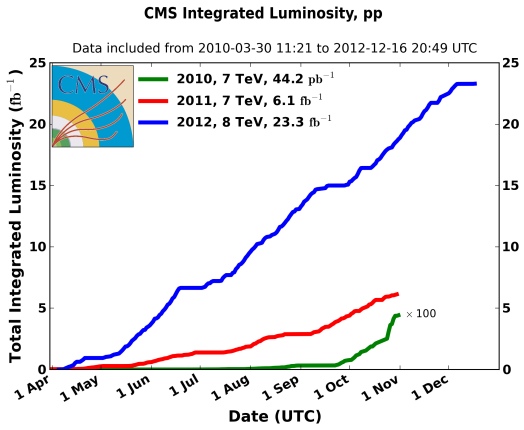


Figure 4.4: CMS Integrated Luminosity

In 2011 the LHC was operated with a center of mass energy of 7 TeV and a peak luminosity of $2 \times 10^{33} cm^{-2} s^{-1}$ While in 2012 the center of mass energy was 8 TeV and the peak luminosity was also increased to $7.7 \times 10^{33} cm^{-2} s^{-1}$. In both 2011 and 2012 the bunch spacing was 50 ns and the total number of bunches was 1380. The

total integrated luminosity delivered by the LHC detector as a function of time is shown in figure 4.4.

Chapter 5

The Compact Muon Solenoid Experiment

Located in the shadow of the Jura mountains at LHC interaction Point 5 is the Compact Muon Solenoid (CMS) Detector. The CMS detector is designed with the purpose of detecting and reconstructing proton-proton and Heavy Ion interactions delivered by the LHC accelerator. CMS is designed to collect and analyze the full physics reach of the LHC and therefore will operate at very high luminosity conditions. Due to the very high crossing rate (40MHz) and variety of physics processes, very high speed electronics that have a large set of channels capable of operating in close synchronization are required.

The CMS detector was designed to meet the following requirements:

- Good muon identification, momentum resolution asnd charge identification over a wide range of momentum
- Good charged particle momentum resolution and reconstruction efficiency within the inner tracker. Efficient triggering and offline tagging of τ-leptons and b-jets,

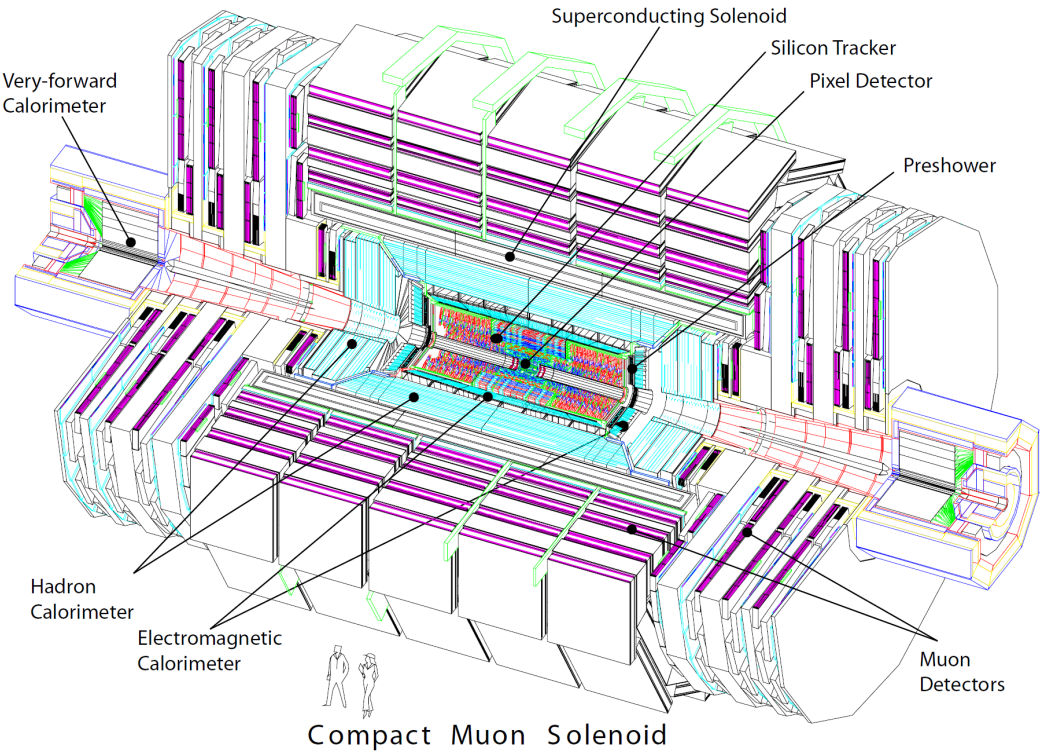


Figure 5.1: CMS Layout

requiring pixel detectors close to the interaction point

- Good electromagnetic energy resolution: diphoton and dielectron mass resolution should be ≈1% at 100 GeV

- Di-jet mass resolution and missing transverse energy measurement: requires hadron calorimeters with a wide coverage and fine segmentation

To achieve these goals CMS uses a large superconducting solenoid magnet which contains the silicon tracker (the largest silicon tracker ever built) as well as the electromagnetic and hadronic calorimeters; outside of the solenoid there is an outer hadronic calorimeter and then the iron magnetic return yoke with multiple layers of muon detectors installed within it. The layout of the detector subsystems and the overall

scale of CMS can be seen in figure 5.1. Their important design features, capabilities and limitations are detailed in the subsequent sections. These detectors and their front-end electronics must be capable of withstanding large magnetic fields on the order of a few Tesla and be able to withstand large doses of radiation (i.e radiation hard).

Finally, an important and unique aspect of CMS is its moving-ring-based structure, which is illustrated in figure 5.2. This design choice allows for very good access

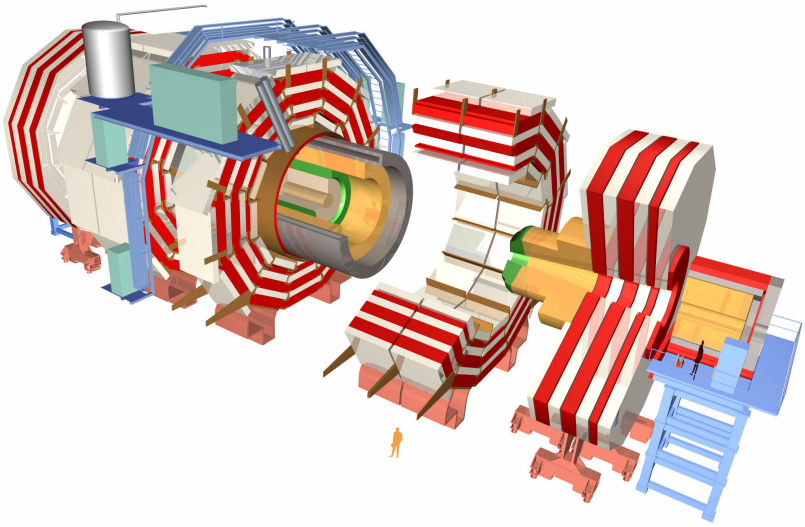


Figure 5.2: CMS moving ring structure

to the detector elements for both maintenance and upgrades which are essential to meeting future running conditions.

5.1 Coordinate System

The CMS detector is located on the LHC ring and is the experiment furthest from the CERN Meyrin site; its position on the LHC ring can be seen in figure 4.3 The CMS coordinate system is oriented such that the x-axis points to the center of the LHC ring. The y-axis points vertically upward, perpendicular to the Earth's surface, the z-axis is in the direction of the beam to the west. The angle ϕ is measured azimuthally starting from the x-axis in the x-y plane, the radial coordinate in this plane is denoted by r. The polar angle θ extends in the r-z plane and, importantly, is used in the definition of pseudorapidity, η, as,

$$\eta = -\ln \tan \frac{\theta}{2}.$$

The spatial coordinate η is preferred over the coordinate ϕ for defining the angle of a particle relative to the beam axis since the particle production in minimum bias collisions is constant as a function of η. The LHC is a hadron collider, therefore, the energy in the parton-parton interaction that initiates interesting physics cannot be known. However, it is known that the momentum in the direction transverse to the z-axis is 0. Therefore, the interesting observables (energy and momentum) are defined as transverse to the beam by measuring their x and y components and denoted as transverse momentum, p_T, and transverse energy, E_T.

5.2 Superconducting Magnet

When designing a particle detector, choice of magnet configuration drives much of the detector layout and design; large bending power is required to precisely measure the momentum of high-energy charged particles. This forced the CMS collaboration to choose superconducting technology for the magnets. The superconducting solenoid magnet used in CMS was designed to reach a 4-T field; during the 2011 and 2012 runs it was operated at a central magnetic flux density of 3.8-T to reduce the effects of aging on the coil [16]. The purpose of this magnetic field is to allow precise measurement of the momentum of charged particles, across a wide range of energies, originating from LHC collisions.

For a charged particle in a uniform magnetic field the momentum of the charged particle is given by,

$$p = qBr$$

Where p is the momentum of the particle, q is its charge, B is the magnetic field strength and r is the radius of the particle's trajectory. The transverse momentum resolution depends on the magnetic field and solenoid radius as

$$\frac{dp}{p} \propto \frac{p}{BL^2}$$

The super conducting magnet consists of two main parts, the superconducting solenoid and the return yoke. Due to structural constraints, the solenoid itself is 6.3m in diameter and 12.5 m in length, this is large enough to allow the tracker, electromagnetic calorimeter and the hadronic calorimeter to be contained within the solenoid. This is a desirable characteristic as it limits the number of radiation lengths of material between the nominal collision point and these detectors. The solenoid is composed of 4 layers of NbTi windings due to the large number of ampere-turns

required. The windings and support structure of the solenoid are detailed in Figure 5.3. The magnetic field is given by,

$$B = \mu_0 n I$$

where μ_0 is the magnetic constant, n is the number of turns per unit length and I is the current. The flux is returned by a 10,000 ton iron yoke which is composed of 11 elements, 6 endcap disks and 5 barrel wheels. The yoke also was designed to contain four muon stations. Both the solenoid cryostat system and the yoke also serve a dual purpose as a structural support for the CMS experiment.

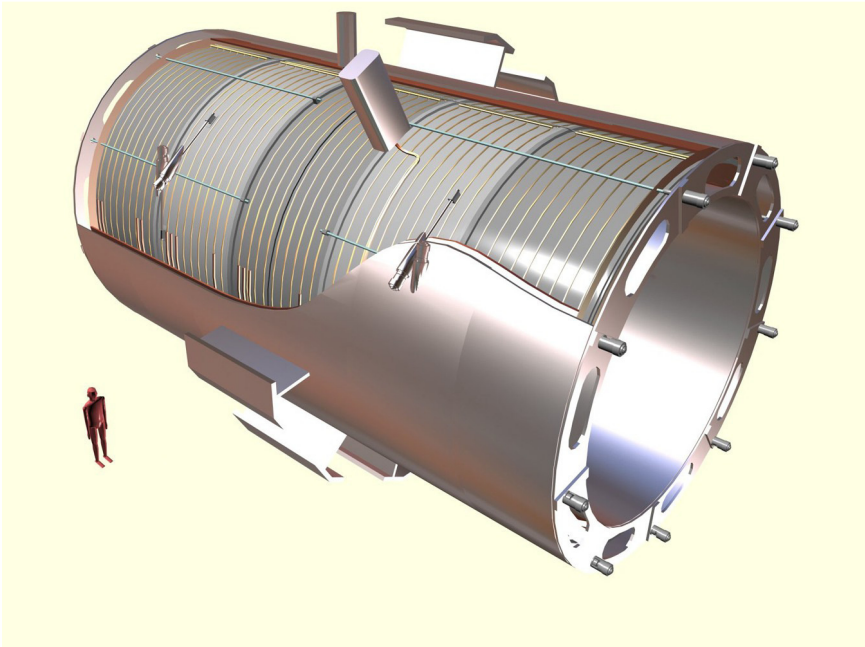


Figure 5.3: CMS Solenoid Magnet Layout

5.3 Tracking System

The tracker is CMS's inner-most detector. The purpose of the tracker is to reconstruct the trajectories of charged particles coming from the LHC collisions and measure the charged particle momenta. These charged particles leave a path in the tracker material referred to as a 'track'. These tracks are then used in the reconstruction of electrons, muons, taus, hadrons and jets and are also used to determine the primary vertex of an interaction. Additionally, the tracker can be used in the identification of displaced vertices which are located away from the primary vertex; a displaced vertex (or 'secondary vertex') is a decay signature that is often present in heavy (b or c -flavored) jets. As can be seen in Figure 5.4, the CMS tracker consists of two main detectors: an inner silicon pixel detector and an outer silicon strip detector. Efficient reconstruction of collisions require a low hit occupancy, a high hit redun-

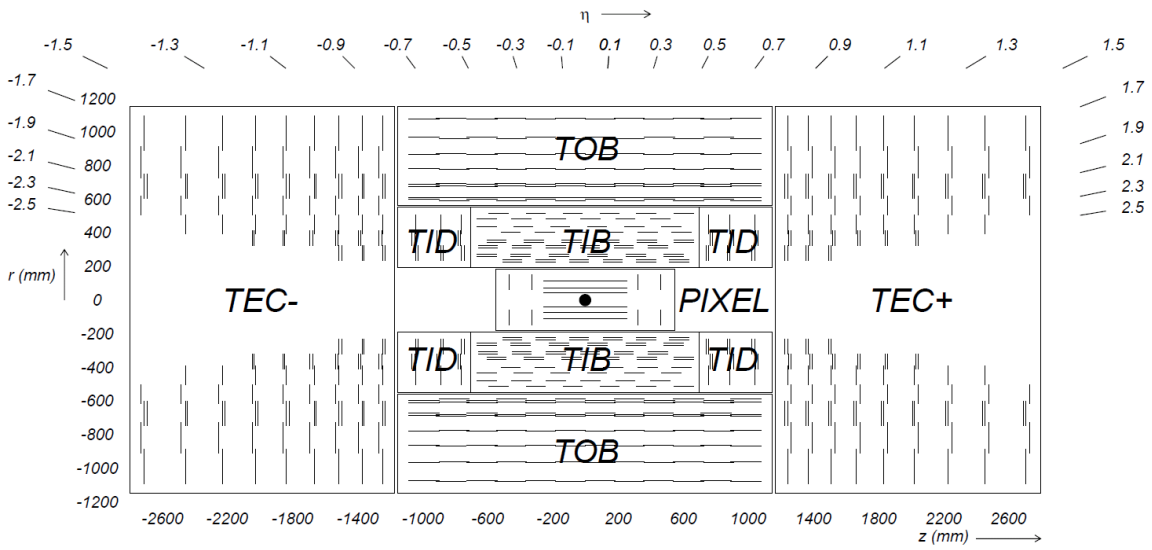


Figure 5.4: CMS Tracker layout

dancy and a fast response such that tracks can be identified reliably and attributed to the correct bunch crossing. Low hit occupancy can be achieved with high granularity while high hit redundancy requires many detector layers. These former two

requirements are only achieved with a high power density of on detector electronics which require efficient cooling. Unfortunately, this directly conflicts with the goal to limit the material budget of the tracker; interaction with material results in Coulomb scattering, bremsstrahlung, nuclear interactions and photon conversion. Finally, an extremely high particle flux results in radiation damage to the silicon sensors mainly in the form of modifications to the silicon crystal lattice. The aforementioned objectives and constraints resulted in a tracker design based entirely on silicon detector technology. At 5.4 m in length and 2.2 m in diameter, the CMS tracker is the largest inner silicon detector ever built in a high energy physics experiment.

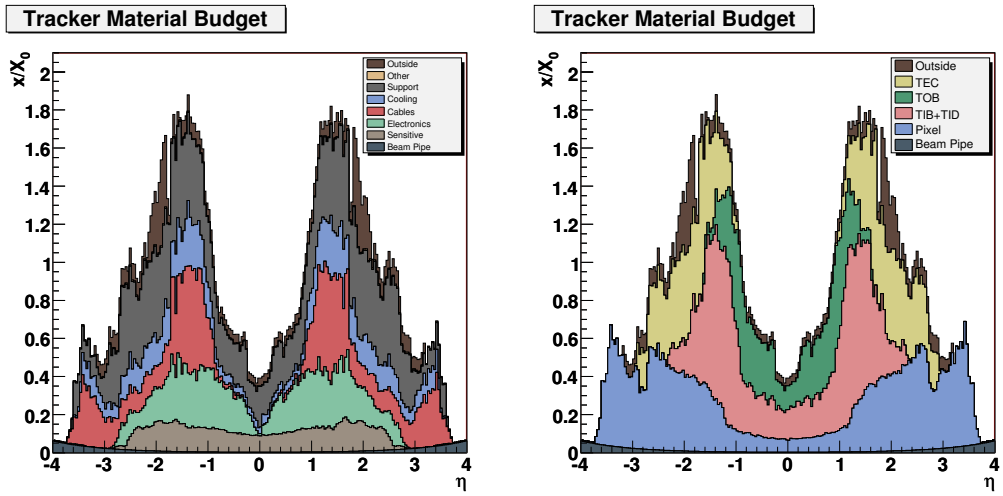


Figure 5.5: Tracker material budget, x/X_0 vs. η

The material budget of the tracker can be characterized by the rate at which a particle passing through it loses energy. The radiation length, X_0, is the mean distance over which a high-energy electron loses all but $1/e$ of its energy by bremsstrahlung and $\frac{7}{9}$ of the mean free path for pair production by a high-energy photon. Figure 5.5 shows the material budget of the CMS tracker in terms of radiation length as a function of η. Due to the location of cabling, electronics and other services, the material budget of the tracker is at a minimum of 0.4 X_0 at $\eta \approx 0$ and increases to

approximately 1.8 X_0 at $|\eta| \approx 1.4$, after which it decreases.

Pixel Detector

The pixel detector is the inner most detector of the tracking system and, covering the region from 4 to 15 cm in radius, is the closest detector to the interaction point. It has a high granularity and contributes precise tracking points in $r-\phi$ and z and is therefore responsible for a small impact parameter resolution that is important for b and c-jet secondary vertex reconstruction and τ-lepton secondary vertex reconstruction.

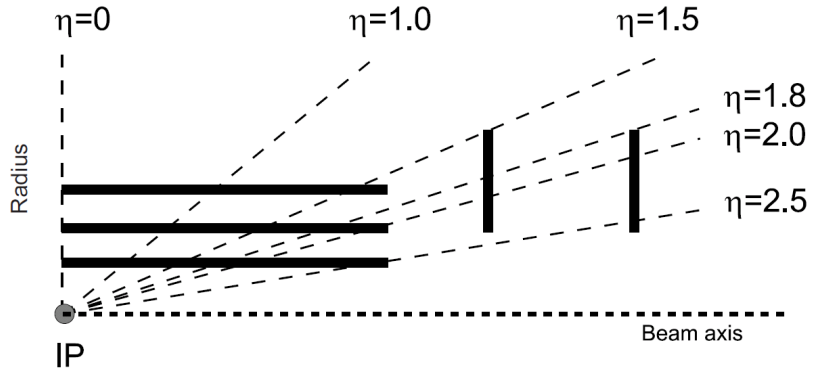


Figure 5.6: Pixel System Layout

The pixel detector is made up of individual pixel cells with a size of $100 \times 150\ \mu m^2$; it has 66 million active elements and covers a surface area of 1 m^2. The pixel detector is composed of three barrel layers and two endcap disks for which the pseudorapidity range from $-2.5 < \eta < 2.5$. The three barrel layers are located at mean radii of 4.3, 7.3, and 10.2 cm. The endcap disks extend from 4.8 to 14.4 cm in radii are located at a mean distance of $z = \pm 35.5$ cm and $z = \pm 48.5$ cm from the interaction point. As can be seen in figure 5.6 this arrangement allows for 3 tracking points over almost the full η-range of the pixel system. Due to particles entering the detector at an average angle of 20° charge-sharing between pixels is achieved which improves position resolution.

The pixel system has a zero-suppressed read out scheme with analog pulse height read-out. This improves the position resolution due to charge sharing, it helps to separate signal and noise hits as well as to identify large hit clusters from overlapping tracks. A position resolution on the order of 10 μm is achieved.

Silicon Strip Tracker

The silicon strip detector is located outside the inner pixel detector, It extends from 25 cm to 110 cm in radius and pseudorapidity up to $|\eta| < 2.5$. This region has a particle flux on the order of 100 times less than what is seen by the inner most layers of the pixel detector. It is a complementary system to the inner pixel detector and has a lower granularity. The silicon strip detector has 9.3 million active elements over a total surface area of 198 m^2 and consists of 3 large subsystems. As can be seen in figure 5.4, the Tracker Inner Barrel and Disks (TIB/TID) extend in radius to 55cm and are composed of four barrel layers with three disks at each. The Tracker Outer Barrel (TOB) consists of six barrel layers and extends to $\pm$118 cm in z. Extending beyond this in the z-direction, the Tracker EndCaps (TEC+ and TEX- where the plus and minus indicate the direction in z) are located from 124cm $< |z| <$280 cm. They are composed of nine disks which are populated with up to seven rings of radial-strip silicon detectors. The combined layouts of the pixel detector and silicon strip detector result in 8 to 14 high precision measurements of track impact points for $|\eta| < 2.4$.

The tracker resolution of muon p_T as a function of p_T and pseudorapidity are shown in figure 5.7. At high momentum (100 GeV), the p_T resolution is approximately 2-3%. The degradation at $|\eta| \approx 1$ and beyond is due to the gap between barrel and the end-cap disks and due to inferior hit resolution of the last hits of the track measured in TEC ring 7. In the barrel, tracker resolution ranges from $\approx$1-2% from 1 to 100 GeV while in the endcap tracker resolution ranges from $\approx$2-7% from 1 to 100 GeV.

High p_T muons benefit from a combined fit between tracks found in the tracker and tracks found in the muon system for an improved p_T resolution.

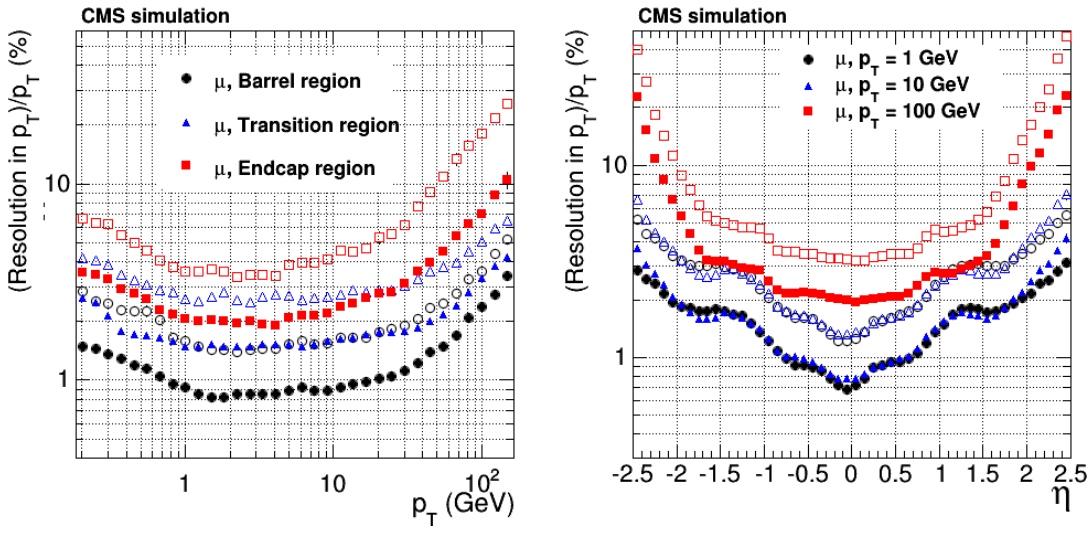


Figure 5.7: Tracker p_T resolution as a function of p_T and η. The sold (open) markers correspond to the half-widths for the 68% (90%) intervals centered on the peak of the distribution [17].

5.4 Electromagnetic Calorimeter

Directly outside of the tracking system lies the electromagnetic calorimeter (ECAL) of CMS. The driving design criteria of the ECAL is to detect and measure the decay to two photons of the higgs boson. Therefore, the ECAL is required to have a fast response time, a fine granularity and resistance to the effects of radiation. Recently advanced lead tungstate ($PbWO_4$) crystal technology was chosen. A preshower detector is placed in front of the endcap. Avalanche photodiodes (APDs) are used as photodetectors in the barrel and vacuum phototriodes (VPTs) in the endcaps. The layout of the CMS ECAL is shown in Figure 5.8. The barrel part of the ECAL (EB) covers the pseudorapidity range $|\eta| < 1.479$ the endcap part covers the pseudorapidity range $1.479 < |\eta| < 3.0$.

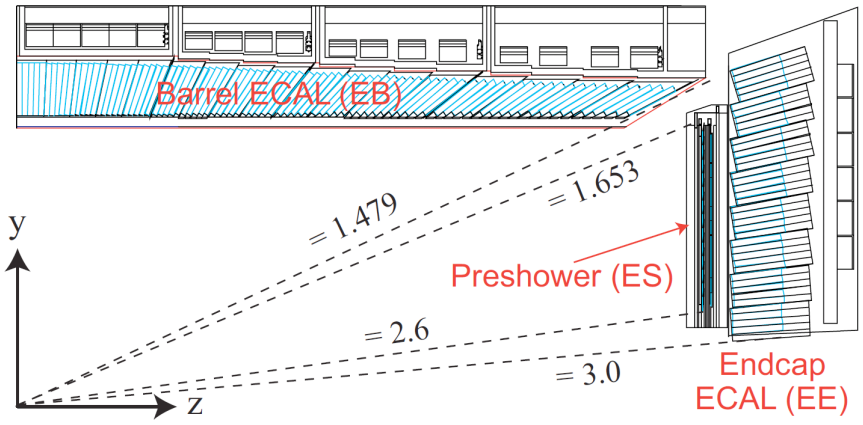


Figure 5.8: CMS Electromagnetic Calorimeter Layout

Lead Tungstate Crystals

The ECAL is composed of 75,848 $PbWO_4$ crystals: 61,200 mounted in the central barrel part, closed by 7,324 crystals in each of the two endcaps. They have a high density of 8.28 g/cm^3 and a short radiation length of 0.89 cm. The Molière radius

(radius of a cylinder containing 90% of the shower's energy deposit) is only 2.2 cm. These characteristics result in a fine granularity and a compact calorimeter. Furthermore, the scintillation decay time of these crystals is of the same order of magnitude as the LHC bunch crossing time whereby 80% of the light is emitted in 25 ns.

ECAL Energy Resolution

The energy resolution in the ECAL can be parameterized as in the following equation:

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

where S is the stochastic term, N the noise term, and C the constant term. The individual contributions are described in the following paragraphs.

The stochastic term in equation 5.4 has four primary contributions: A random event-to-event fluctuation in the lateral containment of the electromagnetic decay and subsequent photon and electron-positron production, a photoelectron statistics contribution, fluctuations in the energy deposited in the preshower absorber and dead material in front of the calorimeter. This shower containment term is expected to be 1.5% when energy is reconstructed by summing an array of 5×5 crystals and 2% when using 3×3 crystals. The noise term, N, has three main contributions: electronics noise, digitization noise, and pileup noise. The digitization and electronics noise (or Electromagnetic Interference) is a consequence of any electronic system and creates small amounts of interference in the output; this interference was measured in the test beam and found to be ≈ 40 MeV/channel. Pileup noise occurs if additional particles from pileup events reach the calorimeter causing signals that overlap. Finally, the primary contributions to the constant term, C, is non-uniformity of the longitudinal light collection, intercalibration errors and leakage from the back of the crystal.

In 2004 an electron test beam with momenta between 20 and 250 GeV was used to measure the parameters in equation 5.4 results by measuring the energy in a 3 $\times$ 3 crystal region. A typical energy resolution was found to be: The energy resolution in the ECAL can be parameterized as in the following equation:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}})^2\right) + \left(\frac{0.12}{E}^2\right) + (0.30\%)^2$$

A later test in 2006 showed a 10% improvement of the noise performance.

5.5 Hadronic Calorimeter

The CMS detector is designed to study a wide range of high energy physics processes. Measurement of hadronic jets and events which result in missing transverse energy require a hadronic calorimeter. The hadronic calorimeter (HCAL) is located outside the Electromagnetic Calorimeter but still within the superconducting magnet volume. The HCAL is a sampling calorimeter. It consists of layered sheets of scintillators interleaved with brass absorber plates. Figure 5.9 shows the placement of the four regions of the HCAL, the HCAL Barrel (HB), the HCAL Endcap (HE), the HCAL Forward (HF) and the outer HCAL (HO). The length scale of the hadronic calorimeter is the interaction length (λ_I), this corresponds to the mean free path of a hadron before undergoing an interaction.

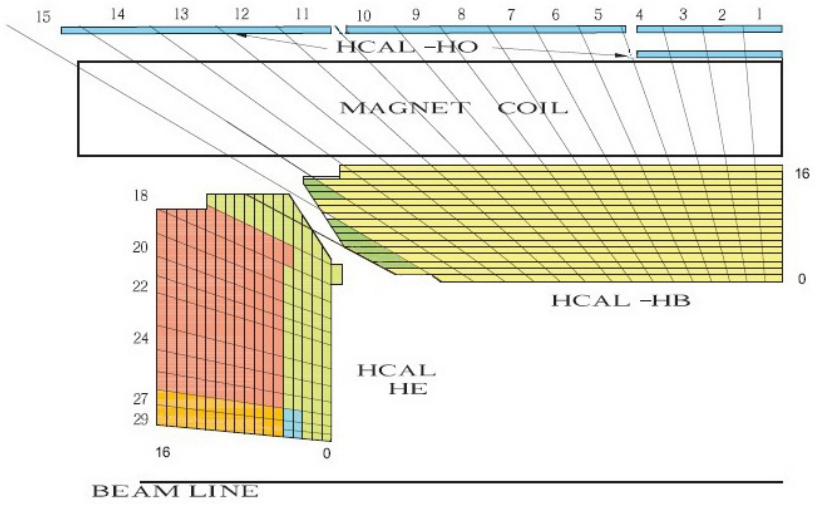


Figure 5.9: CMS Hadronic Calorimeter Layout

The barrel and endcap HCALs, HB and HE, are located within the solenoid magnet. The HB extends across $|\eta| < 1.3$. The HB absorber consists of a 40 mm thick front steel plate (which also adds structural integrity), followed by eight 50.5 mm thick brass plates, six 56.5 mm thick brass plates and and a 75 mm thick steel

plate. At an incident angle of 90° this corresponds to 5.82 λ_I while at $\eta = 1.3$ it corresponds to 10.6 λ_I. The electromagnetic calorimeter adds another 1.1 λ_I of additional material. The HE has a similar design, except the plates have a thickness of 79 mm. The plastic scintillator tiles are layered between the absorber plates. The CMS hadron calorimeter consists of approximately 70,000 tiles which are grouped into mechanical tray units to ease testing and installation. The granularity of the HCAL in $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ for $|\eta| < 1.6$ and $\Delta\eta \times \Delta\phi \approx 0.17 \times 0.17$ for $|\eta| \geq 1.6$ The HO is located outside of the solenoid; this means the magnet then acts as an absorbing medium which contributes an additional 1.4 λ_I. The HO consists of two scintillator layers that have the same granularity as the HB. The light produced in the scintillators is collected by optical wavelength shifting fibers and transferred to Hybrid Photo Diodes (HPDs). HPDs were chosen as the photodetectors due to their low sensitivity to magnetic fields.

The HF is located in the far forward region at $3 < |\eta| < 5$ and a distance of 11.2 meters from the interaction point, on average 760 GeV per proton-proton interaction is deposited into the forward calorimeters, compared to an average of 100 GeV in the rest of the detector. This environment presented a considerable challenge to the calorimetry design. Due to these environmental demands the HF is based on radiation hard Cherenkov Quartz technology which uses quartz fibers as the active medium that are embedded in a steel absorber. A signal is generated when charged particles above the Cherenkov threshold generate light that is then captured by photomultipliers. Therefore, the HF is more sensitive to electromagnetic showers and relativistic charged pions.

5.6 Muon System

Good muon detection and resolution is of central importance to the CMS detector due the muon's appearance in the final state of many important processes (for example H→ZZ→ $\mu\mu\mu\mu$ and H→ $\tau\tau$ with $\tau \rightarrow \mu\bar{\nu_\mu}\nu_\tau$). For the $\mathrm{W} + \mathrm{b\bar{b}}$ cross section measurement, which is detailed in this thesis, the W is identified via its decay to $\mu + \nu$; the $\mathrm{W} + \mathrm{b\bar{b}}$ analysis uses a single muon trigger therefore high μ detection efficiency at a low p_T and a low fake rate is essential. The muon is a relatively easy particle to detect; due to its long lifetime and heavy mass it is less affected by radiative losses and the CMS detector is capable of reconstructing muon momentum and charge over the entire kinematic range of the LHC. The muon system covers the region in pseudorapidity $|\eta| < 2.4$, consists of about 25,000 m^2 of 3 different types of gaseous particle detectors: drift tube (DT) chambers, cathode strip chambers (CSC) and resistive plate chambers (RPCs). Their selection and location are determined by the relative occupancy of muons in the detector and the uniformity of the magnetic field. For example, DTs are more sensitive to a nonuniform magnetic field, therefore, they are used only in the CMS barrel. The location of the muon subsystems can be seen in Figure 5.10.

Drift Tube System

The barrel region has a low muon rate as well as a low flux of neutrons (which are left over from hadronic decays) and a uniform 3.8 T magnetic field. Drift tube chambers with standard rectangular drift cells, which are more sensitive to these environmental effects but are relatively cheaper and easier to produce, are used in this region. They cover the pseudorapidity region $|\eta| < 1.2$ and are stationed within the iron yoke. The barrel muon detector consists of 4 concentric cylinders of drift tube chambers around

the beam line. The 3 inner cylinders have 60 DT chambers while the outer has 70 DT chambers, these combine to a total of approximately 172,000 sensitive wires. The DT Chambers are made up of three (or two) super-layers (SL), the SL is further divided into individual drift tube cells. A diagram of a drift cell is shown in figure 5.11, each drift cell contains an anode wire that is 50 μm in diameter and two electrode plates that create the drift electric field. The walls of the cell are grounded, acting as cathodes. Each cell is filled with a gas mixture of 85% Argon and 15% Carbon Dioxide; the wire and the electrodes are operated with a voltage difference of 1.8 kV. Each SL is made of four layers of drift cells staggered by half a cell. The wires of the inner SLs are aligned perpendicular to the beam line to provide a measurement of the z position of the track while the outer SLs have wires aligned parallel to the beam line to to provide a measurement in the transverse plane.

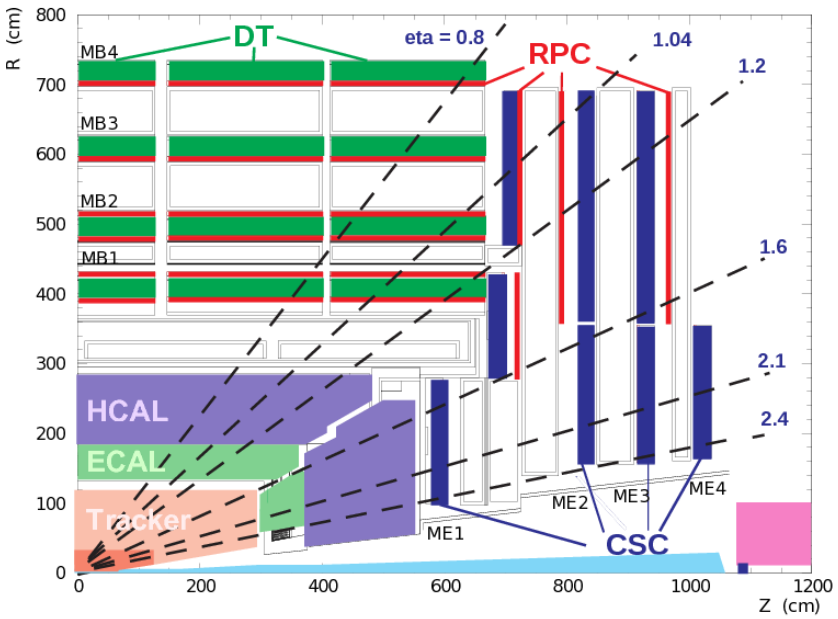


Figure 5.10: Detailed view of the placement of muon detectors within the muon system.

Cathode Strip Chambers

The Cathode Strip Chambers (CSC) are installed in the CMS endcaps. Due to their granularity and fast signal response time they are able to provide precision muon position measurement and a muon trigger signal in one device. CSCs do not require precise gas, temperature or pressure control and are also capable of operating in non-uniform magnetic fields, which make them an ideal candidate for higher ranges of pseudorapidity. As is illustrated in Figure 5.12, the CSCs consist of 7 trapezoidal panels with cathode strips which are interweaved with 6 planes of anode wires. Wires run azimuthally and define a track's radial component. The path of a charged muon is found by interpolating charges on the cathode strips by the avalanche of positive ions which is catalyzed by the charged muon; the signal is then generated by dipole moment of the ionized atoms when electrons are pulled off. Muons from $1.2 < |\eta| < 2.4$ cross 3 or 4 CSCs. In total the CSC system provides a combined 5000 m^2 of sensitive planes and has over 2 million wires. The trapezoidal CSCs cover either 10^{ϕ} or 20^{ϕ} in ϕ with overlap to form continuous ϕ coverage (excluding the ME1/3 ring).

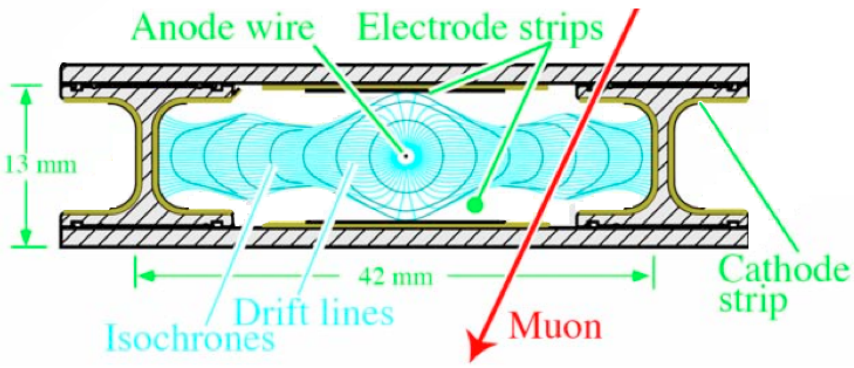


Figure 5.11: An Individual Drift Cell of the Drift Tube Chamber System

Resistive Plate Chambers

Resistive plate chambers (RPCs) are installed up to $|\eta|$ <1.6. RPCs are much faster than the 25ns bunch crossing time and, therefore, are used for triggering as well as muon kinematic measurements. As can be seen in Figure 5.13, RPCs are composed of two parallel plates which are separated by gas and they are operated in avalanche mode with readout strips in between. RPCs need intensive monitoring of temperature, humidity and pressure to ensure stability of conditions for proper operation. During 2011 and 2012 the CMS detector had six layers of RPC chambers in the barrel iron yoke, 2 located in each of the first and second muon stations and 1 in each of the last stations.

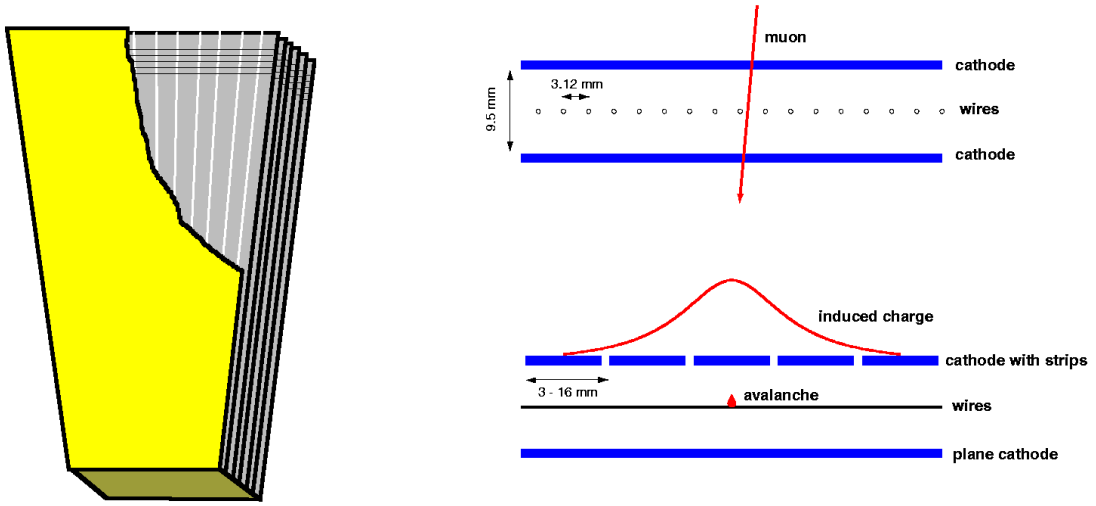


Figure 5.12: Front view of CSC and cross-sectional view of CSC

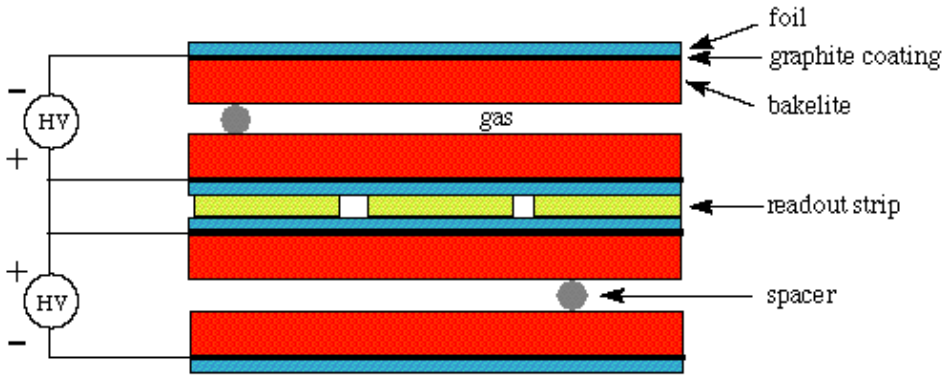


Figure 5.13: Layout of a Double Gap RPC

5.7 Trigger

The LHC delivers proton-proton collisions at the beam crossing interval of 25 ns, this corresponds to a crossing frequency of 40 MHz. The data for each crossing occupies 0.5 to 1 megabytes of storage space. As it is responsible for all data acquisition, the trigger system is the most important subsystem of CMS.

The CMS trigger system is divided into two parts: the Level 1 trigger and the High Level Trigger (HLT). The Level 1 trigger uses coarsely segmented data from only the calorimeters and the muon systems to make initial data selections while holding high resolution data in pipelined memories in front-end electronics. The Level 1 Trigger hardware is implemented in customized programmable memory look up tables (LUTs), FPGAs and ASICs. The Level 1 trigger must reduce the maximum proton interaction rate of 10^9 Hz by at least a factor of 10^6, resulting in a final output of approximately 100kHz, before high resolution data is passed to the HLT. The HLT has access to the complete readout data of the event and can use complex algorithms to filter events. For particularly interesting events, the algorithms used in this process are often similar to what is done in offline analysis.

Level 1 Trigger

The purpose of the Level 1 trigger is to reduce rates from the input crossing rate of 40MHz to a maximum output rate of 100kHz. As can be seen in figure 5.14 information from the muon trigger remains separate from the calorimeter trigger until it is combined in the Global Trigger. The level 1 trigger is described in more detail below.

The first step in the calorimeter trigger process is to create calorimeter trigger primitives using a trigger primitive generator (TPG). For the purpose of providing

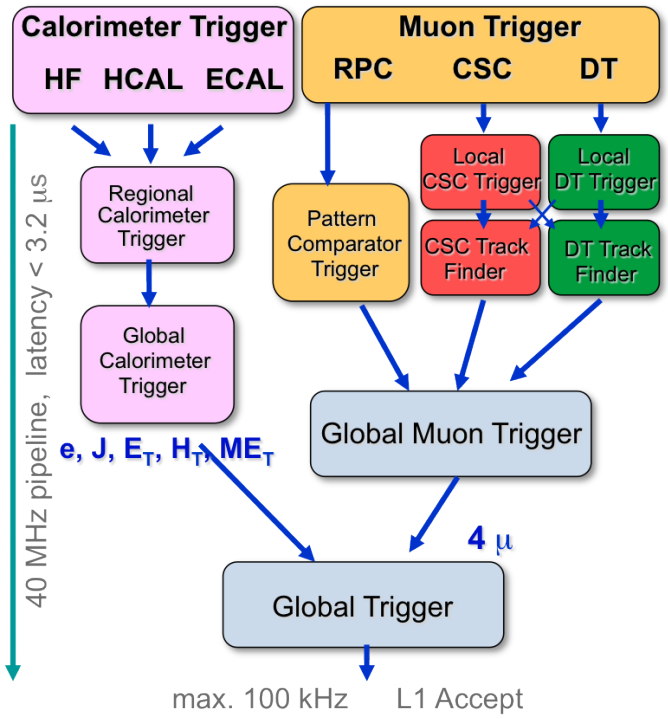


Figure 5.14: Level 1 Trigger Architecture

compact, fast information the calorimeters are divided into trigger towers; the TPG sums the transverse energy the ECAL crystals or HCAL read-out towers to obtain a trigger tower E_T. For $|\eta|$ <2.1, trigger towers have a granularity in $\eta \times \phi$ of 0.087×0.087, the granularity decreases at higher η values. The TPG electronics are integrated with the calorimeter readout and the trigger primitives are passed to the Regional Calorimeter Trigger (RCT) using high-speed serial links. An electron or photon candidate is expected to be narrow in η and broader in ϕ. An additional bit, known as a fine-grain veto bit, is set for each trigger tower to indicate whether an electron candidate's highest two adjacent strips in the tower contain less than 90% of the total E_T. After that, the RCT combines tower information to form electron and photon candidates. The RCT also sums ECAL and HCAL towers into broader regions which are used in the GCT to form jets. These regions also have a τ-veto bit which determines how compatible a region is with a τ lepton. After the GCT has

combined regions to form jets, the GCT then sorts the electron/photon candidates and jets and passes them onto the Global Trigger.

The Muon Trigger uses all three muon systems (DTs,RPCs,CSCs) for triggering. The DT system has fairly good timing resolution and is used to reconstruct tracks and associate tracks to bunch crossings. At a regional level, the DT and CSC track finders combine segments to build muon tracks and assign transverse momentum using LUTs. All regional information from the three muon systems are then forwarded to the Global Muon Trigger (GMT) where it is then combined to proved track information with equal or better resolution than the regional systems. The tracks are then ranked and forwarded to the Global Trigger (GT).

Regional Calorimeter Trigger

The Regional Calorimeter Trigger is responsible for processing the level 1 trigger data of the Electromagnetic and Hadronic calorimeters in barrel, endcap and the HF forward detector. The RCT consists of 18 crates installed in nine racks in the CMS underground counting room, which is adjacent to and shielded from the CMS cavern. For triggering purposes, the HCAL and ECAL are subdivided into regions in (η, ϕ). As illustrated in figure 5.15 a region consists of of 4×4 trigger towers; the RCT has 7 receiver cards per crate and one jet summary card for each of the 18 crates. A total of 24 bits is received from each trigger tower: two 8-bit calorimeter energies each from the ECAL and the HCAL, two energy characterization bits, 5 bits of error detection and 1 LHC bunch crossing bit.

Good e/γ selection efficiency at a low rate is essential. e/γ identification in the RCT takes place as follows: starting from the trigger towers the RCT sums of ECAL and HCAL energies and also calculates the the ratio of HCAL to ECAL energy, (H/E). The (H/E) ratio is used to discriminate electrons and photons from pions and

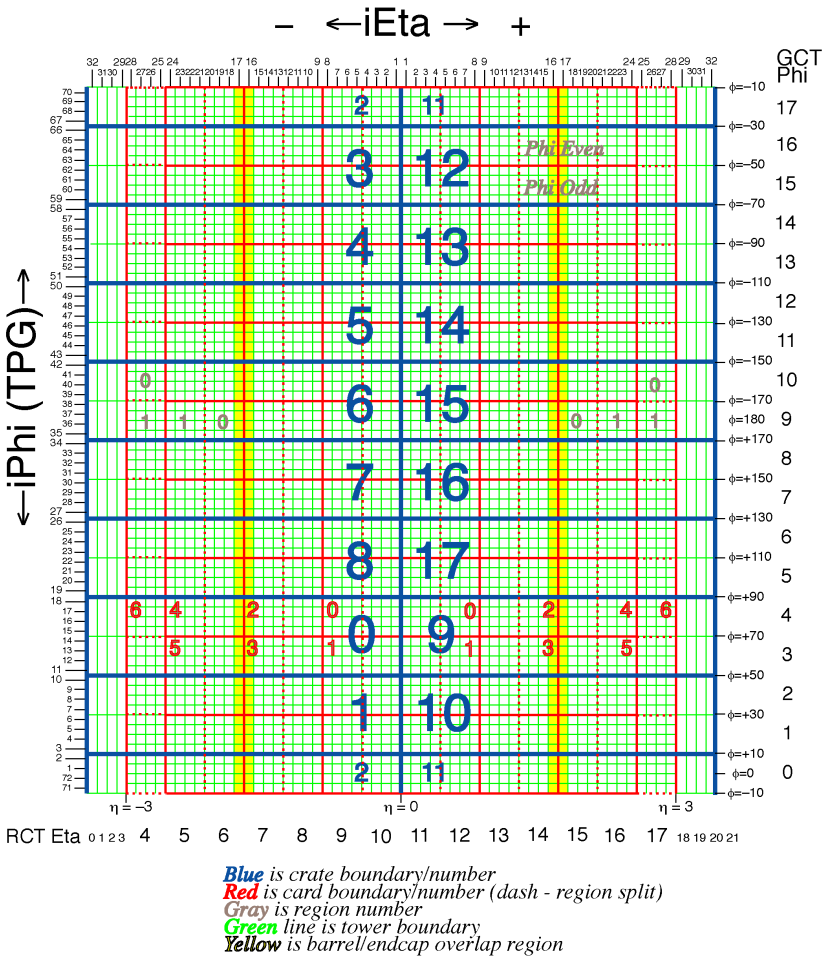


Figure 5.15: RCT layout

electromagnetic deposits in jets. The RCT e/γ algorithm is applied to ECAL trigger towers that have an energy higher than their four immediate neighbors. The seeding tower must also have an energy deposit which is contained within 2×5 crystals, this is known as the fine grain requirement and, if this is true, a fine grain bit is set by the TPGs. Then, the tower energy is combined with the nearest highest energy neighbor, this accounts for particles that deposit energy in two towers; this sum is then associated with the e/γ candidate. To further separate from jets, the RCT also determines if the e/γ candidate is isolated. As can be seen in figure 5.16. the 8 towers around the central are divided into L-shaped quiet corners, for an electron

to be considered isolated two of these must have an energy below a configurable threshold; those towers must also not have the fine grain bit set. The RCT then ranks the isolated and the non-isolated e/γ candidates and passes the highest four to the GCT. These collections of isolated and non-isolated candidates are mutually exclusive. Due to having a sufficiently low rate, all isolated and electrons with a transverse momentum greater than 63.5 GeV are accepted as isolated for the 2011 and 2012 runs.

For the purpose of jet reconstruction, the RCT also sums ECAL and HCAL energies in clusters of 4×4 towers (as can be seen in figure 5.15) in to regions. During this process the RCT examines the energy profile to see if the region is compatible with a hadronic τ lepton. A τ-veto bit is set unless the pattern of active towers corresponds to at most 2×2 contiguous towers. The patterns are detailed in figure 5.17 This τ veto bit is included when the jet transverse energy is sent to the GCT. The RCT is an essential system for data acquisition at a high efficiency and a low

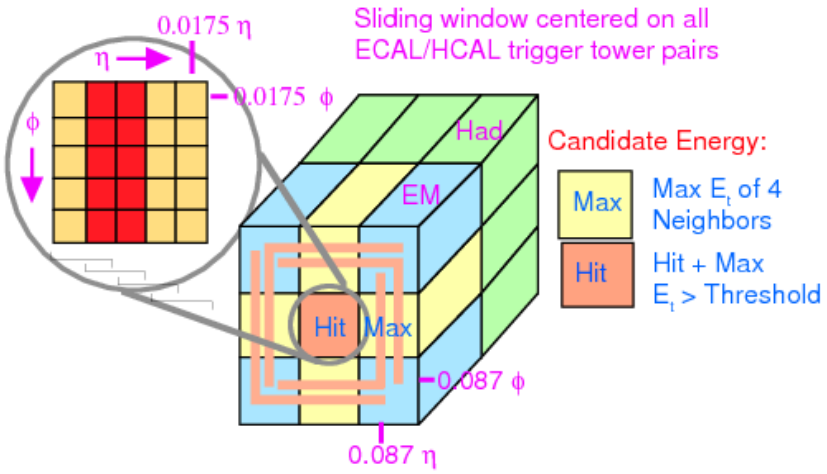


Figure 5.16: The fine grain veto bit is set if an e/γ deposit is not contained within a 2×5 ECAL crystal area (as can be seen in red). An Isolated electron is required to have a at least twi of the L-Shaped corners (in tan) to below a certain threshold

rate, therefore good performance must be continually monitored during and between runs. The RCT emulator as a high level software package designed to emulate RCT hardware response given any online input. Using the RCT emulator, the performance of the RCT is monitored during and after data-taking by constantly comparing the trigger output with the expected output from the RCT Emulator. The RCT data quality monitoring software package runs the RCT emulator on a subset of events and compares the emulator response.

High Level Trigger

After events pass the Level 1 trigger they are sent to the High Level Trigger (HLT). The HLT uses a filter farm of processors to reduce rates by a factor of 10^6 from 100 kHz to a final output rate of less than 300 Hz. This filter farm takes as input data that is processed by the Data Acquisition (DAQ) system. The DAQ system is responsible for taking data which is pushed into the DAQ by Front-End Drivers (FEDs). All sub-detectors at CMS operate with analog outputs, the FEDs are responsible for data digitization and signal output. The maximum input rate to the DAQ system is

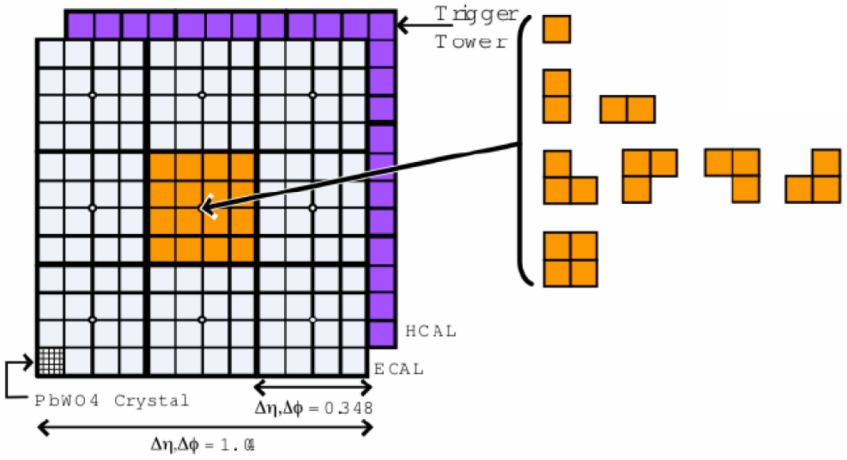


Figure 5.17: RCT Tau Identification Algorithm

100kHz which corresponds to a data flow of ≈100 GByte/s from approximately 650 data sources. An event builder assembles the event fragments from all FEDs which belong to the same Level 1 output; this data is then sent from underground to the surface building, SCX where the HLT is located. Algorithms for data selection are implemented in the HLT are written in C++ and are similar to or exactly the same as algorithms that are used in offline selection. This further helps to ensure a high efficiency of interesting physics events. The HLT processing time depends primarily on the algorithm complexity, therefore, simple algorithms are processed first and if an event fails a selection step then a node is free for the next event. In this way, the HLT is able to optimize its processing capacity.

Chapter 6

Event Reconstruction

The raw data format is used for online reconstruction in the high level trigger or offline reconstruction. In the following section offline event reconstruction is described; during this stage a number of algorithms are used to identify and construct the kinematic properties of tracks, track vertices, electrons, muons, taus, jets, missing transverse energy and event variables which aid in physics analysis. Where appropriate, configuration for HLT is indicated.

6.1 Track and Primary Vertex Reconstruction

Good track reconstruction provides a means for very accurately determining the momentum of charged particles. Identification of the primary vertices, signifies a proton-proton interaction in the detector, and measurement of the sum of their associated tracks, provides discrimination of the hard scattering process from pileup vertices. Furthermore, using tracker tracks to search for secondary vertices is crucial for accurate identification of heavy flavored jets. Tracking and vertexing of charged particles is a crucial part of CMS reconstruction and physics analysis and is described in the

following sections.

Track Reconstruction

Track reconstruction at CMS follows multiple iterations of a track finding sequence called the Combined Track Finding (CTF) sequence [18]. The CTF sequence consists of seed generation through hit clustering, track finding via a filtering technique, track trajectory fitting and finally selection of tracks that pass quality requirements. In CTF, the first iteration reconstructs tracks that have the highest quality. Subsequent iterations reconstruct tracks with gradually less stringent requirements to reconstruct tracks which are lower p_t, have missing hits or which are greatly displaced. After each iteration of the CTF track reconstruction sequence, the hits associated with tracks are removed. This reduces the combinatorial complexity and simplifies subsequent iterations.

The CTF track reconstruction proceeds as follows:

- First, seeds are generated which provide the initial track candidates. Charged particles follow helical paths in the magnetic field. Therefore, five parameters (including curvature) are needed to define a trajectory. To determine these five parameters at least 3 hits (or 2 hits in the pixel detector and a constraint on the origin of the track trajectory from the beam spot) are required. Seeds are built in the inner part of the tracker and track candidates are reconstructed outwards. The choice to begin seeding in the central region of the tracker and then move outwards towards the endcaps is due to a finer granularity and hence lower occupancy per sensor in the center of the pixel detector. Each iteration of CTF uses independent quality parameters for seeding layers.

- Next, track finding is performed based on the Kalman filter method [19]. Track

finding starts by using the seed trajectories to define and then search for adjacent layers of the detector with a hit. After successful identification of adjacent layers with valid hits the track finding algorithm updates the trajectories of the tracks.

- Track fitting is then performed. Constraints, such as a beam spot matching requirement, applied during the track finding stage can introduce bias to the track trajectory. Track fitting removes bias and provides full information about the track kinematic trajectory[17]. The trajectory is refitted using a Kalman filter and smoother with a Runge-Kutta propagator that takes into account both material effects and accommodates the inhomogeneous magnetic field.

- The final step is track selection, where tracks are required to pass a number of quality based selection criteria. The selection criteria is placed on a track's number of layers with valid hits, the fit-based χ^2/dof, and the track's compatibility with a primary vertex. In addition to these, several requirements are imposed as a function of track p_T, η, and the number of layers with valid hits.

The performance of the track reconstruction is evaluated in $Z \rightarrow \mu\mu$ events using a 'tag and probe' method whereby a well identified 'tag' muon track is found and the efficiency of reconstructing a second 'probe' muon is measured; the reconstruction efficiency of the 'probe' muon is shown in figure 6.1.

Primary Vertex Reconstruction

Primary vertices (PVs) are reconstructed in order to locate and determine the associated uncertainty of each proton-proton interaction vertices regardless of whether it is a 'signal' or 'background' vertex. Primary vertex reconstruction proceeds in three steps. The first step is to select tracks based on their association with a primary

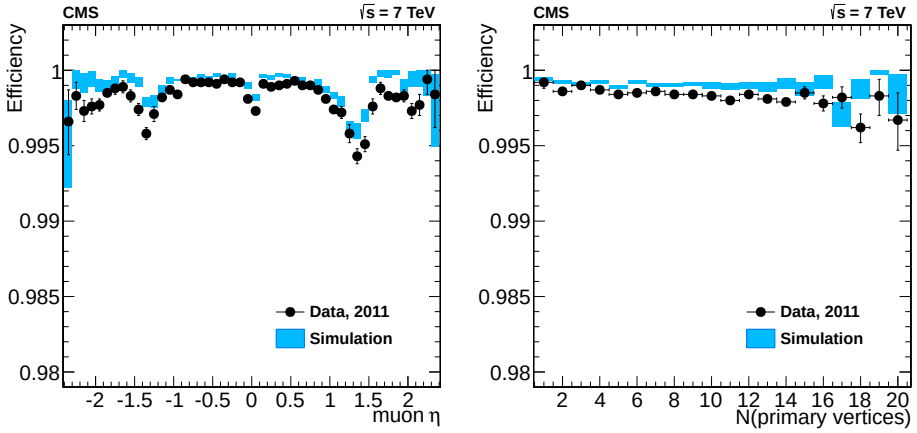


Figure 6.1: Tracker efficiency of muons in $Z \to \mu\mu$ decays. Measured using Tag and Probe as a function of η (left) and number of vertices (right)[17]

interaction region. To do this, a number of quality selections are imposed, based upon the significance of the transverse impact parameter (d_{xy}), the number of strip and pixel hits that are associated with a track and the normalized χ^2 from the fit to the trajectory. In selecting tracks, there is no requirement on the p_T of the track; this is important so that all PVs including ones from minimum bias events are reconstructed. The second step is to cluster the selected tracks based on their z coordinate at their point of closest approach to the beam spot. This is done using a Deterministic Annealing (DA) algorithm which finds a global minimum given many degrees of freedom [20]. The DA algorithm is based upon an algorithm developed for physical chemistry processes whereby a system can be driven to its lowest energy state by a gradual reduction of temperature, or annealing. The DA algorithm begins by calculating a global χ^2 whereby each reconstructed track is assigned to a vertex prototype. The assignment of track to vertex is given a weight between 0 and 1 based on the compatibility of vertex and track. This assignment is controlled by a temperature parameter where an infinite temperature means all weights are equal and T=0 means all the weights are 0 or 1. The clustering starts with a single prototype vertex at high

T and in subsequent iterations the prototype vertex is split into multiple vertices. If all tracks are compatible with a single vertex then two prototype vertices will become one. Each prototype vertex becomes a vertex candidate with all associated tracks that have a weight greater than 0.5. The third and final step is to take candidate vertices based on DA clustering in z and use an 'adaptive vertex fitter' to compute vertex parameters [19]. These parameters include the 3-D position and covariance matrix, as well as indicators for the success of the fit such as the number of degrees (n_{dof}) of freedom for each vertex and weights of the track used in each vertex. The adaptive vertex fitter uses a modified definition of n_{dof} where,

$$n_{dof} = -3 + 2 \sum_{i=1}^{nTracks} w_i, \tag{6.1}$$

Here, w_i is the weight of the i^{th} track. This implies that n_{dof} is strongly correlated with the number of tracks that are compatible with arising from the interaction region which means that n_{dof} can also be used to select true proton-proton interactions. For a vertex to be selected at the analysis level, it is required to have a z position smaller than 24 cm with respected to the origin of the detector, a ρ position less than 2 cm from the IP and more than 4 degrees of freedom. Of the vertices that pass this criteria the signal primary vertex is defined as the one that maximizes

$$\sum_i p_{T,i}^2, \tag{6.2}$$

where the sum extends over all associated tracks and $p_{T,i}$ is the transverse momentum of the i^{th} track.

6.2 Electron ID and Reconstruction

Electrons in CMS are reconstructed using tracker or ECAL seeds. Electrons with a low transverse momentum have a higher efficiency for reconstruction using tracker

seeds; whereas electrons with a higher transverse momentum have a higher efficiency if reconstructed using ECAL seeds. When traversing the tracker, electrons pass through material equivalent to 0.4-0.8 X_0 this causes them to lose a significant portion of energy to radiating bremsstrahlung photons. This will in turn cause the electron energy deposit to spread in ϕ. Accurate measurement of electron energy at the initial interaction point requires collection of photons produced via bremsstrahlung.

To begin reconstruction, clusters are formed in the ECAL. Next, electron reconstruction requires seeds to be produced using either an ECAL driven (1) or a tracker driven approach (2). (1) is more efficient for electrons with $E_T > 4$ GeV; in this method, the clusters produced in the ECAL are extended in ϕ to form super clusters (SCs). SCs are then selected and matched back to tracks in the inner tracker layers. Approach (2) uses hits in the pixel detector to seed low energy electrons; in this case the energy deposit in the calorimeter is very broad. All reconstructed tracks are considered as seeds and the bremsstrahlung hypothesis is tested by extrapolating a straight line from the track position to the corresponding ECAL cluster. The process is repeated for all layers and a supercluster is defined by summing all linked electromagnetic cluster deposits.

Next, trajectories are reconstructed using a dedicated model of the electron energy loss and fitted using a Gaussian Sum Filter (GSF) algorithm [21]. For electron reconstruction, the GSF algorithm is preferred over the typical Kalman Filter algorithm. This choice is motivated by the fact that bremsstrahlung energy loss, as described by Bethe and Heitler, does not following a Gaussian distribution. So, while the Kalman Filter algorithm is optimal in cases where all probability densities encountered during track reconstruction are Gaussian, the GSF algorithm distributions are weighted sums of Gaussians and appropriately describes electron energy loss. The electron trajectory builder constructs all possible trajectories for a given seed and the best fit

is chosen using a χ^2 approach. Finally, a trajectory smoother is applied.

To improve the electron selection and reject the large QCD backgrounds in the di-tau analysis a Boosted Decision Tree (BDT) based discriminator is used. The training has been performed in two bins of p_T and three bins of η as shown in Table 6.1. The BDT has the following 19 variables as input:

- The normalized χ^2 of the common track fit, the number of valid hits in the track fit, the normalized χ^2 of the *GSFTrack* fit.

- The distance in η ($\Delta\eta_{SC}(\text{Track}_{vtx})$) and ϕ ($\Delta\phi_{SC}(\text{Track}_{vtx})$) between the reconstructed super cluster in the calorimeter and the track evaluated at the primary vertex position, the distance in η between the super cluster seed and the track evaluated at the calorimeter surface.

- The cluster shape variables $\sigma_{i\eta,i\eta}$ and $\sigma_{i\phi,i\phi}$, where $i\eta$ ($i\phi$) indicate the integer label of the electromagnetic calorimeter cell in η (ϕ), the cluster shape variable $f_e = 1 - e1X5/e5X5$, where $e1X5$ ($e5X5$) indicate the energy deposition in an array of 1×5 (5×5) cells in the vicinity of the super cluster seed, the cluster shape variable $R9 = e3x3/E_{SC}$, where $e3x3$ and E_{SC} indicate the energy in an array of 3×3 cells in the vicinity of the super cluster seed and the raw energy of the reconstructed super cluster.

- The ratio of the hadronic energy over electromagnetic energy of the super cluster (H/E), the ratio of the super cluster energy over the momentum of the associated track evaluated at the selected primary vertex (E/P), the variable $1/E_e - 1/p_e$, where E_e and p_e indicate the reconstructed energy and momentum of the electron candidate, the ratio of the electron cluster over the momentum of the associated track and the ratio of the seed cluster over the associated track,

where each time the track momentum has been evaluated at the surface of the calorimeter.

- The ratio of the energy that has been reconstructed in the pre-shower detector over the raw energy of the reconstructed super cluster. The momentum and η of the reconstructed electron candidate.

An electron is considered as well identified if the BDT discriminator falls above the thresholds shown in Table 6.1. In addition the electron candidate is required to have a distance from the selected primary vertex of $d_z < 0.1$ cm along the z direction of the experiment and $d_0 < 0.02(0.045)$ cm in the plane perpendicular to z in the $e\mu$ ($\mu\tau_h$ / $e\tau_h$) decay channel. Furthermore, there should be no missing hits in the inner layers of the pixel detector, no hits before the selected primary vertex and a vertex fit probability of more than $P > 10^{-6}$ to minimize the probability that the electron candidate originates from a photon conversion.

	BDT Discriminator Value (>)		
	$\|\eta\| < 0.8$	$0.8 \leq \|\eta\| < 1.479$	$1.479 \leq \|\eta\|$
$p_T \leq 20$ GeV	0.925	0.915	0.965
$p_T > 20$ GeV	0.925	0.975	0.985

Table 6.1: Thresholds for the BDT discriminator to identify electrons. For electrons with $p_T > 20$ GeV the values in braces correspond to the Tight ID working point.

Rejection of Electrons from converted photons

In the di-tau analysis a non-negligible background contribution is due to γ + jets production where a high energy photon converts to an electron-positron pair and a jet is misidentified as a hadronic tau. To reject this background electrons from photon conversions are identified and then a veto is placed on these electrons. Photon

conversions are reconstructed by combining opposite sign track pairs and performing a vertex fit of those tracks to identify conversions. To reject electrons coming from conversions, a requirement is placed on the minimum number of hits in the pixel detector given the track position and direction.

6.3 Muon ID and Reconstruction

Muons at CMS are reconstructed using information from both the tracker and the muon detectors. Muons which come from the decay of an on shell W boson are typically higher in transverse momentum than muons which are from tau semi-leptonic decays. Three muon reconstruction approaches are used: standalone, global and tracker muon reconstruction. Depending on the energy of the muon the reconstructed trajectory is either solely dependent on the tracker or dependent on the tracker and the muon system.

Standalone Muon Reconstruction

Standalone muon reconstruction uses only tracks from the muon system to reconstruct tracks using a Kalman filter technique which is seeded by track segments or Level-1 trigger electronics. Tracks are propagated in iterative steps taking into account the magnetic field, muon energy loss in the material, multiple scattering and missing hits in the muon system. Next a suitable χ^2 cut is applied to reject bad hits due to showering, delta rays and pair production. A backward Kalman filter is then applied, working from outside in and finally the track is extrapolated to the beam-spot and a vertex-constrained fit to the track parameters is performed.

Global Muon Reconstruction

Global muon reconstruction matches standalone tracks to tracks in the tracking system. Tracks are selected which roughly correspond in momentum and position to the standalone muon tracks. This is performed in two steps: First, tracks are selected in a defined $\eta \times \phi$ region which is centered on the standalone track. Next, spatial and momentum matching is used to select the best matching track. Compatibility of the standalone muon track and tracker track with the primary vertex is also required. Finally a new 'global track' is created combining tracker and muon hits; at this stage, no new hits are selected, instead, the selected hits are refitted as a global track. If more than one candidate track pair is matched then the candidate with the best χ^2 value is selected. For muons with $p_T < 200$GeV the p_T measurement is driven by the tracker resolution, for muons with $p_T > 200$GeV, the global-muon fit can improve the momentum resolution compared to the tracker-only fit. In muons with a higher p_T, as might be found in the boosted topologies of $\mathrm{W} + \mathrm{h} \rightarrow \mathrm{b\bar{b}}$ or $\mathrm{W} + \mathrm{b\bar{b}}$ it is useful to require that muons are globally reconstructed.

Tracker Muon Reconstruction

In tracker muon reconstruction, all tracks with $p_T > 0.5$ GeV and $p > 2.5$ GeV are considered as possible muon candidates. Track reconstruction is outlined in Section 6.1. After tracks are reconstructed their trajectories are extrapolated to the muon system. If there exists one hit in the muon system which 'matches' the extrapolated track then that track is defined as a tracker muon. A 'match' requires the distance in x between the segment and the extrapolated track is less than 3 cm or that the pull for x is less than 4. The pull is defined as the difference in the position of the matched segment and the position of the extrapolated track divided by the sum of

the uncertainties on the position of the matched segment and the position of the extrapolated track.

6.4 Electron and Muon Isolation

To discriminate signal muons and electrons from leptons which are created in QCD interactions an isolation requirement is essential. As pileup of interactions in the detector increases, the performance of standard combined relative isolation, which sums the energy deposited by all PF candidates in a cone of $\Delta R = 0.4$ around the central lepton and divides the sum by the p_T of the candidate, degrades. Charged PF particles are associated with a vertex using the deterministic annealing algorithm; non charged PF particles are not associated with any vertex by the vertex algorithm, instead, they are associated with vertices by using their closest distance in the z axis after they are extrapolated to the beam line. Using this separation algorithm, charged particles can be properly associated with a given primary vertex and used to calculate isolation. However, this algorithm does not properly account for neutral particles produced in pile-up interactions. Therefore, a specific correction, known as a $\Delta\beta$ correction, is used to account for the neutral energy from other interactions. A charged particle's transverse momentum sum is created by summing over the charged particles inside the isolation cone of the lepton while requiring that those charged particles do not originate from the primary vertex. The charged particle sum is converted into an expected neutral deposit by assuming that the average charged to neutral particle ratio is 2:1. A relative combined isolation variable is then defined as:

$$I_{\text{rel}} = \frac{\sum \text{p}_\text{T}(\text{charged}) + \max\left(\sum E_T(\text{neutral}) + \sum E_T(\text{photon}) - \Delta\beta, 0\right)}{\text{p}_\text{T}(\mu \text{ or } e)} \tag{6.3}$$

where p_T(charged) corresponds of the p_T of all charged particle candidates, p_T(photon) and p_T(neutral) correspond to the transverse energy of the photon and neutral hadron

candidates and $\Delta\beta$ corresponds to the energy estimate of neutral particles due to pile-up. In the $\mu\tau_h$ and $e\tau_h$ channels of the MSSM di-tau analysis $I_{\text{rel}} < 0.1$ is required, both for muons and electrons. In the $\text{W} + \text{b}\bar{\text{b}}$ analysis $I_{\text{rel}} < 0.12$ is required.

6.5 τ ID and Reconstruction

The τ lepton's high mass means that the τ plays a very important roll in the search for the SM higgs boson, and MSSM higgs bosons. The lifetime of the τ is short enough that it decays before reaching the inner most detector. This short lifetime makes τ reconstruction particularly challenging; the solution is to reconstruct the decay products of the τ. The dominant hadronic τ decays (τ_h) and any intermediate resonances are outlined in Table 6.2. These decays consist of one or three charged π mesons and up to two π^0 mesons. This thesis uses the hadron plus strips (HPS) algorithm for the reconstruction of τ_h's.

The HPS algorithm reconstructs photons into 'strips': these are objects which are built out of charged particles within a window of size $\Delta\eta = 0.05$ and $\Delta\phi = 0.20$. The algorithm starts by centering a strip on the most energetic electromagnetic particle within the PF jet and then it searches for other electromagnetic particles within the window. If another electromagnetic particle is found then the object gets associated

Decay mode	Resonance	Mass (MeV)	Branching fraction (%)
$\tau^- \to h^-\nu_\tau$			11.6%
$\tau^- \to h^-\pi^0\nu_\tau$	ρ^-	770	26.0%
$\tau^- \to h^-\pi^0\pi^0\nu_\tau$	a_1^-	1200	9.5%
$\tau^- \to h^-h^+h^-\nu_\tau$	a_1^-	1200	9.8%
$\tau^- \to h^-h^+h^-\pi^0\nu_\tau$			4.8%

Table 6.2: Branching fractions of dominant hadronic τ decays and mass of any intermediate resonance.

with the strip and the four-momentum is recalculated. This procedure is repeated until no other particles are found. Strips which satisfy the requirement of $p_T^{strip} > 1$ GeV are finally combined with the charged hadrons to reconstruct individual τ_h decay modes [13].

The following decay topologies are considered by the HPS τ ID algorithm where h stands for a charged hadron:

- **Single hadron** corresponds to $h^-\nu_\tau$ and $h^-\pi^0\nu_\tau$ decays in which the neutral pions have too little energy to be reconstructed as strips.
- **One hadron + one strip** reconstructs the decay mode $h^-\pi^0\nu_\tau$ in events in which the photons from π^0 decay are close together on the calorimeter surface.
- **One hadron + two strips** corresponds to the decay mode $h^-\pi^0\nu_\tau$ in events in which photons from π^0 decays are well separated.
- **Three hadrons** corresponds to the decay mode $h^-h^+h^-\nu_\tau$. The three charged hadrons are required to come from the same secondary vertex.

Charged hadrons and strips are required to be contained within a cone of size $\Delta R = (2.8\text{GeV}/p_T^{\tau_h})$ where $p_T^{\tau_h}$ is the transverse momentum of the τ hadron $p_T^{\tau_h}$ and is required to match the (η, ϕ) direction of the original PF jet within a radius of $\Delta R = 0.1$.

Tau Isolation

The isolation of τ_{had} candidates is computed by summing the transverse momenta of charged particles of $p_T > 0.5$ GeV plus photons of $E_T > 0.5$ GeV reconstructed by the PF algorithm within a cone of size $\Delta R = 0.5$ centered on the τ_{had}. Charged hadrons considered in the isolation p_T sum are required to satisfy $\Delta z < 2$ mm with

respect to the τ_{had} primary vertex. Charged hadrons and photons used to build the τ_{had} candidate are excluded from the isolation p_T sum. The contribution of pileup to the τ_{had} isolation is accounted for by applying $\Delta\beta$ corrections:

$$I_{\tau_{had}} = \sum P_T^{charged}(\Delta z < 2 \text{ mm}) + \max\left(P_T^{\gamma} - \Delta\beta, 0\right).$$

The $\Delta\beta$ corrections are computed by summing the transverse momenta of charged particles that have a longitudinal impact parameters $\Delta z > 2$ mm with respect to the τ_{had} production vertex and are within a cone of size $\Delta R = 0.8$ around the τ_{had}. The sum is scaled by a factor 0.4576, which is chosen to make the τ_{had} identification pileup insensitive:

$$\Delta\beta = 0.4576 \cdot \sum P_T^{charged}(\Delta z > 2 \text{ mm}).$$

6.6 `SVFit` Algorithm

The tau pair mass, $M_{\tau\tau}$, is reconstructed in a likelihood based algorithm, `SVFit`[22]. The kinematics of tau lepton decays is underdetermined by experimental measurements. 5 parameters are needed to specify hadronic tau decays in this model: momentum, polar and azimuthal angles of the tau lepton in the laboratory (CMS detector) frame, plus two decay angles in the rest–frame of the tau lepton.

These parameters are chosen to be:

- x, the fraction of tau lepton energy (in the laboratory frame) carried by visible decay products.
- ϕ, the azimuthal angle of the tau lepton in the laboratory frame.
- $m_{\nu\nu}$, the mass of the neutrino system.

E_x^{miss} and E_y^{miss}, reconstructed in the plane transverse to the beam–axis provide 2 further constraints.

The reconstruction of tau decay kinematics is underconstrained by measured observables; this is addressed by the likelihood approach implemented in the SV-fit algorithm. The model makes a prediction for the probability density $p(\vec{x}|\vec{y},\vec{a})$ to observe the values $\vec{x} = (E_x^{miss}, E_y^{miss})$ measured in an event, given that the unknown parameters specifying the kinematics of the tau pair decay have values $\vec{a} = (x_1, \phi_1, m_{\nu\nu}^1, x_2, \phi_2, m_{\nu\nu}^2)$ and the momenta of the visible decay products are equal to the observed $\vec{y} = (p_1^{vis}, p_2^{vis})$. The following likelihood model is used to compute probabilities:

$$P(M_{\tau\tau}^i) = \int \delta\left(M_{\tau\tau}^i - M_{\tau\tau}(\vec{y},\vec{a})\right) p(\vec{x}|\vec{y},\vec{a}) d\vec{a} \tag{6.4}$$

for a series of mass hypotheses $M_{\tau\tau}^i$. The best estimate $\hat{M}_{\tau\tau}$ for the tau pair mass is taken to be the maximum of $M_{\tau\tau}^i$ within the series. Lower (upper) limits on the reconstructed mass $\hat{M}_{\tau\tau}$ are determined for every event by the 0.16 (0.84) quantiles of the series of mass hypotheses $M_{\tau\tau}^i$ and associated probability values $P(M_{\tau\tau}^i)$.

6.7 Particle Flow Reconstruction

In proton-proton collisions, even at energies on the order of a few TeV, most stable decay products have a low p_T. At CMS, to identify and reconstruct these stable particles a particle flow (PF) event reconstruction technique has been developed [23][24]. PF event reconstruction combines reconstructed tracks with calorimeter deposits to form object collections of electrons, muons, photons, charged and neutral hadrons, HF hadrons and HF EM particles for each event. These collections can then be used to build taus, Jets, $E_{\mathrm{T}}^{\mathrm{miss}}$ and to quantify lepton isolation and tag jets originating from heavy quarks.

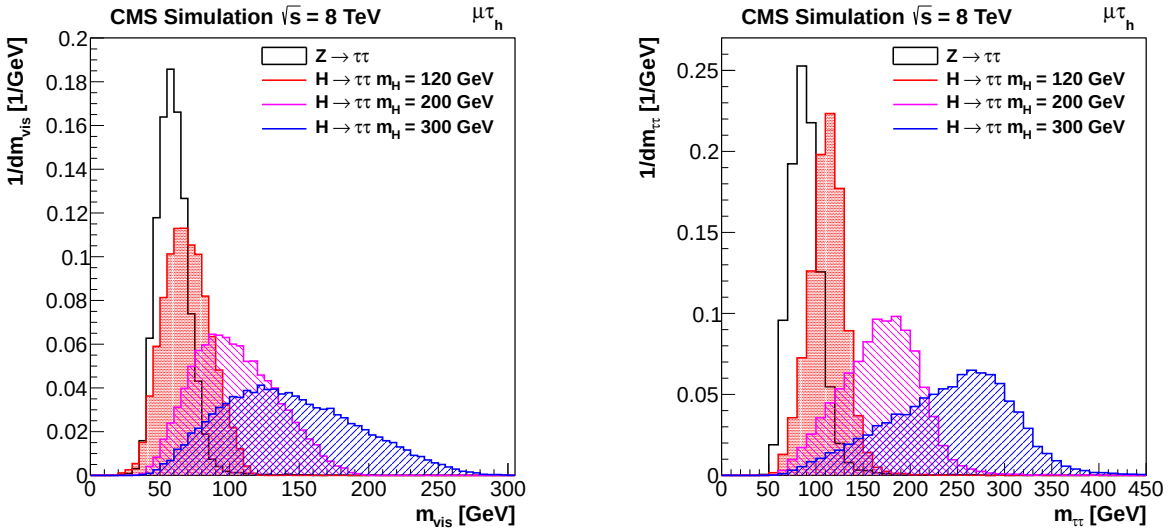


Figure 6.2: Distribution of visible mass (left) and $M_{\tau\tau}$ reconstructed reconstructed by the SVfit algorithm (right)

The first step of PF reconstruction is to perform iterative tracking. Here, tracks are seeded and then reconstructed with tight criteria where the emphasis is on achieving a low fake rate. After a track has been identified its hits in the tracker are removed and successive iterative steps then loosen track identification criteria.

The second step of PF reconstruction is to produce clusters in the ECAL and the HCAL. In this step, cluster seeds are produced from local calorimeter cell energy maxima, then topological clusters are grown by combining cells with at least one common side with a cell already in a cluster.

The final step in PF reconstruction, is to apply a 'link algorithm' which links hits in the ECAL, HCAL, tracker and muon system. In the end-cap the link algorithm begins with the outer most hit in the tracker and extrapolates first to the 2 layers of the ECAL pre-shower. Next, it searches for topological clusters in the ECAL barrel and end-cap that correspond to a maximal depth expected of a typical electron energy deposit profile. A search for topological clusters in the HCAL is then performed

at a depth corresponding to 1 X_0. The track is linked to any given depth if the extrapolated position in the calorimeter is within cluster boundaries. The cluster can be enlarged by up to one cell to account for discontinuities in the detector elements, radiation via bremsstrahlung or pair production. Finally, a link between a charged-particle track in the tracker and a muon track in the muon system is established when a global fit returns an acceptable χ^2 value.

After links have been established PF reconstruction is performed which can be summarized into three steps: First, a collection of electrons is created using GSF filter, this is further described in section 6.2. Electron tracks and calorimetric deposits are then removed. Next, PF charged hadrons are constructed by identifying links between tracks in the tracker and clusters in the ECAL. If the energy deposit in the ECAL is the same as the total p_T in the tracker within calorimetric uncertainty a PF charged hadron is created. If the energy in the calorimeters is much higher than a PF photon or a PF neutral hadron might also be created. If the energy in the calorimeters is too small then a relaxed search for hits in the muon system is performed. Finally, remaining clusters of ECAL and HCAL clusters give rise to PF photons and PF neutral hadrons. Figure 6.3 shows relative fractions of charged hadrons, photons, neutral hadrons, electrons, hf hadrons and em particles in jets.

6.8 Jet ID and Reconstruction

Jet identification and reconstruction is of central importance to the understanding and identification of many physics processes. Furthermore, efficient b jet identification is of central importance in the measurement of the $\mathrm{W} + \mathrm{b\bar{b}}$ cross section. In the search for an MSSM higgs boson, production in association with b jets is enhanced. Therefore, efficient jet and b jet identification is crucial.

Jet Reconstruction

Jet clustering algorithms have become a crucial part of physics analysis. Jet clustering in this thesis is performed using the anti-k_T jet clustering algorithm [25]. In order to perform comparisons with perturbative effects in theoretical calculations jet clustering algorithm must be both infrared and collinear (IRC) safe. When applied at particle level, an algorithm is collinear safe if the hardest particle will not be easily changed by the presence of another collinear particle. Two illustrate infrared safety, consider two hard particles on opposite edges of a jet cone with no particles between them; in this case two cones are formed, each centered on a hard particle. An algorithm is infrared unsafe if the added presence of a soft particle between the two hard particles causes a third jet to be formed. The difference with and without the soft particle results in non-cancellation of divergent real soft production and corresponding virtual contributions [26]. A jet clustering algorithm is infrared unsafe if the number of jets created through clustering increases when including an additional soft particle within the cone of a previously formed jet. A jet clustering algorithm is considered collinear safe if a hard

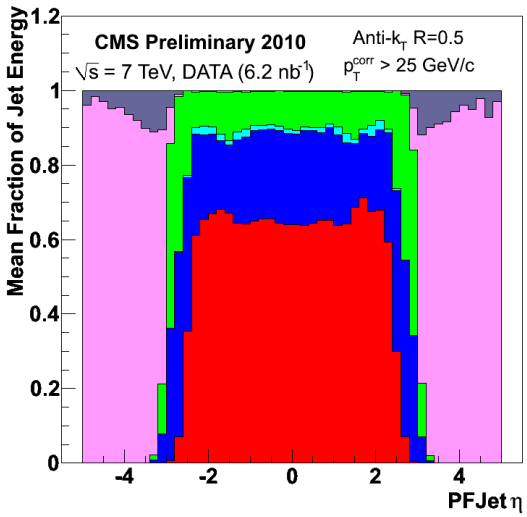


Figure 6.3: Particle Flow jet composition

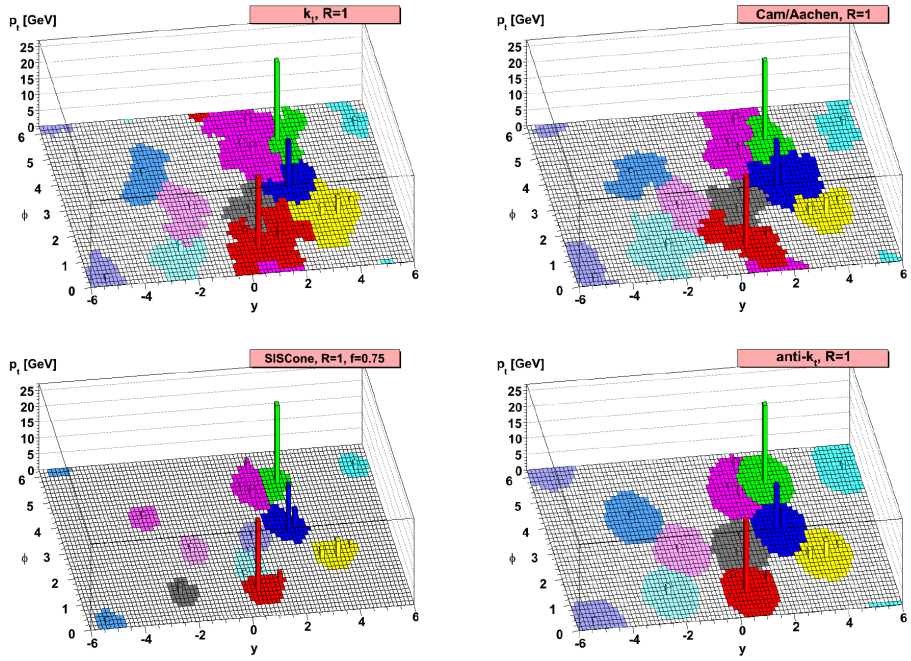


Figure 6.4: Comparison of Jet Clustering Algorithms [25]

particle is able to split collinearly and that splitting does not influence the shape of stable jet cone [27]. The development and subsequent choice of the anti-k_T jet clustering algorithm was stimulated by questions of sensitivity to non-perturbative effects like hadronization and underlying event contamination. Previously used jet clustering algorithms such as the k_T [28] and Cambridge/Aachen [29] jet clustering algorithms were IRC safe, however, they had the property that soft radiation could provoke irregularities in the boundaries of final jets. Algorithms such as SIScone [30] that were soft-resilient were not IRC safe. To perform jet clustering in the anti-k_T, k_T and Cambridge/Aachen jet algorithm a distance d_{ij} is introduced between PF particles i and j and d_{iB} between the particle entity and the beam (B). The clustering proceeds by identifying the smallest distances between candidate particles in the cluster. If the minimum distance is d_{ij} the entities i and j are combined. When the minimum distance is d_{iB} then i is considered a jet and all its entities in the jet are removed from the list of PF particles. The procedure is repeated until no entities

are left. The definition of d_{ij} and d_{iB} are,

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2} \tag{6.5}$$

$$d_{iB} = k_{ti}^{2p} \tag{6.6}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, k_{ti} is the transverse momentum, y_i is the rapidity and ϕ_i is the azimuthal angle of the particle i. The case where $p = 1$ is the k_T algorithm, $p = 0$ corresponds to the Cambridge/Aachen algorithm and $p = -1$ is the anti-k_T algorithm. The effects of each of these algorithms on an event with a few well-separated hard particles and many soft particles is shown in figure 6.4. As can be seen in the figure, the key feature of the anti-k_T algorithm is that soft particle do not modify the shape of the jet while hard particles do. The anti-k_T clustered jets in this thesis use R=0.5.

Jet Energy Corrections

Jet energy corrections (JEC) are applied to improve the accuracy of the jet p_T measurement and to flatten the jet energy response as a function of η and p_T [31]. JEC are applied by taking the dot product of a vector-factor, C, to each component of the jet four-momentum,

$$p_\mu^{cor} = C \cdot p_\mu^{raw} \tag{6.7}$$

where p_μ^{cor} is the corrected jet four-momentum vector and p_μ^{raw} is the raw jet four-momentum[32]. C itself is composed of an offset correction C_{offset}, the MC calibration factor C_{MC}, and the residual calibrations C_{rel} and C_{abs} for the relative and absolute energy scales. The correction C_{offset} removes the extra energy due to noise and pile-up, factor C_{MC} removes the non-uniformity in η and the non-linearity in p_T and the residual corrections account for the small differences between data and simulation.

These are combined into the overall correction factor C by,

$$C = C_{offset}(p_T^{raw}) \cdot C_{MC}(p'_T, \eta) \cdot C_{rel}(\eta) \cdot C_{abs}(p''_T). \tag{6.8}$$

Here, p'_T is the jet p_T after applying the offset correction and p''_T is the jet p_T after applying all previous corrections. Jets are required to pass identification criteria that eliminate jets originating or being seeded by noisy channels in the calorimeter [33]. Jets which fall within $\Delta R < 0.5$ from a lepton candidate are not included in the jet collection.

6.9 b-Jet ID and Secondary Vertices

The Combined Secondary Vertex (CSV) b-tagging algorithm makes use of the long lifetime and heavy-flavor of b-hadrons. The CSV b-tagging algorithm combines the following variables into a single discriminating variable using a likelihood ratio technique: secondary vertex mass, multiplicity of charged particles associated to the secondary vertex, the flight significance associated to the secondary vertex, the energy of charged particles associated to the SV divided by the energy of all charged particles associated to the jet, the rapidities of charged particle tracks associated to the secondary vertex, and the track impact parameter significance exceeding the charm threshold. This algorithm was tuned using b, c and non-heavy flavour jets from QCD and top samples and provides extreme discrimination of heavy and udsg jets. At a b-tagging efficiency of 60%, jets from light quarks can be reduced by a factor of 100! [34].

Secondary vertices are reconstructed inside the jet using the Trimmed Kalman Vertex Finder [19]. This algorithm begins by taking as input all tracks that are inside of the jet and performing a compatibility fit. It then removes the least compatible track and refits the vertex; this procedure is repeated until the fit is below a given

threshold. The secondary vertex studied in this thesis is that which has the greatest significance of flight distance.

6.10 Missing Transverse Energy, Recoil Corrections and MVA $E_{\mathrm{T}}^{\mathrm{miss}}$

Missing transverse energy is defined using PF candidates as,

$$E_{\mathrm{T}}^{\mathrm{miss}} = -\sum_{i} p_T \tag{6.9}$$

where i runs over all reconstructed PF candidates. To improve the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution in events with jets and neutrinos in the decay products which are expected to have a high $E_{\mathrm{T}}^{\mathrm{miss}}$, such as $\mathrm{W} + \mathrm{b\bar{b}}$, recoil corrections are applied to the $E_{\mathrm{T}}^{\mathrm{miss}}$.

To improve the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution a correction to the recoil of the generated bosons is applied. Momentum conservation in the transverse plane requires,

$$E_{\mathrm{T}}^{\mathrm{miss}} + q_T + u_T = 0 \tag{6.10}$$

where q_T is the vector boson transverse momentum which is measured in $Z \to \mu\mu$ data and matched to simulation in $W \to \mu\nu$. Finally, u_T is the transverse momentum of the hadronic recoil.

$$u_T \equiv \sum_{j} p_{j,T}, \tag{6.11}$$

where the index j runs over all particles initiating at the interaction point excluding the vector boson.

Multivariate Analysis (MVA) $E_{\mathrm{T}}^{\mathrm{miss}}$ is used in the search for a MSSM $h \to \tau\tau$. MVA $E_{\mathrm{T}}^{\mathrm{miss}}$ aims at improving further the $E_{\mathrm{T}}^{\mathrm{miss}}$ and makes use of a series of multivariate regressions based on a boosted decision tree (BDT) [35]. A decision tree requires as input both a signal and background monte carlo sample. These samples

are split into two parts: one part is used for training (creation of the decision tree) the second part is used for testing (to test the final classifier after training). Using the training sample, a decision tree is created by defining a number of discriminatory variables from an event and then splitting each variable by some value which optimizes the separation of signal and background. If the final variable selection has a purity of signal over signal plus background above a given threshold then it is considered a 'leaf' and used in the selection of a signal and background events. A score of 1 is assigned if the event falls on a signal leaf and 0 if it falls on a background leaf. Boosting is then performed by taking all signal events that were misclassified as background or background events misclassified as signal and increasing the weight of that event (boosting). A new tree is built using the new weights and the final score is calculated using these weights. A similar method is used to determine a BDT regression scale factor.

The MVA $E_{\mathrm{T}}^{\mathrm{miss}}$ uses the BDT regression model to determine correction factors to u_T. This correction is performed in two steps: first, compute a correction to the azimuthal angle of u_T by training a BDT with the true hadronic recoil, $-q_T$, as the target. Next a separate BDT is trained to predict the magnitude of u_T. This corrected u_T is then used in equation 6.10 to calculate a new MVA $E_{\mathrm{T}}^{\mathrm{miss}}$.

Chapter 7

Event and Detector Simulation

Event simulation is crucial in order to measure detector acceptances and efficiencies for various physics processes and to estimate backgrounds due to known but SM physics processes. Event simulation requires prediction of a physics process expected at the LHC, production of physics final states and distribution in phase space of final state particles and then simulation of particle interactions within the detector[36]. The physics process is simulated using a monte carlo (MC) generator while the simulation of the detector is performed using a dedicated software package which models in interactions of particles in their passage through matter. The output of simulation should be similar if not the same as the output of the detector; in this way the simulation and data can undergo the same reconstruction and analysis. Simulations are initiated using a program called an 'event generator'. The purpose of an event generator is to produce hypothetical events with a distribution of observables as predicted by theory. Event simulation can be used for the following:

- Aid in the design of detectors and experiments.
- Provide a prediction of what physics processes can be expected and at what rates

- As a tool for devising analysis strategies to optimize signal significance and reduce unwanted backgrounds.
- As a framework to in which to test signal significance.
- As a means to test current physics theory/model and aid in the confirmation of new physics discovery.

In the study of particle physics it is important to not trust completely a single generator; if particle physics where perfectly understood and well modeled then there would be no point in building a detector! Instead, it is important to understand the ingredients that go into a simulation and choose a generator which describes well the desired process. A number of physics generators are in common use; a few of these are outlined in section 7.2. Physics simulation can be divided into a few key components: The hard scattering matrix elements, which define the process(es) under study, the structure function, which models the momentum distributions of the partons which are involved in the collision, final and initial state radiation and underlying event processes including multiple interactions in the detector.

7.1 Hard Scattering Process

Hard scattering matrix element generation of a hard scattering process begins typically with a well defined subprocess. Considering the simulation of the process $W \to \mu\nu$ we start with the leading order differential cross section defined as,

$$d\sigma(\bar{q}q' \to W^- \to \mu^-\bar{\nu_\mu}) = f(x,\mu_F)f'(x',\mu_F)\frac{1}{2\hat{s}}|\mathcal{M}(\bar{q}q' \to W^- \to \mu^-\bar{\nu_\mu})|^2\frac{d\cos\theta d\phi}{8(2\pi)^2} \tag{7.1}$$

where the decay angles (θ, ϕ) of the W^- are two degrees of freedom. $\mathcal{M}$ is the relevant matrix element, q and q' are the initial partons, $1/(2\hat{s})$ is the parton flux and $\hat{s} = xx's$

where s is the center of mass energy squared. The factorization scale is μ_F and has the physical interpertation as a collinear cutoff[37]. Equation (7.1) is used to write an event generator. To do this first a sampling of the relevant phase space must be performed using a random number generator which selects momenta of the colliding constituents by sampling the parton distribution function of the proton at the energy scale of the subprocess.

Next, the events must be unweighted using the hit-and-miss technique to produce events with observables. The relevant phase space for this process is two dimensional: -1 $< \cos\theta <$ 1, 0 $< \phi <$ 2π. The values of $cos\theta$ and ϕ can be chosen using a uniform random number generator. Evaluating the equation (7.1) at $d\sigma(\theta_i, \phi_i)$ gives that candidate event's differential cross section which is equivalent to the probability of this event occurring. The average of many candidate event's differential cross section, $\langle d\sigma \rangle$, is an approximation to $\int d\sigma$ and converges to the measured cross section. So far, the candidate events are distributed flat in phase space and there is no physics information in the distributions. The next step then is to unweight the event using the 'hit-and-miss' technique so that they are distributed according to theoretical prediction. By unweighting the events are generated with the frequency predicted by the theory being modeled and individual events represent what might be observed in a trial experiment.

Parton Showering, Hadronization and the Underlying Event

To successfully simulate a process in a hadron hadron collider the simulated event must include parton showering, hadronization and simulation of the underlying event. Parton showering is a form of correction to the hard scattering subprocess where it is not feasible to calculate these processes exactly. Programs that use the parton shower approach are also able to simulate a wide variety of initial and final state processes.

The Lund String model for parton showering is based on models from lattice QCD simulations which show that at large distance the potential energy of color sources increases linearly with their separation. A schematic of parton showering using the Lund String model is shown in figure 7.1. Considering the production of a quark-anti-quark pair, at first the quark and anti-quark are moving apart very quickly. A gluonic string is stretched between them. Its potential energy grows causing a reduction in the kinetic energy of the quark-anti-quark system. When the gluon's potential energy reaches the order of a hadron mass the gluon string breaks apart and a quark-anti-quark pair is produced.

For a lepton-lepton collider, the energies of the initial state leptons is determined by the accelerator and is well known. For a hadron-hadron collider each hadron is regarded as a collection of quarks, antiquarks and gluons each of which carries some fraction x of the hadron's momentum at a scale Q with a number density

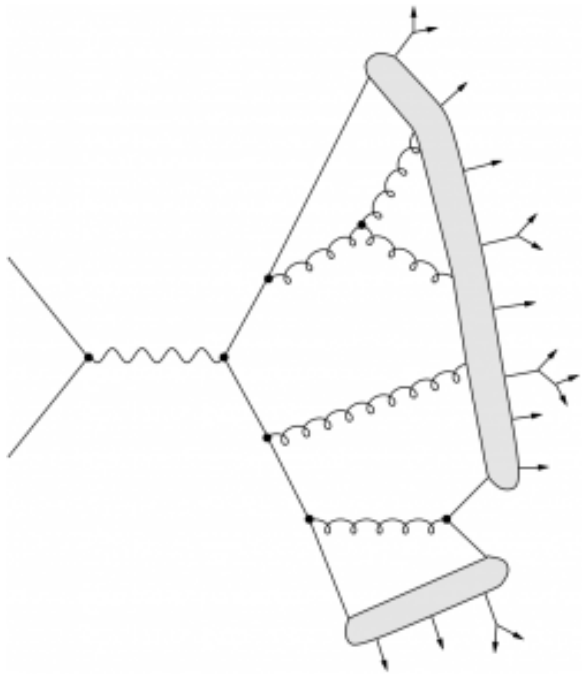


Figure 7.1: Lund String Model for Parton Showering

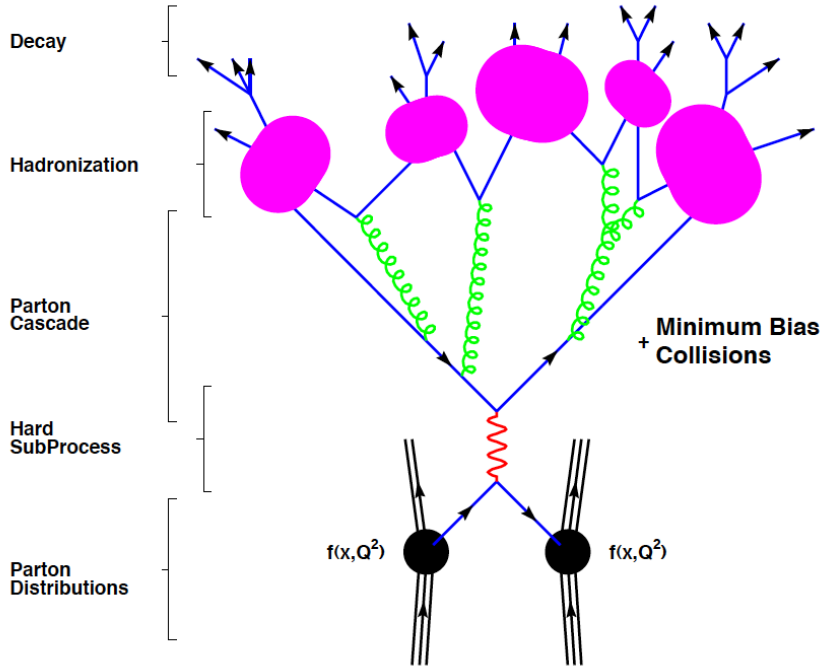


Figure 7.2: The basic simulation structure for an event in a hadron collision including parton distribution, hard subprocess, parton showering, hadronization, final decay and pileup (Minimum Bias Collisions) [38].

$f(x, Q^2)$. Q characterizes the hard scattering ($Q^2 = M^2_{l^+l^-}, p^2_T$, ect.) and changes on the order of $O(1)$ are equivalent[39]. Figure 7.2 shows a proton-proton collision where a valence quark is separated from one of the colliding protons and a a sea anti-quark is separated from the other of the colliding protons. The phenomenology of the parton interaction is encoded in the parton distribution function $f(x, Q^2)$. The parton distribution function represents the probability densities to find a parton carrying a momentum faction x at a squared energy scale Q^2[40]. Since the hadronization scale is much smaller than the hard scale the impact of the hadronization model choice on the final state is typically small for most physics analyses. Parton distribution functions used in this thesis are `CTEQ6.6` and `MSTWNNLO` . `CTEQ6.6` parton distribution functions are extracted from a global analysis of hard-scattering data from a variety

of fixed-traget and collider experiments. In the 6.6 version of CTEQ has incorporated a comprehensive treatment of heavy-quark effects along with including recent HERA charm production cross sections. MSTWNNLO also obtains parton distribution functions via a global analysis. The data sets fitted in MSTWNNLO include CCFR/NuTeV dimuon cross sections, which constrain the strange quark and antiquark distributions. They also include Tevatron Run II data on inclusive jet production, the lepton charge asymmetry from W decays and the Z rapidity distribution.

The beam remnants are the colored remains of the proton which are left behind when the parton which participates in the hard subprocess is pulled out. As the beam remnants are color connected to the hard subprocess they should be included in the hadronization system. Multiple parton interactions (MPI) where more than one pair of beam partons interact are also simulated. Due to the composite nature of hadrons and the low impact parameter of spectator partons MPI have an increase rate of occurrence in the presence of a hard scattering process. Figure 7.3 shows a comparison of feynman diagrams for W+jj interaction from double parton scattering (an MPI interaction) and single parton scattering. In the $\mathrm{W} + \mathrm{b\bar{b}}$ analysis the effects from double parton scattering must be taken into account when calculating the measured cross section. MPI agreement between data and simulation is tested by observing the relative p_T balance of the two jets via the ratio of $(p_{T,j_1} + p_{T,j_2})/p_{T,j_1,j_2}$ [41].

The LHC is a high luminosity experiment. Therefore, an obseved event will have as background interactions from other colliding protons. Successful event simulation must include pileup from other proton-proton collisions.

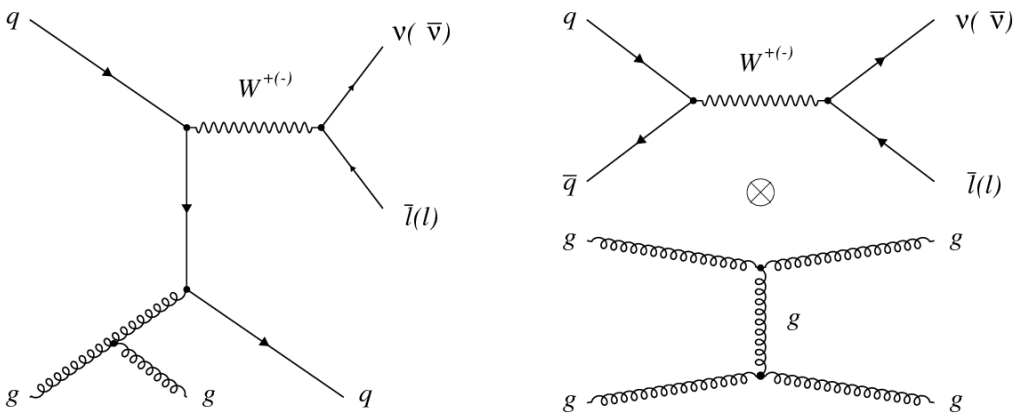


Figure 7.3: Single Parton Scattering and Double Parton Scattering in a W+jj event [41]

7.2 Monte Carlo Generator Programs

A number of monte carlo (MC) generators are used to simulate physics processes. The MC method makes use of a random number generator to simulate event to event fluctuations which are intrinsic quantum processes. The choice of MC generator is largely dependent on the type of process to be simulated.

MadGraph

MadGraph[42] is a matrix element generator for SM processes at any collider. It provides a computation of tree-level matrix elements with a fixed number of partons in the final state. For a user-specified process MadGraph generates the amplitudes for all relevant subprocesses and produces the mappings for the integration over the phase space. The final state only events may be passed directly to a shower MC program.

Tauloa

Tauloa[43] is a package which is used for generation of tau lepton decays including spin polarization. For each tau decay mode there is a individual phase space generator, modeling of the weak decay including first order QED corrections for leptonic decays and a part describing the hadronic current.

PYTHIA

At leading order `PYTHIA`[44] contains approximately 240 different $2 \to n$ subprocesses. The initial state shower is based on backwards evolution whereby the hard scattering process is simulated and is evolved backwards in time to the shower initiators. Partons radiated in the initial state can initiate final-state showers of their own. The Lund string model is used in hadronization; it is based on linear confinement where quarks are located at the ends of the string and gluons are energy and momentum carrying strings. The production of a new $q\bar{q}$ pair causes the string to break. A quark from one break can combine with an antiquark from an adjacent quark to form a meson.

POWHEG

`POWHEG` (Positive Weight Hardest Emission Generator) [45, 46] generate the hardest radiation first and then feed the event to a shower gnerator for subsequent softer radiation. For a shower generator the hardest emission with highest transverse momentum is always generated first; `POWHEG` simply replaces the hardest emission with its own NLO accurate emission.

MCFM

Monte Carlo for FeMtobarn (MCFM)[47] processes at hadron colliders is a parton-level Monte Carlo program which gives next to leading order predictions for a range of processes at hadron colliders. The difference between leading and next to leading order is described in section 1.3. MCFM produces weighted events and therefore can give very accurate predictions for event rates at NLO with cuts included; however, it gives very little information about regions of phase space which are dominated by multiple parton interactions. Since the final state contains partons rather than hadrons, full detector simulation cannot be performed using the MCFM output. Furthermore, a parton to hadron correction factor must be included when comparing (for instance) with `PYTHIA`, a generator which does include hadrons in the final state. The parton to hadron correction factor is estimated for the $\mathrm{W} + \mathrm{b\bar{b}}$ cross section measurement detailed later in this thesis and is shown to give significant correction to the overall simulated cross section.

7.3 Detector Simulation

`GEANT4`[48] is a toolkit for simulating the passage of photons, muons, electrons, hadrons and ions through matter at energies from 250 eV to several PeV. The simulation package models the detector geometry, electric and magnetic fields and a wide range of initial to final state physics processes. Geometry simulation includes a detailed model of the experiment including layout of the detectors and the location of absorbers (cables, cooling systems, ect.) Event management for recording details of each event as well as a number of options for the visualization of the simulation output. Magnetic field simulation makes use of detailed measurements of the magnetic field throughout the CMS detector. `GEANT4` uses decay tables to properly simulate de-

cay rates for particles such as π, K mesons and resonant baryons based on data from the Particle Data Group[49]. The models used for decay of various physics processes are either theory or data driven and the toolkit offers a large variety of complementary and sometimes alternative physics models. GEANT4 uses a simple model for simulating the lifetime of a particle whereby the probability of a particle surviving a distance l is given as,

$$P(l) = e^{-\eta_\lambda}. \tag{7.2}$$

In the above equation, η_λ is an exponential distribution which is independent of material and energy; therefore, $\eta_\lambda = -\ln(\eta)$ and η is a random number on the interval $(0, 1)$. GEANT4 has been an essential and very successful tool in the simulation of detector effects at CMS.

Chapter 8

Measurement of $\mathrm{W} + \mathrm{b\bar{b}}$ Production

This chapter describes a study of the production of a W boson and two b jets in proton-proton collisions, where the W boson is observed via its decay to a muon and a neutrino, and each b jet is identified by the presence of a b hadron with a displaced decay vertex. This production channel provides an important testing ground for standard model (SM) predictions. A key feature of this analysis is the $\mathrm{b\bar{b}}$ phase space covered. Previous measurements have concentrated on W-boson production with at least one observed b-quark jet, for which the predictions differ from the experimental results. Previous measurements of vector boson production with associated b-quark jets have shown varying levels of agreement with theoretical calculations [50, 51, 52]. This difference is larger in the production of events with a collinear $\mathrm{b\bar{b}}$ pair that is reconstructed as one jet [53, 54], a topology afflicted by significant theoretical uncertainties. Recently, a measurement of W plus one b jet production in proton-proton collisions at center of mass energy of 7 TeV presented by the ATLAS collaboration[55] has found the difference between data and MCFM NLO is on average 1.5 σ. A comparison between their measured and predicted cross section is shown in figure 8.1. It is important to mention that the ATLAS W + 2 jet category is, in fact, completely

orthogonal to the W + 2 b jet measurement that is the subject of this chapter. In their W + 2 jet category a passing event is required to have 2 jets where exactly one of the jets are b tagged; events with 2 b tags are excluded to reduce the significant $t\bar{t}$ background.

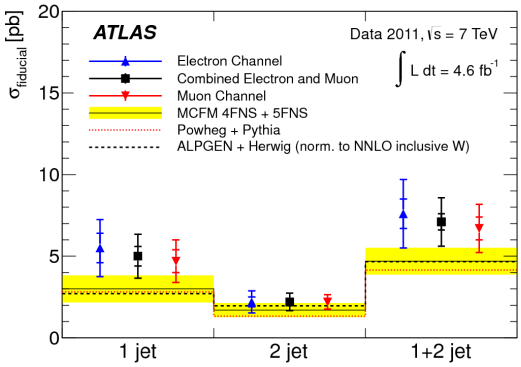


Figure 8.1: ATLAS W+b cross section measurement at 7 TeV [55]

By focusing on the observation of W-boson production with two well-separated b-quark jets, this analysis provides an essential study to analyses where $\mathrm{W} + \mathrm{b\bar{b}}$ is an irreducible background such as SM higgs boson production in association with an electroweak gauge boson and subsequent decay to $\mathrm{b\bar{b}}$.

Other SM processes produce events with an experimental signature similar to the one studied here; these SM processes must be carefully understood and modeled to obtain an accurate $\mathrm{W} + \mathrm{b\bar{b}}$ cross section measurement. These backgrounds include production of top quark-antiquark pairs ($t\bar{t}$), associated production of a W boson with light jets misidentified as b-quark jets, single-top-quark production, multijet production (henceforth labeled "QCD multijet"), Drell–Yan production associated with jets, and electroweak diboson production.

The production mechanism of $\mathrm{b\bar{b}}$ pairs together with W or Z bosons has been the subject of extensive theoretical studies and is included in different simulation programs [46, 56, 57] but is still not thoroughly understood. According to the SM,

the primary contribution for $b\bar{b}$ production in association with a W boson is due to the splitting of a gluon into a $b\bar{b}$ pair. Two different models for b-quark production are available, depending on whether there are four or five quark flavors in the proton parton distribution functions (PDFs) [58]. In the 4 flavor scheme the production of b–quarks is computed by perturbative QCD. The 5 flavor performs the calculation in the framework of bottom parton densities within the proton. An implicit approximation made by the 5 flavor scheme is that the produced b–quarks carry small transverse momenta. Finite P_T effects are accounted for in the 5 flavor scheme at higher orders.

The analysis uses a sample of proton-proton collisions at a center-of-mass energy of $\sqrt{s}$=7 TeVcollected in 2011 with the CMS experiment at the LHC, corresponding to an integrated luminosity of $5.0 fb^{-1}$.

A number of Monte Carlo (MC) event generators are used to simulate the signal and backgrounds. The events with W or Z boson production, or with $t\bar{t}$ production, are generated at leading order (LO) with `MADGRAPH`5.1 [57], which is interfaced with `PYTHIA`6.4 [59] (also LO) for hadronization. Single top samples are generated at next-to-leading order (NLO) with `POWHEG`2.0 [60, 61, 45]. Diboson (WW, WZ, ZZ) and multijet samples are generated with `PYTHIA`6.4 [59]. For LO generators, the default set of parton distribution functions (PDF) used to produce these samples is CTEQ6L [62], while MSTWNNLO [40] is used for NLO generators. For all processes, the detector response is simulated using a detailed description of the CMS detector, based on the `GEANT4` package [63], and event reconstruction is performed with the same algorithms as used for data. The simulated samples include additional interactions per bunch crossing (pileup), the distribution for which comes from data. Parton showering is simulated with PYTHIA using the Z2 tune [64].

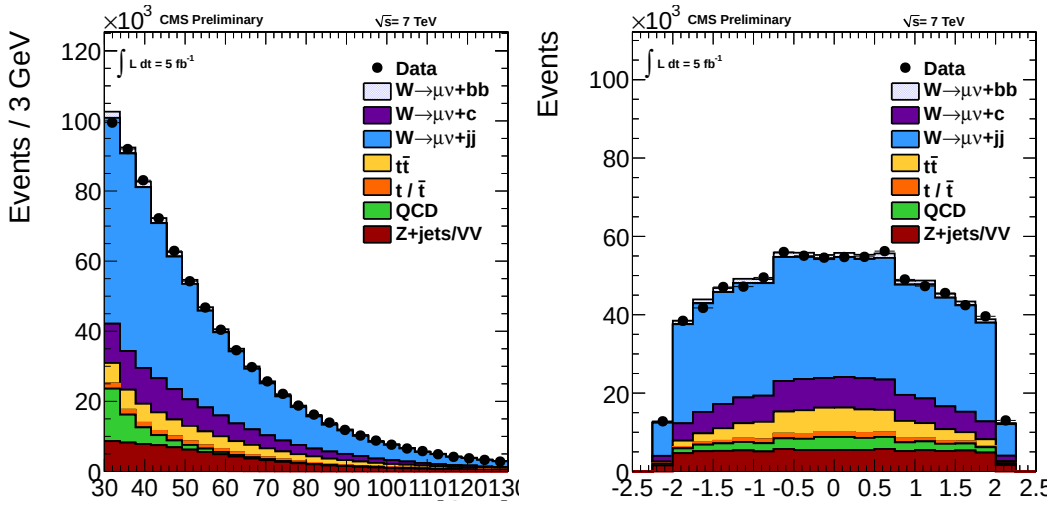
A description of muon reconstruction as well as a description of isolation is described in chapter 6. Muon candidates are required to be compatible with the primary

vertex of the event and have an isolation (I^{rel}) less than 0.12. Figure 8.2 shows the muon transverse momentum, pseudorapidity and Isolation (I^{rel}) is shown in figure 8.5 (left).

These identified, isolated muons are then combined with the missing transverse energy $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ of the event to form a leptonic W candidate. The missing transverse energy $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ is defined as the negative vector sum of the transverse momenta of all reconstructed particles in the event, with $E_{\mathrm{T}}^{\mathrm{miss}} = |\vec{E}_{\mathrm{T}}^{\mathrm{miss}}|$, and is corrected using the procedure described in Ref. [65]. The reconstructed transverse mass of the system is built from the transverse momentum of the isolated muon and the missing transverse energy of the event,

$$M_{\mathrm{T}} = \sqrt{2 \mathrm{p}_{\mathrm{T}}{}^{\ell} E_{\mathrm{T}}^{\mathrm{miss}} \, (1 - \cos \Delta\phi)} \; ,$$

where $\Delta\phi$ is the difference in azimuth between $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ and $\vec{P}_{\mathrm{T}}^{\mu}$. In $\mathrm{W} \to \ell\nu$ decays the M_T distribution presents a Jacobian peak with an edge at the W mass. Therefore it is a natural topological discriminator against non-W final states which yield a lepton candidate and missing transverse energy, such as QCD multijet processes, in which



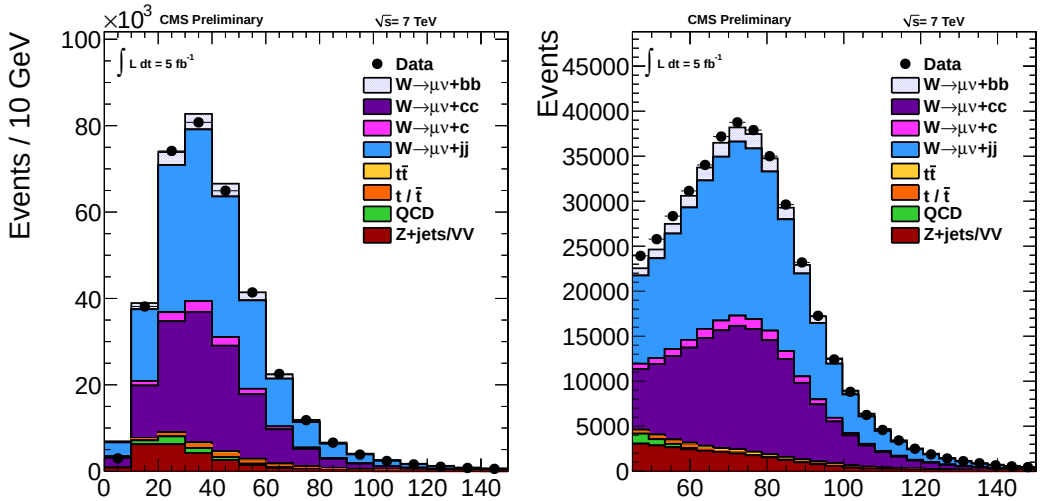
(a) μ p_T [GeV] (b) μ η

Figure 8.2: Muon p_T and η are shown above.

the events concentrate at low values of the M_T variable. The M_{T} and M_T variables are shown in figure 8.3.

Jets are reconstructed from the PF candidates. The anti-$\mathrm{k_T}$ clustering algorithm [25] with distance parameter of 0.5 is used, as implemented in the FASTJET package [66, 67]. Jet reconstruction is described in more detail in section 6.8. Jets are required to pass identification criteria that eliminate jets originating or being seeded by noisy channels in the calorimeter [33]. In addition to this, jets originating from pileup interactions are rejected by requiring compatibility of the jets with the primary interaction vertex. Jet energy corrections are also applied as a function of the jet $\mathrm{p_T}$ and η [68].

Secondary vertices (SV) are reconstructed inside each jet. This study makes use of the combined secondary vertex (CSV) b-tagging algorithm [34]; this algorithm makes use of the long lifetime and high mass of b hadrons to provide optimized b-quark jet discrimination; the following variables are combined into a single discriminating variable using a likelihood ratio technique: secondary vertex mass, multiplicity of



(a) E_T^{miss} [GeV] (b) $M_T(\mu, E_T^{miss})$ [GeV]

Figure 8.3: The missing transverse energy (E_T^{miss}) and transverse momentum of muon and E_T^{miss} (M_T) are shown above.

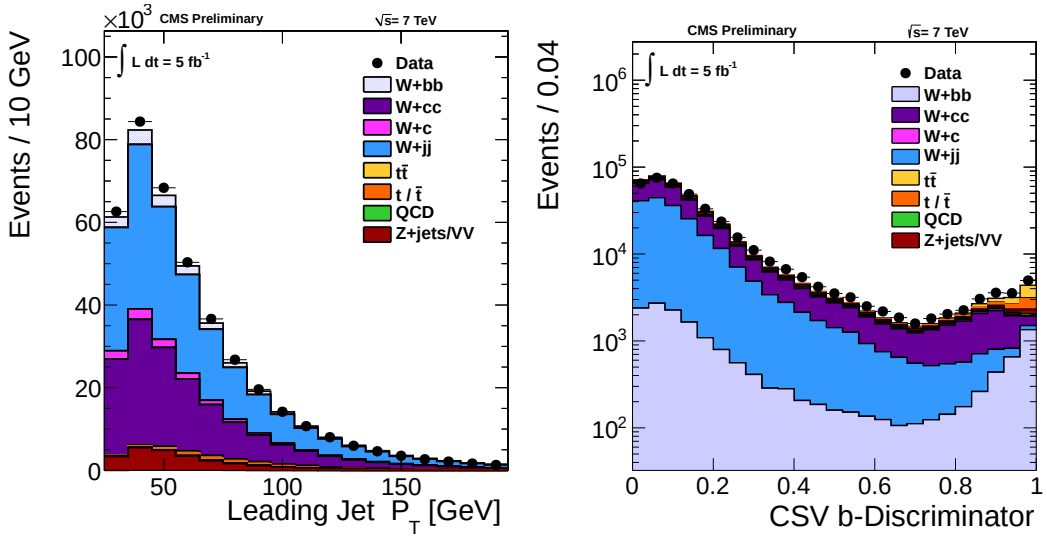


Figure 8.4: (left) The highest-p_T jet (J_1) before applying b-tagging. (right) The CSV b-discriminator for J_1.

charged particles associated to the secondary vertex, the flight signicance associated to the secondary vertex, the energy of charged particles associated to the SV divided by the energy of all charged particles associated to the jet, the rapidities of charged particle tracks associated to the secondary vertex, and the track impact parameter signicance exceeding the charm threshold. If a jet does not have an SV then CSV algorithm computes 'pseudo Vertex' and 'No-Vertex' values. Low values of the CSV discriminator mean the jet is less b jet-like, while values close to 1 are more b jet-like. B-tagged jets are selected by imposing a minimum threshold on the CSV discriminator value. The analysis is based on a CSV discriminator threshold which provides an efficiency of approximately 50% for identifying jets containing b-flavored hadrons while reducing the misidentification probability for light-quark jets to 0.1% [69].

The $W + b\bar{b}$ selected events are required to have an isolated muon with $I^{\text{rel}} < 0.12$, $p_T > 25GeV$, $|\eta| < 2.1$, exactly two jets with $p_T > 25GeV$ and $|\eta| < 2.4$, where both selected jets must contain a secondary vertex and pass the b-tagging CSV requirement. To reduce the contribution from Z-boson production, the event

is rejected if there is a second muon, without any requirements on the isolation and p_T, which builds with the isolated muon a dimuon system with invariant mass $m_{\mu\mu} > 60GeV$. The $t\bar{t}$ background is reduced by requiring that there are no additional isolated electrons or muons with $p_T > 20GeV$ in the event and no jets with $p_T > 25GeV$ and $2.4 < |\eta| < 4.5$. To reduce the contribution from QCD multijet events $M_T > 45GeV$ is also required. An evolution of the yields in monte carlo and data due to selection is shown in table (8.1).

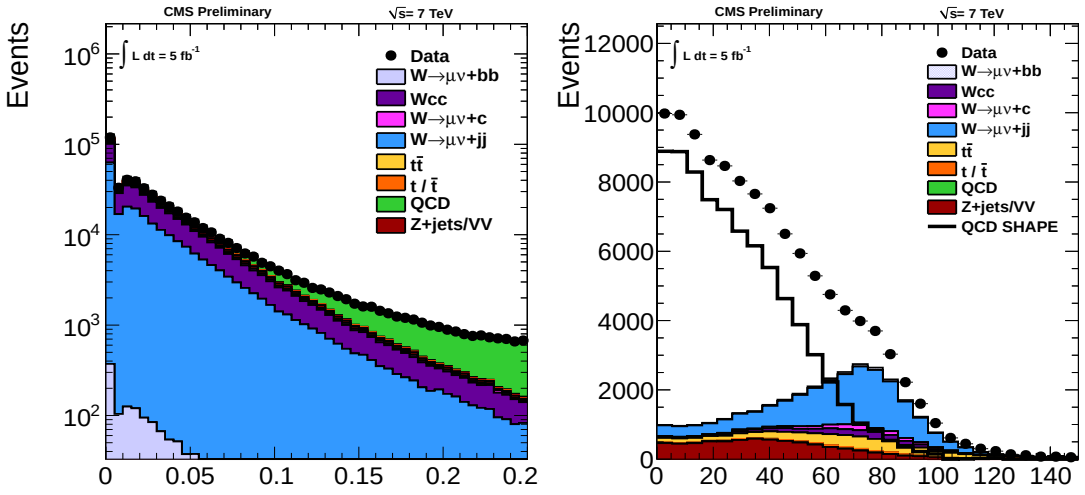
After all the selection requirements the significant background contributions are: $t\bar{t}$, single top, W+jets (u,d,c,s,g), Z+jets (u,d,c,s,b,g) and QCD multijet. Final contributions of these backgrounds in the signal region are computed via a simultaneous fit. This fit is described in section 8.6 and provides the final estimate for the signal and background yields. The initial yields are taken either from data, in estimates based on the control regions, or from simulation, normalized to the NNLO predictions. The shapes and normalizations of the background distributions are validated in data with a set of control regions, as described in the next sections.

8.1 QCD multijet background

With the exception of QCD, the shapes of the background distributions are taken from simulation. A shape for the QCD template is obtained directly from a multijet-enriched control region in data. To create a control region enriched in QCD events all selection requirements, including lepton and jet vetos, are applied, except the isolation requirements of the main lepton are inverted. The muon is required to have $I_{rel} > 0.2$; figure 8.2 shows the muon isolation distribution. Good agreement between data and monte carlo is observed. After selecting these non-isolated events, there remains a significant number of top, W and Z events estimated from MC simulation.

Process	Wbb	W+l	W+c	W + c$\bar{c}$	Z+jets	$t\bar{t}$	Single-t	VV	QCD	Total MC	Total Data
W+2jets	39333.1	378197.2	23502.8	284441.3	94169.7	74082.9	15880.3	10195.0	42276.7	962079.0	928445.0
$M_T > 45$ GeV	30381.9	291990.5	18282.5	221570.1	35144.1	54095.4	11909.0	7780.7	8393.9	672138.2	642674.0
JetVeto $\eta < 2.4$	21605.7	237049.8	14498.1	175253.2	26490.8	9027.4	6599.4	5744.0	6520.0	498028.4	478315.0
JetVeto $\eta < 4.5$	17152.9	196618.2	11853.4	142711.5	21440.5	5390.2	4356.3	4727.3	5298	405888.5	408705.0
Lepton Veto	17125.9	196213.9	11836.3	142517.5	18994.8	4707.9	4155.8	4446.5	5284.3	401973.7	402742.0
2 CSV Tight Tags	356.5	47.3	1.7	56.3	34.5	620.6	168.0	19.9	36.3	1249.3	1387.0
>=1 SV in each jet	332.3	1.5 ±0.2	1.3 ±0.3	21.0 ±4.3	30.9 ±2.9	595.5 ±32.8	160.3 ±7.8	18.9 ±0.8	33.1 ±2.7	1194 ±78.0	1230.0 ±35.0
2 CSV Med Tags	697.4	227.2	30.6	340.8	83.4	1039.6	327.4	42.3	63.5	2822.0	3065.0
>= 1 SV in each jet	557.5	18.9 ±1.9	12.7 ±1.3	156.5 ±31.2	54.8 ±5.4	893.8 ±53.6	273.9 ±14.5	33.1 ±1.8	18.9 ±5.6	2020.2 ±153.3	2227.0 ±47.2

Table 8.1: Evolution of event selection



(a) μ I^{rel} (b) M_T [GeV]

Figure 8.5: Muon Isolation (left). The contributions of individual backgrounds in the anti-isolated M_T distribution (right). The template for the QCD shape is taken as the difference between simulation and data in the non-isolated region.

These events make up to approximately 10% of all events in the control region and are subtracted from data to construct the QCD shape in the signal region; this background subtraction is illustrated in Fig. 8.5. The initial normalization of the QCD yield (before final signal extraction) is found by performing a fit in the $M_T(W)$ distribution in the region where $M_T(W) < 45$. The QCD uncertainty in the final fit is taken to be $\pm 50\%$. This uncertainty is sufficient to provide coverage for normalization and shape mismodelings of the small QCD contribution in the final selection.

8.2 W+jets: light and charm component

The W+jets (u,d,c,s,g) process, where the jets are not initiated by b quarks, is the dominant background before applying the selection requirements on the secondary vertex and b-tagging. Figure 8.4 (left) shows the p_T of the leading jet at this preselected stage. The CSV algorithm working point which provides maximum reduction of W+jets (u,d,c,s,g) is used. The CSV b-tagging discriminant for the leading jet is

shown in Figure 8.4 (right).

The presence of light and charm jets in the sample is very small at the higher values of the discriminant. Furthermore, to increase the purity of the sample a secondary vertex is required to be reconstructed in each of the selected jets. Figure 8.7 shows the mass of the secondary vertex of the leading jet (J_1, right) and the sub-leading jet (J_2, left), for the final selection in the signal region. These selection requirements have been validated in the $t\bar{t}$ and Z+jets control regions described below.

A powerful discrimination variable to distinguish $W + b\bar{b}$ and $W + c\bar{c}$ is the sum of the masses of the two secondary vertices ('J1SVMass + J2SVMass') corresponding to the two b jets. The correlation between J1SVMass + J2SVMass is shown in Figure 8.6 for the $W + b\bar{b}$ and $W + c\bar{c}$ samples. It is clear from this distribution that $W + b\bar{b}$ populates the high mass region of the phase-space (J1SVMass + J2SVMass 3 GeV) while the $W + c\bar{c}$ contribution populates the low mass region of the phase-space. This distribution is shown in figure 8.11 for jets passing the CSV medium working point.

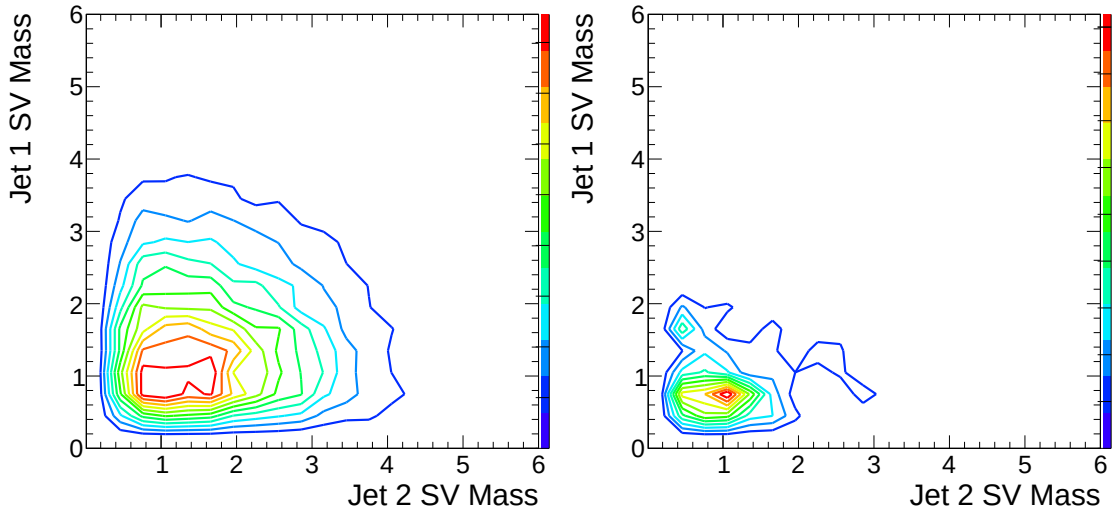


Figure 8.6: Contour plot high p_T vs. second highest p_T secondary vertex mass. The W+c secondary vertex mass is concentrated in the region where the secondary vertex masses of the jets sum is less than 3.

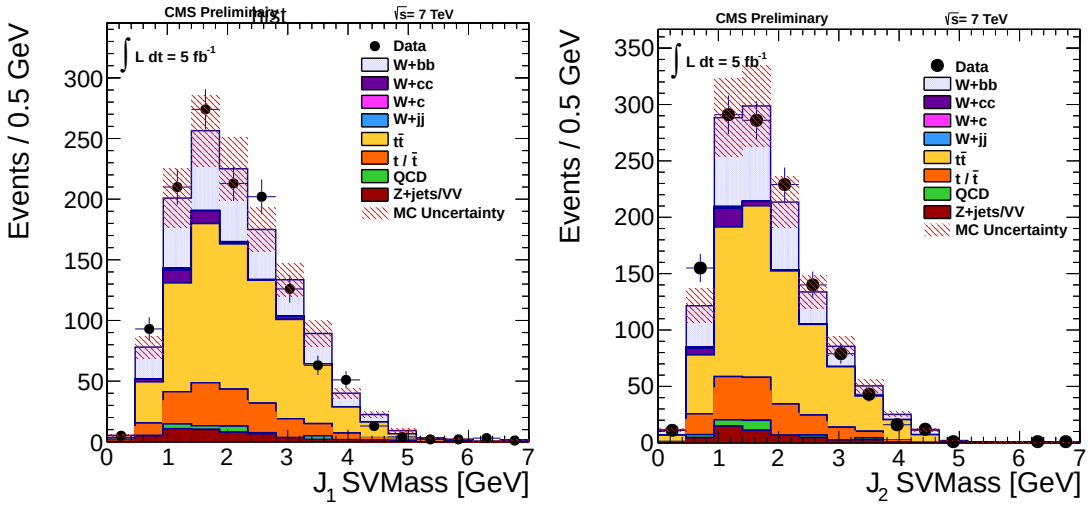
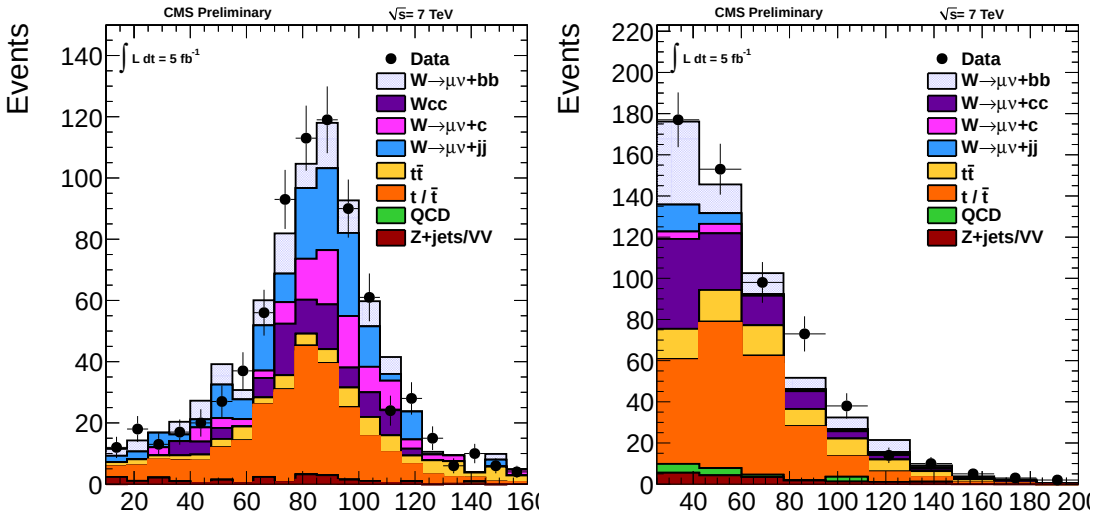


Figure 8.7: Mass of the secondary vertex for the highest-p_T jet (J_1, right) and for the second jet (J_2, left) in the signal region.

8.3 Top backgrounds

The events selected for the $t\bar{t}$ control region pass the selection requirements, with no restrictions on the number of leptons in the event. In addition to the two highest-p_T b-tag jets, the events are required to have at least two extra light jets. This higher jet multiplicity requirement selects a sample that is dominated by $t\bar{t}$ events. Figure 8.10 (right) shows the invariant mass of the two highest-p_T additional jets (3rd and 4th highest-p_T in the event, $m_{J_3J_4}$). In $t\bar{t}$ events this distribution reconstructs the mass of the hadronically decaying W boson. It is used in the final fit to extract the $t\bar{t}$ background normalization. The simulation describes the observed distributions well both in shape and normalization.

A single-top-quark control region is defined by selecting events in which the W boson is accompanied by exactly one b jet passing the described tagging criteria, and an additional forward jet with $|\eta|$ >2.8. No additional vetoes on extra light jets or leptons are applied. As seen in figure 8.8, the simulation describes the single-top production control region well and therefore it is used to estimate the yield and shape



(a) M_T [GeV] (b) Leading Jet p_T [GeV]

Figure 8.8: The single top control region defined by one (b)-tagged jet and one forward jet at (η >2.8). Both the transverse mass of the selected W in the $t \to b\mu\nu$ process (left) and the pt of the leading jet (right) show agreement with the Monte Carlo prediction.

of the single top contribution in the signal region.

8.4 Z Background

The Z+jets background estimate is validated in a control region where the standard selection is applied except that a second muon is required, $70 < m_{\mu\mu} < 100 GeV$. As seen in figure 8.9, agreement between the observed distributions and simulation is observed in this region.

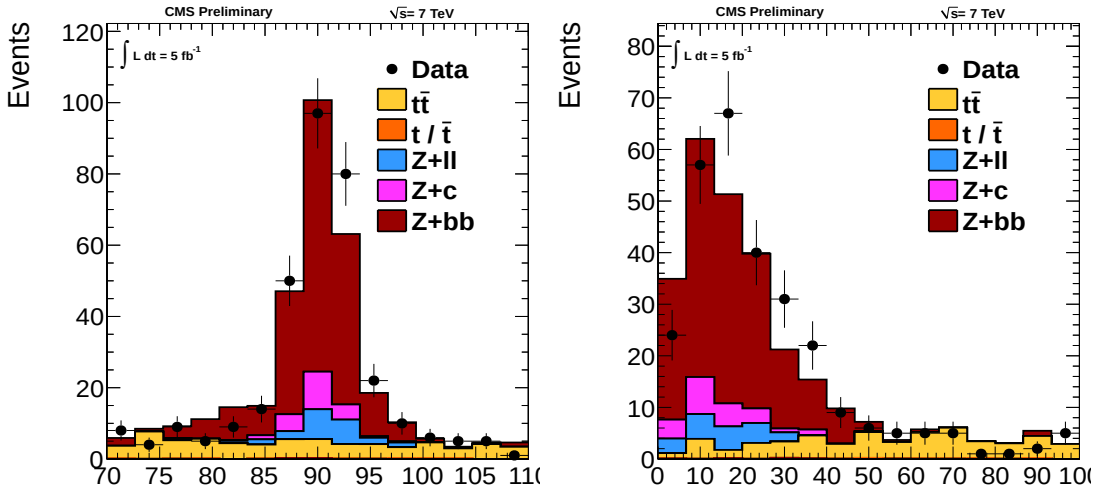
8.5 Systematic Uncertainties

One of the dominant systematic uncertainties comes from the relative uncertainty on the b-tagging efficiency (6% per jet) is taken from Ref. [69] along with the uncertainty of the light and charm jet mistagging efficiencies are. The jet and muon energy scales

are allowed to vary up and down by one standard deviation and are added to the fit as a binned shape variation. The uncertainty associated to the pileup description in Monte Carlo is estimated by shifting the overall mean of the number of vertices up or down by 0.6 bunch crossings; it has a negligible effect on the analysis. To account for the $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainty the component of $E_{\mathrm{T}}^{\mathrm{miss}}$ that is not clustered in jets is scaled by $\pm 10\%$. Uncertainties on the muon efficiency estimation (triggering, identification, isolation) are estimated to be 1%. Background normalization is also taken into account, with an uncertainty assigned to each process according to previous CMS measurements or to the described control regions. The luminosity uncertainty, 2.2%, is taken from Ref. [70].

8.6 Final Yield Extraction

The final yields are extracted via a binned maximum likelihood fit. To constrain the most prominent backgrounds and reduce the final systematic uncertainty the fit is



(a) M_{ll} [GeV] (b) E_T^{miss} [GeV]

Figure 8.9: In the Z+jets control region the invariant mass of the Z (left), and the E_T^{miss} (right)

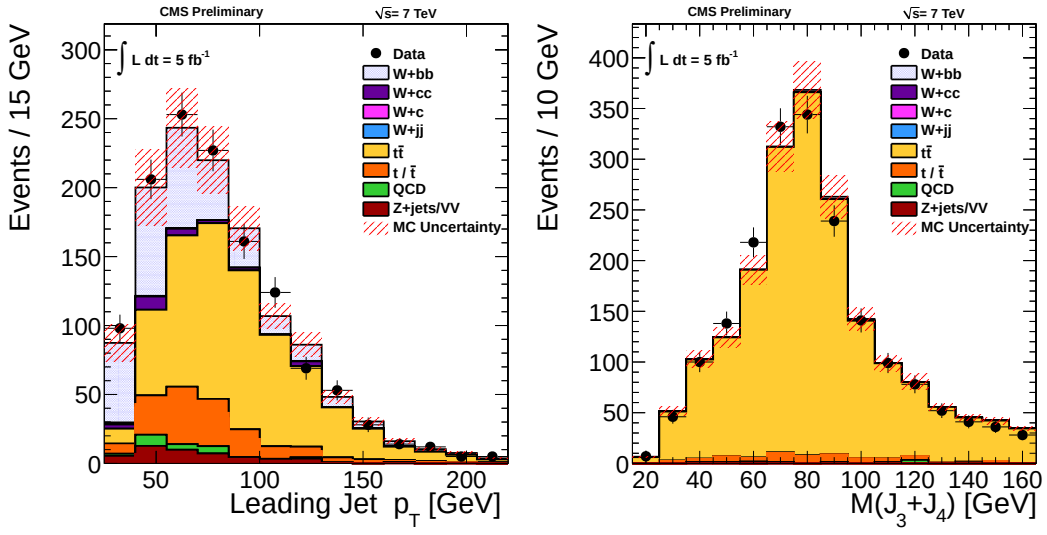


Figure 8.10: (left) The p_T distribution of the highest-p_T jet in the signal region, normalized to the result of the binned maximum likelihood fit. (right) The invariant mass of the two additional light jets in the $t\bar{t}$ control region, also normalized to the results of the fit.

performed simultaneously on the p_T of the leading jet (J_1) in the signal region after all selection requirements have been applied, and on the $m_{J_3J_4}$ distribution obtained from the $t\bar{t}$ control region. The J_1 p_T is chosen as the final fit variable due to its discrimination power against top-related backgrounds. Figure 8.10 shows the two fitted distributions, ${p_T}^{J_1}$ in the signal region (left) and $m_{J_3J_4}$ in the $t\bar{t}$ control region (right), normalized to the results of the fit.

The statistical and systematic uncertainties are introduced in the form of nuisance parameters via log-normal distributions around the estimated central values. The fitted yields for each one of the processes can be found in Table 8.2, compared to the Monte Carlo predictions. To calculate the final uncertainties, first the total errors are calculated by taking both the statistical and systematic errors into account in the fit. Then the systematic nuisance parameters are removed, the fit is re-run and the statistical uncertainty is obtained. The total systematic error is calculated by subtracting in quadrature the statistical uncertainty from the total error.

The observed number of events in data after selection in the signal region is $N(S+B)_{data} = 1230 \pm 35$. The number of signal events obtained in the binned maximum likelihood fit is 300 ± 60.

8.7 Alternate Approach

To show the robustness of the $\mathrm{W} + \mathrm{b\bar{b}}$ fit result a separate study was performed with two selected b-tagged jets that require each jet to fulfill a looser CSV b-tagging criterion, corresponding to an efficiency of 70% for jets containing b-flavored hadrons, while the misidentification probability for light-quark jets is 1%. With the exception of the modification to the CSV threshold, all other selections for the signal and control region remain unchanged. The $\mathrm{W} + \mathrm{c\bar{c}}$ contribution is non-negligible with this selection, therefore, the sum of the invariant mass of the secondary vertex found in each selected jet is used to distinguish between $\mathrm{W} + \mathrm{b\bar{b}}$ and $\mathrm{W} + \mathrm{c\bar{c}}$. The scalar sum of the transverse momenta of the muon, the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ and the jets, H_{T}, is used

Process	Prediction	Fitted Yield
$\mathrm{W} + \mathrm{b\bar{b}}$	332 ± 66	300 ± 60
$\mathrm{W} + \mathrm{c}$, $\mathrm{W} + \mathrm{c\bar{c}}$	21 ± 4	20 ± 4
W+usdg	1.5 ± 0.2	1 ± 1
Z+jets	31 ± 3	32 ± 3
$t\bar{t}$	596 ± 35	647 ± 52
Single top	160 ± 13	170 ± 13
WW, WZ	19 ± 3	17 ± 3
QCD	33 ± 17	33 ± 16
Total	1194 ± 78	1220 ± 82
Observed Events	1230 ± 35	

Table 8.2: Comparison of the expected (before the fit) and measured (after the fit) yields for each of the processes. The uncertainty on the Monte Carlo prediction takes into account the variation allowed to the nuisance parameters in the fit. The uncertainty in the fitted yields corresponds to the full uncertainty after the fit.

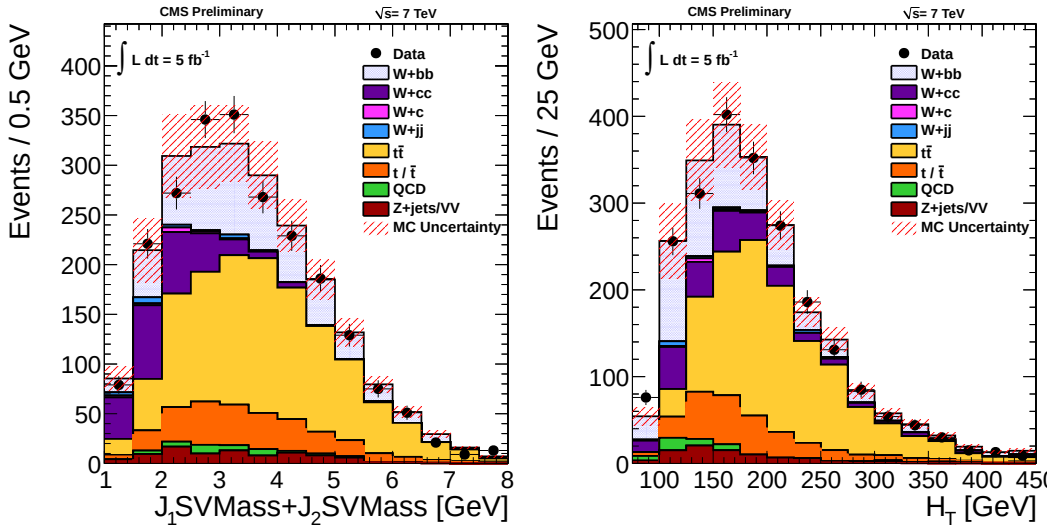


Figure 8.11: The distribution of the sum of the masses of the two secondary vertices (J_1 SV mass + J_2 SV mass) (left) and H_T of the system (right) in the alternative medium b-tag selection, normalized to the results of the cross-check fit.

to distinguish W+jets from top contributions. The $W + b\bar{b}$ signal is extracted via a two dimensional fit of H_T versus the sum of the the secondary vertex masses of the highest- (J_1) and second-highest-p_T (J_2) jets. An equivalent $t\bar{t}$ control region to the one described in the tighter selection, based on the reconstruction of the W mass using two light jets, is also used in this case. The variables J_1 SV mass + J_2 SV mass and H_T are shown in Fig. 8.11, with yields normalized to the results of the fit. The cross section value computed with this alternative method is found to be consistent with the primary fit results quoted above.

8.8 Final Cross Section Measurement

The $W + b\bar{b}$ cross section within the reference fiducial phase space is obtained using the expression

$$\sigma(pp \to W + b\bar{b}) \times \mathcal{B}(W \to \mu\nu) = \frac{N_S}{\int L dt \; \epsilon_{sel}},$$

where the efficiency of the selection requirements, $\epsilon_{\text{sel}} = (11.2 \pm 1.0)\%$, is computed using the MADGRAPH+ PYTHIA MC sample. The uncertainty in this selection efficiency comes from the PDF and scale variation uncertainties mentioned above.

The fiducial volume is defined by requiring a final-state muon with $\mathrm{p_T} > 25\text{GeV}$ and $|\eta| < 2.1$ and exactly two final-state particle jets, reconstructed using the anti-k_T jet algorithm with a distance parameter of 0.5, with $\mathrm{p_T} > 25\text{GeV}$ and $|\eta| < 2.4$ and with each containing at least one b hadron with $\mathrm{p_T} > 5\text{GeV}$. Events with extra jets are vetoed. The measured fiducial cross section is

$$\sigma(\mathrm{pp} \to \mathrm{W} + \mathrm{b\bar{b}}) \times \mathcal{B}(\mathrm{W} \to \mu\nu) = 0.53 \pm 0.05\,(\text{stat.}) \pm 0.09\,(\text{syst.}) \pm 0.06\,(\text{theo.}) \pm 0.01 lum.pb.$$

This measured value cannot be directly compared to the SM NLO cross section calculated with MCFM [8, 9] because the latter pertains to jets of partons, not jets of hadrons, and does not include the production of $\mathrm{b\bar{b}}$ pairs from double-parton scattering (DPS).

MCFM predicts a cross section of $0.52 \pm 0.03 pb$ at the parton level, using the MSTW2008 NNLO PDF set and setting the factorization and renormalization scales to $\mu_{\mathrm{F}} = \mu_{\mathrm{R}} = m_{\mathrm{W}} + 2m_b$. The 0.03 pb uncertainty in the theoretical cross section is estimated by varying the scales μ_{F}, μ_{R} simultaneously up and down by a factor of two. This uncertainty also takes into account the PDF uncertainties following the PDF4LHC recommendation. This uncertainty in the theoretical cross section may be underestimated because of the requirement of exactly two jets in the final state. Therefore, a more conservative estimate of this uncertainty in the theoretical prediction is computed, following the procedure described in Ref. [71], and the total theoretical uncertainty is found to be 30%.

Two corrections are needed to link the theoretical prediction to the measurement,

a hadronization correction and a DPS correction. At the parton level, the events are required to have a muon of $p_T > 25GeV$ and $|\eta| < 2.1$ and exactly two parton jets of $p_T > 25GeV$ and $|\eta| < 2.4$, each containing a b quark. The hadronization correction factor $C_{b\to B} = 0.92\pm0.01$, calculated using a five-flavor MADGRAPH + PYTHIA reference MC, is used to extrapolate the cross section computed at the level of parton jets to the level of final-state particle jets. The uncertainty assigned to this correction is obtained by comparing the corresponding factors computed with a four-flavored MADGRAPH MC simulation. The simulated MADGRAPH + PYTHIA events include DPS production of $b\bar{b}$ pairs and they reproduce these processes adequately as measured by CMS [72]. The contribution of DPS events to the cross section at the parton-jet level is estimated to be $\sigma_{\rm DPS} = (\sigma_{\rm W} \times \sigma_{b\bar{b}})/\sigma_{\rm eff} = 0.08 \pm 0.05 pb$. The value of the effective cross section, $\sigma_{\rm eff}$, is taken from Ref. [73], and is assumed to be independent of the process and interaction scale. The uncertainty in $\sigma_{\rm DPS}$ takes into account both the uncertainty in the measurement of $\sigma_{\rm eff}$ and the uncertainty in the fiducial $b\bar{b}$ cross section. The theoretical cross section at hadron level can be extrapolated from the MCFM parton-jet prediction by applying the hadronization correction and adding the DPS contribution, resulting in $0.55 \pm 0.03(\texttt{MCFM}) \pm 0.01(\text{had}) \pm 0.05(\text{DPS}) pb$. This value is in agreement with the measured value.

8.9 Additional $\mathrm{W} + \mathrm{b\bar{b}}$ Kinematic Distributions

In addition to this measurement of the cross section, we have explored the kinematics of the $\mathrm{W} + \mathrm{b\bar{b}}$ system. The angular distance between two selected b jets, $\Delta R(\mathrm{J}_1, \mathrm{J}_2)$ and the $M_{\rm T}$ distribution are compared to Monte Carlo predictions in Fig. 8.12. The shapes are taken from simulation and are normalized to the fit results. Figure 8.13 shows the invariant mass of the two selected b jets system and its p_T. The observed

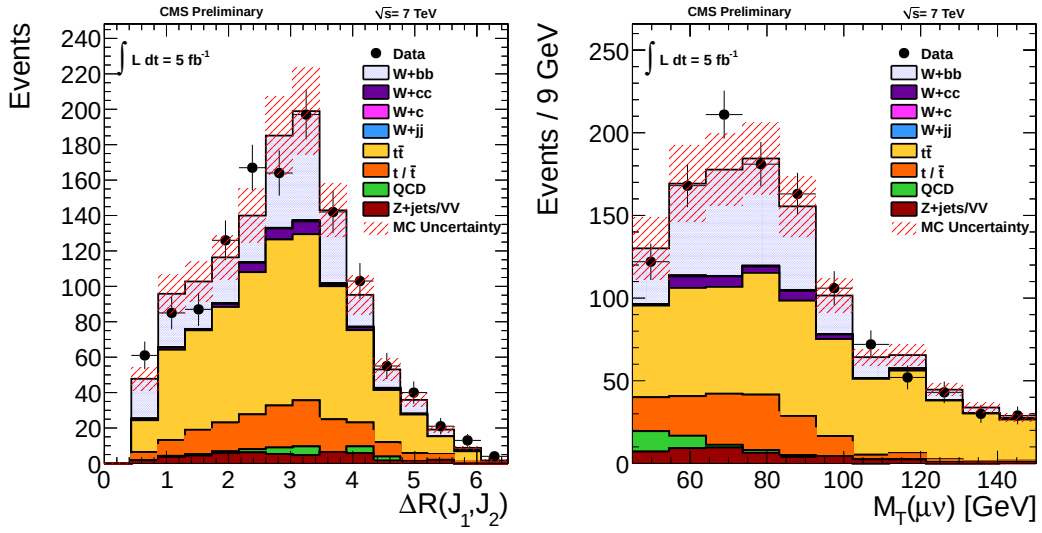


Figure 8.12: (left) The ΔR between the two selected b jets (right) the M_{T} distribution, normalized to the results of the fit.

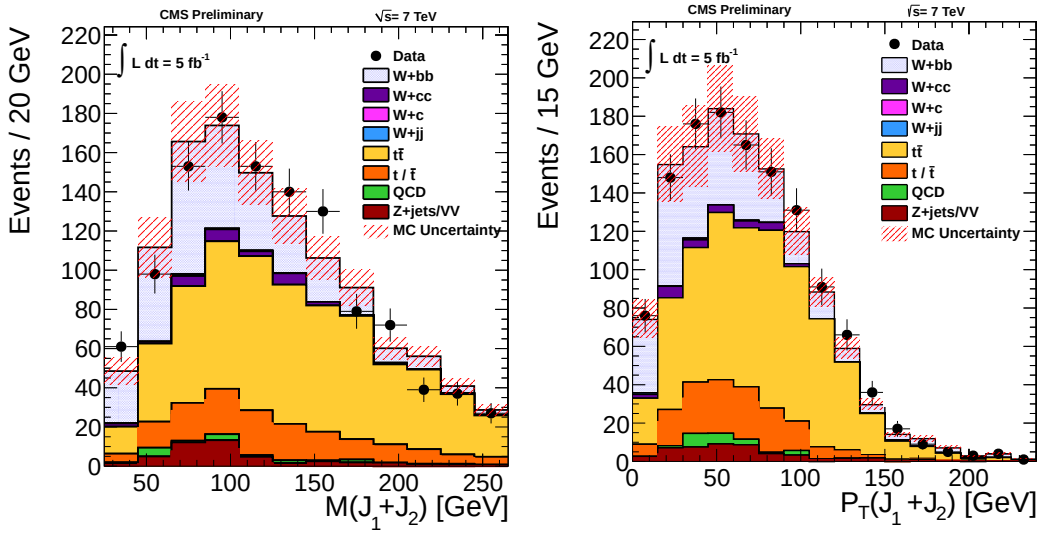


Figure 8.13: (left) The invariant mass, $m_{\mathrm{J_1J_2}}$ of the two selected b jets and (right) the $p_T(J_1J_2)$ distribution, normalized to the results of the fit.

distributions are well described by the simulation.

Chapter 9

Search for an MSSM Higgs Boson

9.1 Introduction

The dominant neutral MSSM Higgs boson production mechanism is the gluon-fusion process, gg $\rightarrow h, H, A$, for small and moderate values of tanβ. At large values of tanβ the b-associated production is the dominant contribution, due to the enhanced bottom Yukawa coupling. In the region of large tanβ the branching ratio to tau leptons is enhanced, making the search for neutral MSSM Higgs bosons in the di-τ final state of particular interest.

In this chapter, a search for neutral MSSM Higgs bosons in pp collisions corresponding to an integrated luminosity of 24.6 fb^{-1}, with 4.9 fb^{-1} at 7 TeVand 19.7 fb^{-1} at 8 TeVis presented. The different $\tau\tau$ final states studied are: $\mu\tau_h$ and eτ_h where μ indicates the muonic decay of the τ, $\tau \rightarrow \mu + \bar{\nu_\mu} + \nu_\tau$, e represents the electronic decay of the τ, $\tau \rightarrow e + \bar{\nu_e} + \nu_\tau$, and τ_h denotes a hadronic decay of a τ. The final states eμ, $\mu\mu$ and $\tau_h\tau_h$ are reported elsewhere [74]. These results are an extension of a previous search by the CMS experiment [22] and are similar to those performed by the the ATLAS experiment [75], the Tevatron [76, 77, 78], and are complimentary to

the MSSM Higgs search at LEP [79]. Figure 9.1 shows a summary of these results with exclusion at 95% CL in the $\tan\beta - M_A$.

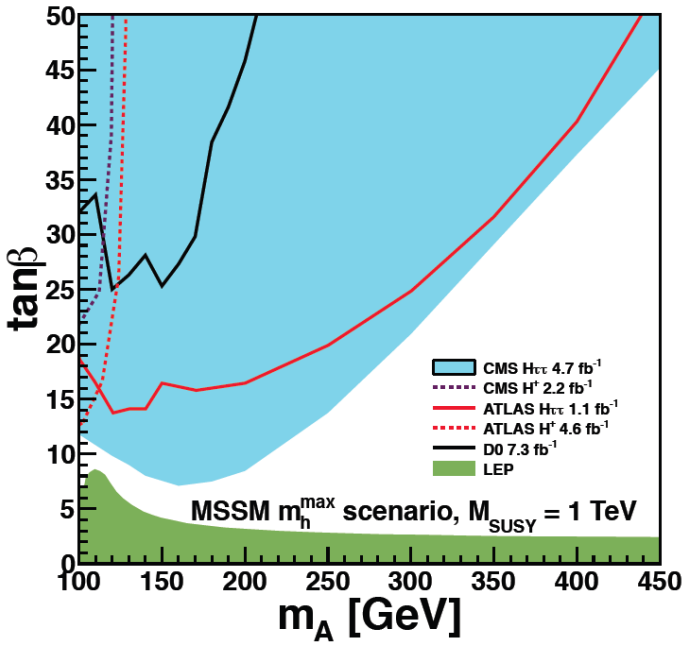


Figure 9.1: A comparison of exclusion at 95% CL in the $\tan\beta - M_A$

Traditionally, searches for MSSM Higgs bosons are expressed in terms of benchmark scenarios where the lowest-order parameters $\tan\beta$ and M_A are varied, while fixing the other parameters that enter through radiative corrections to certain benchmark values, this is further detailed in section 3.3. For one loop corrections the dominant contribution is the $O(\alpha_t)$ term due to top and stop loops. Where α_t is defined as $\alpha_t \equiv h_t^2/(4\pi)$ and h_t is the top-quark Yukawa coupling. While for two loop corrections the dominant contributions are the strong corrections and Yukawa corrections of $O(\alpha_t^2)$ to the one loop $O(\alpha_t)$ term[11]. In this study, the scenario $m_h^{\max}$ [11, 12] is used as it yields conservative expected limits in the $\tan\beta$ and M_A plane. In this scenario, the parameters are set to the following values: $M_{\rm SUSY}$ = 1TeV; X_t =2$M_{\rm SUSY}$; μ = 200 GeV; $M_{\tilde{g}}$ = 800 GeV; M_2 = 200 GeV; and $A_b = A_t$,

where M_{SUSY} is the common soft-SUSY-breaking squark mass of the third generation; $X_t = A_t - \mu/\tan\beta$ is the stop mixing parameter; A_t and A_b are the stop and sbottom trilinear couplings, respectively; μ the Higgsino mass parameter; $M_{\tilde{g}}$ the gluino mass; and M_2 is the SU(2)-gaugino mass parameter.

An indication that the recently discovered higgs boson decays into tau pairs has been reported by CMS [80]. If the new boson is interpreted as the light scalar MSSM Higgs h, part of the tanβ and M_A parameter space in the m_h^{max} scenario is excluded. However, changes in the stop mixing parameter open up a large region of the allowed parameter space [81, 82].

The results are interpreted both in the context of the MSSM m_h^{max} scenario and also in a model independent way, As noted above, the m_h^{max} scenario fixes SUSY parameters in order to estimate higgs production cross sections as function of M_A and $\tan\beta$. The model independent method simply puts an exclusion at 95% CL on the production cross section of $m(\phi)$ This exclusion is performed in terms of upper limits on $\sigma\cdot\mathrm{BR}(\Phi \to \tau\tau)$ for gluon-fusion and b-associated neutral Higgs boson production, where we denote by Φ any of the three neutral MSSM Higgs bosons.

9.2 Trigger and Event Selection

The trigger selection requires a combination of electron, muon and tau trigger objects [83, 84, 85]. The identification criteria and transverse momentum thresholds of these objects were progressively tightened as the LHC instantaneous luminosity increased over the data-taking period.

A particle-flow algorithm [86, 87, 88] is used to combine information from all CMS subdetectors to identify and reconstruct individual particles in the event, namely muons, electrons, photons, and charged and neutral hadrons. From the resulting

particle list jets, hadronically decaying taus, and missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$), defined as the magnitude of the vector sum of the transverse momenta, are reconstructed. The jets are reconstructed using the anti-k_T jet algorithm [89, 90] with a distance parameter of $R = 0.5$. Hadronically-decaying taus are reconstructed using the hadron plus strips (HPS) [13] algorithm, which considers candidates with one charged pion and up to two neutral pions or three charged pions. To tag jets coming from b-quark decays the Combined Secondary Vertex (CSV) algorithm is used, this is further detailed in section 6.9. This algorithm is based on the reconstruction of secondary vertices, together with track-based lifetime information [91].

Events in the eτ_h ($\mu\tau_h$) final state are required to contain an electron of $\mathrm{p_T} > 20$ GeV (muon $\mathrm{p_T} > 17$ GeV) and $|\eta| < 2.1$ plus an oppositely charged τ_h of $\mathrm{p_T} > 20$ GeV and $|\eta| < 2.3$. The $\mathrm{p_T}$ thresholds for electrons (muons) are increased to 24 GeV (20 GeV) in the 2012 dataset, following the raise in trigger thresholds at higher instantaneous luminosity.

Events in the eτ_h final state reject $\mathrm{Z} \rightarrow$ ee background by rejecting events with a second electron of opposite sign that passes the 'loose electron' requirement. To reject the $t\bar{t}$ background an event is rejected if a second electron of either opposite or same sign passing 'tight electron' selection requirements or if an event contains a muon with 'tight muon' selection. $\mathrm{Z} \rightarrow$ ee background in the eτ_h final state is further suppressed by requiring $E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV. Events in the $\mu\tau_h$ final state reject $\mathrm{Z} \rightarrow \mu\mu$ background by rejecting events with a second muon of opposite sign that passes the loose muon requirement. To reject the $t\bar{t}$ background an event is rejected if a second muon of either opposite or same sign passing 'tight muon' selection requirements or if an event contains a electron with 'tight electron' selection. Electron and muon definitions are described as:

- **Tight Electron:** $P_T^e > 10$ GeV and $\eta_e < 2.5$, reconstructed with Electron ID

criteria recommended by the EGamma POG [92] 'loose ID', with loose isolation requirements $I_{rel} < 0.3$ with I_{rel} computed as described in section 6.4.

- **Loose Electron:** $P_T^e > 15$ GeV and $\eta_e < 2.5$, reconstructed with Electron ID criteria recommended by the EGamma POG [92] 'Working Point 95 Veto' and with loose isolation requirements $I_{rel} < 0.3$ with I_{rel} computed as described in section 6.4.

- **Tight Muon:** $P_T^\mu > 10$ GeV and $\eta_e < 2.4$, reconstructed with muon tight ID criteria recommended by the Muon POG reconstructed as global and as tracker muon, passing the particle–flow muon identification criteria [93] and with loose isolation requirements $I_{rel} < 0.3$ with I_{rel} computed as described in section 6.4.

- **Loose Muon:** $P_T^\mu > 15$ GeV and $\eta_\mu < 2.4$, reconstructed as global and as tracker muon, passing the particle–flow muon identification criteria [93] and with loose isolation requirements $I_{rel} < 0.3$ with I_{rel} computed as described in section 6.4.

An average of 10 (20) proton-proton interactions occurred per LHC bunch crossing in 2011 (2012), making the reconstruction of physics objects challenging. For each reconstructed collision vertex the sum of the $\mathrm{p_T}^2$ of all tracks associated with the vertex is computed and the one with the largest value is taken as the primary collision vertex. In order to mitigate the effects of pile–up on the reconstruction of $E_{\mathrm{T}}^{\mathrm{miss}}$, a multivariate regression correction (further detailed in section 6.10) is used where the inputs are separated in those components coming from the primary vertex and those which are not [94]. The correction improves the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution in $\mathrm{Z} \to \mu\mu$ events by roughly a factor of two in the case where 25 additional pile-up events are present.

Taus from Higgs boson decays are expected to be isolated in the detector, while leptons from heavy-flavor (c and b) decays and decays in flight are expected to be

found inside jets. Lepton isolation (described in detail in section 6.4) is used to discriminate the signal from the QCD multijet background, based on the charged hadrons, photons, and neutral hadrons falling within a cone around the lepton momentum direction.

To correct for the contribution to the jet energy due to pile-up, a median energy density (ρ) is determined event by event. The pile-up contribution to the jet energy is estimated as the product of ρ and the area of the jet and subsequently subtracted from the jet transverse energy [95]. In the fiducial region for jets of $|\eta| < 4.7$, jet energy corrections are also applied as a function of the jet $E_{\mathrm{T}}$and η [68].

In order to reject events coming from W+jets background a dedicated selection is applied. In the eτ_h and $\mu\tau_h$ final states, the transverse mass of the electron or muon and the $E_{\mathrm{T}}^{\mathrm{miss}}$ (the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution is shown in figure 9.6), $M_T = \sqrt{2\mathrm{p_T}E_{\mathrm{T}}^{\mathrm{miss}}\,(1-\cos\Delta\phi)}$, is required to be less than 30 GeV, where $\mathrm{p_T}$ is the lepton transverse momentum and $\Delta\phi$ is the difference in ϕ of the lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ vector. Figure 9.7 shows the M_T distribution.

The μ p_T and η distributions for 7 and 8 TeV are shown in figure 9.3, while the electron p_T and η distributions are shown in figure 9.2. The τ p_T and η distributions for eτ_h and $\mu\tau_h$ final states are shown in figure 9.4 and figure 9.5. Figure 9.8 shows the visible mass distribution which is the invariant mass of the light lepton and the visible decay products of the τ. The $m_{\tau,\tau}$ distribution using the SVFit algorithm is shown in figure 9.9 for the inclusive selection, figure 9.10 for the no b-tag selection and 9.11 for the b-tag selection.

To further enhance the sensitivity of the search for Higgs bosons, the sample of selected events is split into two mutually exclusive categories:

- **B-Tag:** This event category is intended to exploit the production of Higgs bosons in association with b-quarks which is enhanced in the MSSM. At least

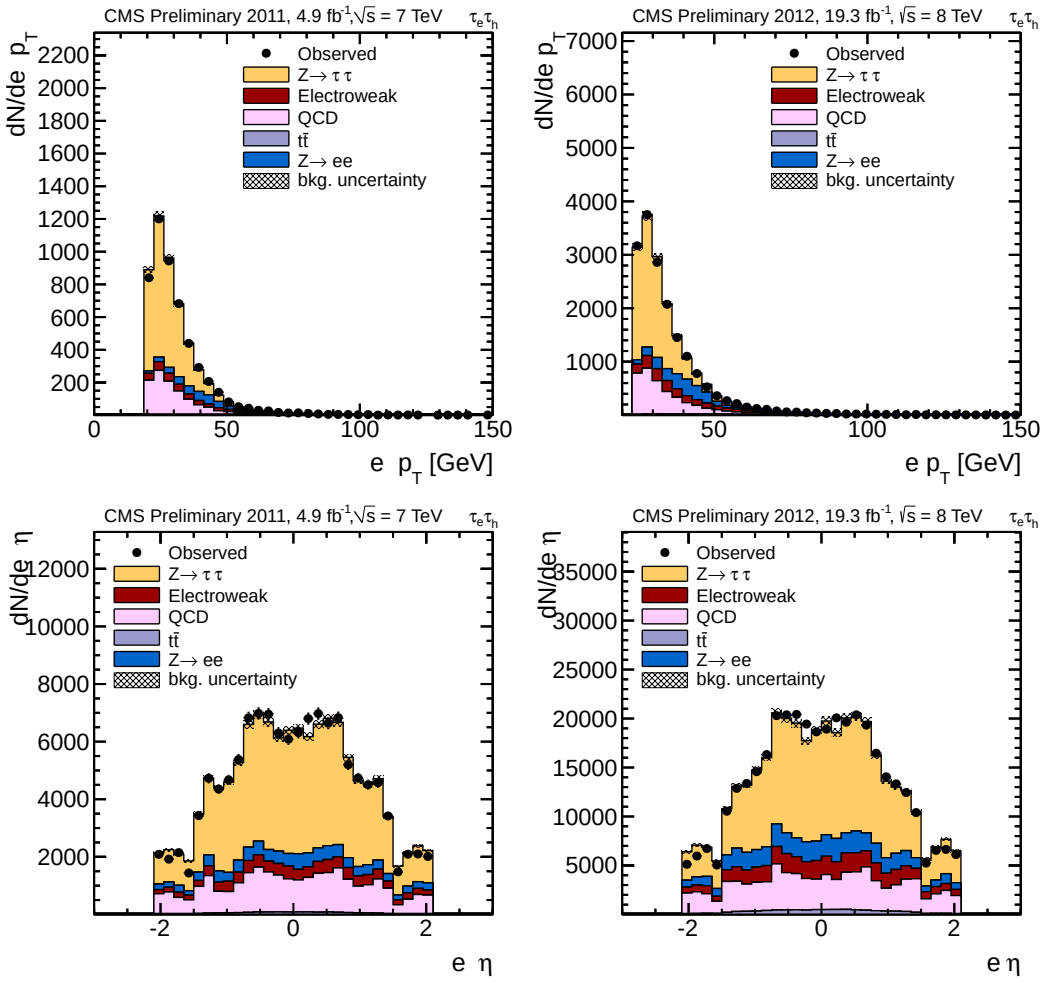


Figure 9.2: Distribution of the electron transverse momentum and psuedorapidity.

one b-tagged jet with $p_T > 20$ GeV is required and not more than one jet with $p_T > 30$ GeV.

- **No-B-Tag:** This event category is mainly sensitive to the gluon-fusion Higgs production mechanism. Events are required to have no b-tagged jets with $p_T > 20$ GeV.

The number of b-tagged jets in the $\tau_\mu\tau_h$ and $\tau_e\tau_h$ channels for both 7 and 8 TeV can be seen in figure (9.15). Figures 9.13 and 9.14 show the highest jet p_T and η

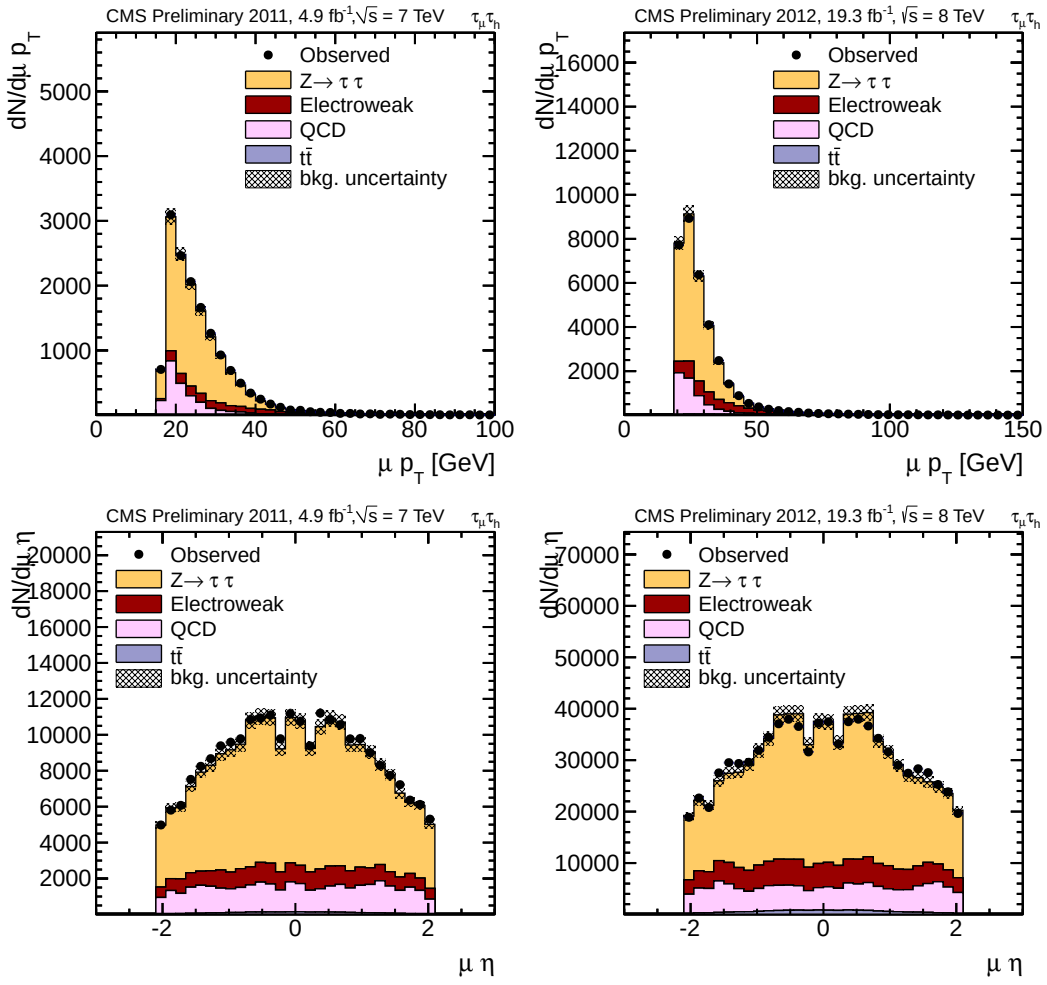


Figure 9.3: Distribution of the muon transverse momentum and psuedorapidity.

distributions for the $\tau_\mu\tau_h$ and $\tau_e\tau_h$ channels respectively.

The observed number of events for each category, as well as the expected number of events from various background processes, are shown in Tables 9.1 and 9.2 together with expected signal yields and efficiencies.

The largest source of background events comes from $Z \to \tau\tau$, which is modeled using a sample of $Z \to \mu\mu$ events where the reconstructed muons are replaced by the reconstructed particles from simulated tau decays. The normalization for this process is determined from the measurement of the $Z \to \mu\mu$ yield in data.

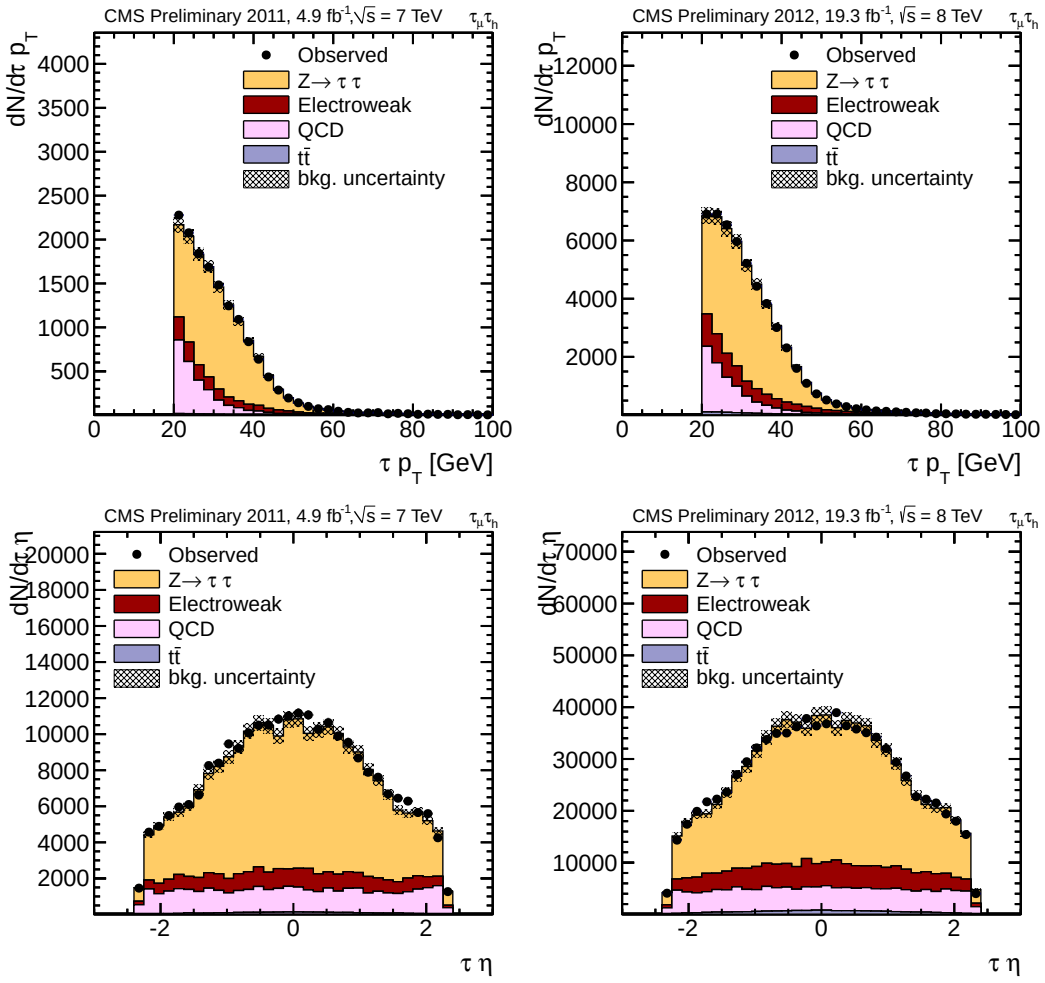


Figure 9.4: Distribution of the τ_h transverse momentum and psuedorapidity.

Another significant source of background is QCD multijet events. QCD events may contribute to the eτ_h and $\mu\tau_h$ channel in case one jet is misidentified as an isolated electron or muon, and a second jet as τ_h. In the eτ_h and $\mu\tau_h$ channels the rate of QCD background is estimated using the number of observed same-sign tau pair events.

Events from W+jets in which there is a jet misidentified as a τ_h are another sizeable source of background in the eτ_h and $\mu\tau_h$ channels. The rate of this background is estimated using a control region of events with large transverse mass, $m_T > 70$ GeV.

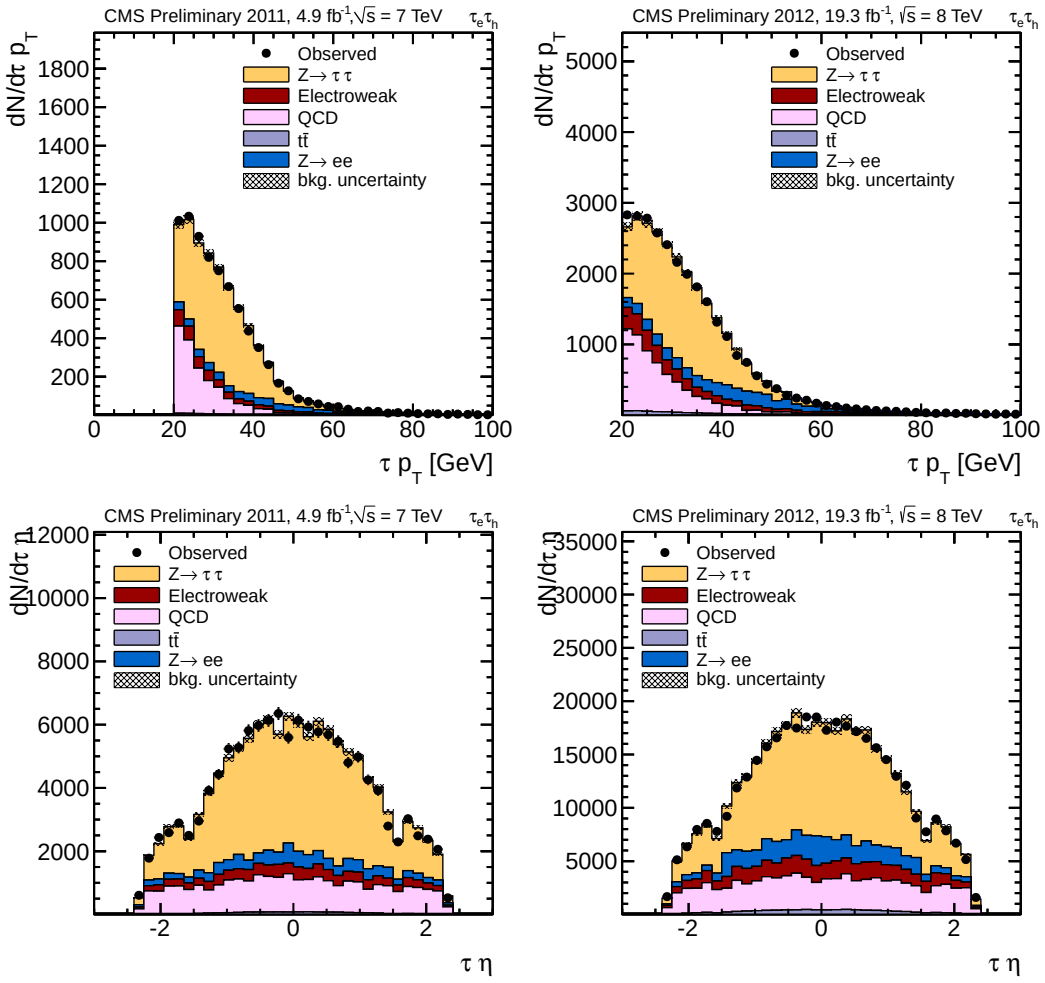


Figure 9.5: Distribution of the τ_h transverse momentum and psuedorapidity.

The pre-fit W+jets normalization Other background processes include $q\bar{q}$ production and $Z \rightarrow ee/\mu\mu$ events, particularly in the $e\tau_h$ channel, as the probability for electrons to be misidentified as τ_h amounts to 1–2%. The shape of the $t\bar{t}$ and di-boson backgrounds are estimated from simulation using `MADGRAPH` [42] and `PYTHIA` [44], respectively. The event yield for $t\bar{t}$ is determined from measurements in a background-enriched region which exhibits high jet multiplicity. The number of jets is shown in figure 9.12, in particular, good agreement between signal and background is seen for high jet multiplicity where $t\bar{t}$ dominates.

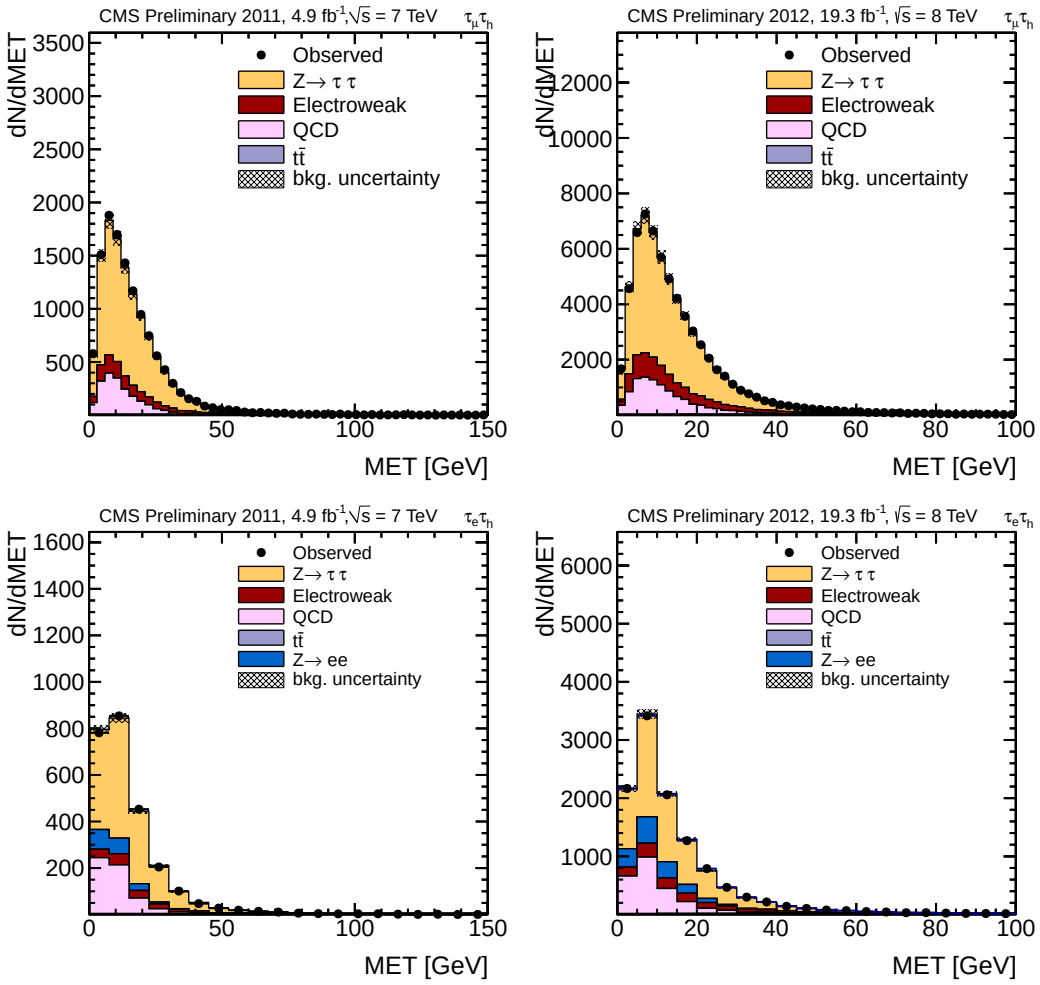


Figure 9.6: Distribution of the missing transverse momentum from the $\tau_\mu\tau_h$ (top) and $\tau_e\tau_h$ (bottom) channels.

The event generator PYTHIA is used to model the MSSM Higgs boson signal. Taus are decayed by the TAUOLA [43] package. In all Monte Carlo samples, additional interactions are simulated and reweighted to the observed pile-up distribution. The missing transverse energy in Monte Carlo simulated events is corrected for the difference between data and simulation measured using a sample of $Z \to \mu\mu$ events [96].

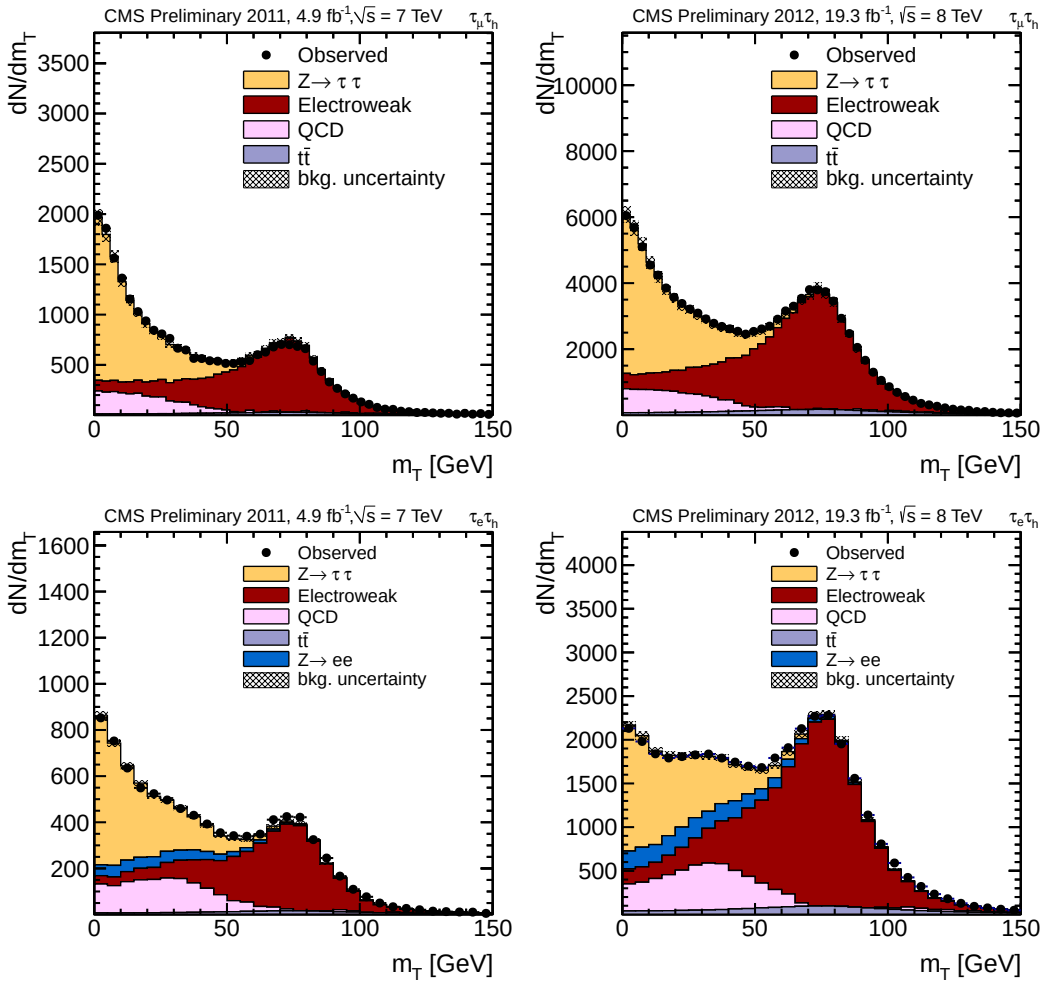


Figure 9.7: Distribution of the transverse mass from the $\tau_\mu\tau_h$ (top) and $\tau_e\tau_h$ (bottom) channels.

9.3 Tau-pair invariant mass reconstruction

To distinguish the signal of Higgs bosons from the background, the tau-pair mass, $M_{\tau\tau}$, is reconstructed using a maximum likelihood technique [22] which is further detailed in section 6.6. The algorithm computes the tau-pair mass that is most compatible with the observed momenta of visible tau decay products and the missing transverse energy reconstructed in the event. Free parameters, corresponding to the missing neutrino momenta, are subject to kinematic constraints and are eliminated

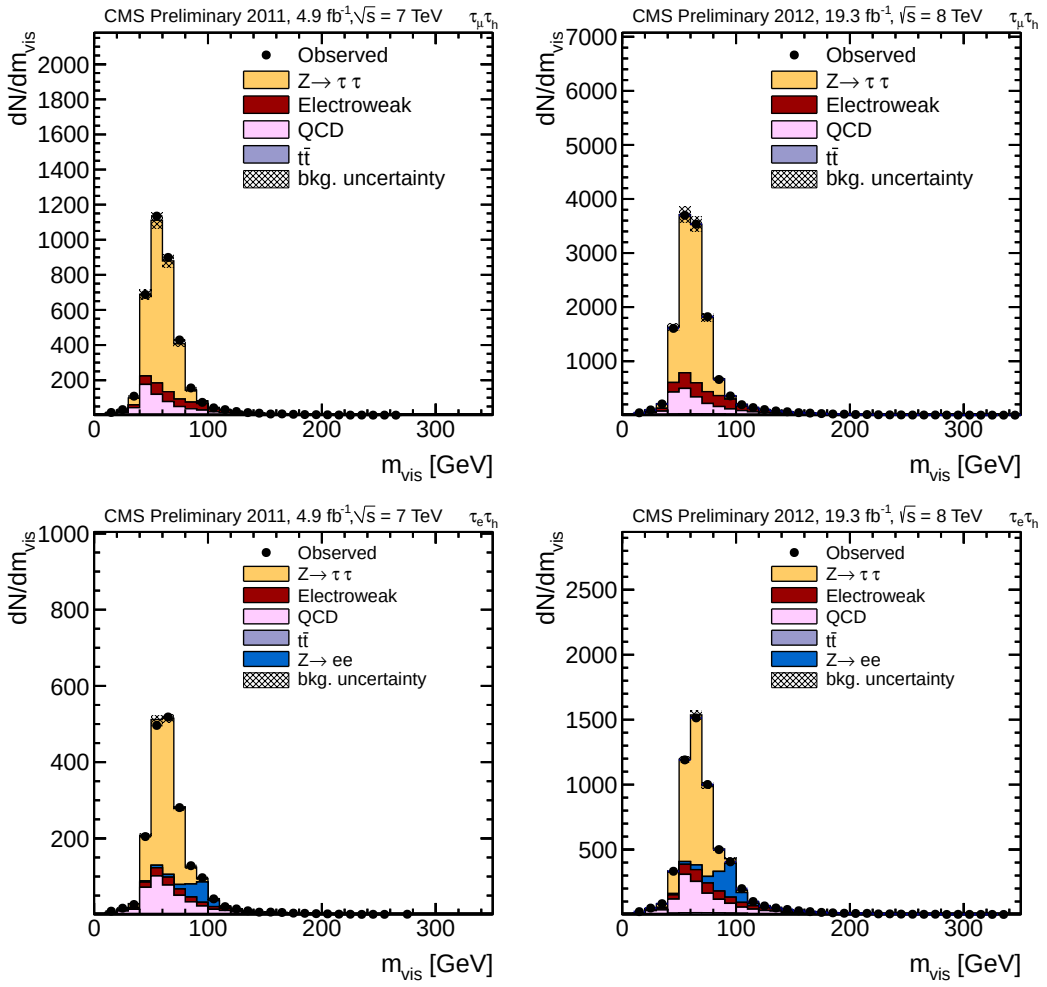


Figure 9.8: Distribution of the Visible Mass for inclusive selection.

by performing a fit to determine a value for these parameters. The algorithm yields a tau-pair mass distribution consistent with the true value and a resolution of 15-20%.

9.4 Systematic uncertainties

Various imperfectly known or simulated effects can alter the shape and normalization of the invariant mass spectrum. The main contributions to the normalization uncertainty include the uncertainty in the total integrated luminosity (4.5% for 2011 and

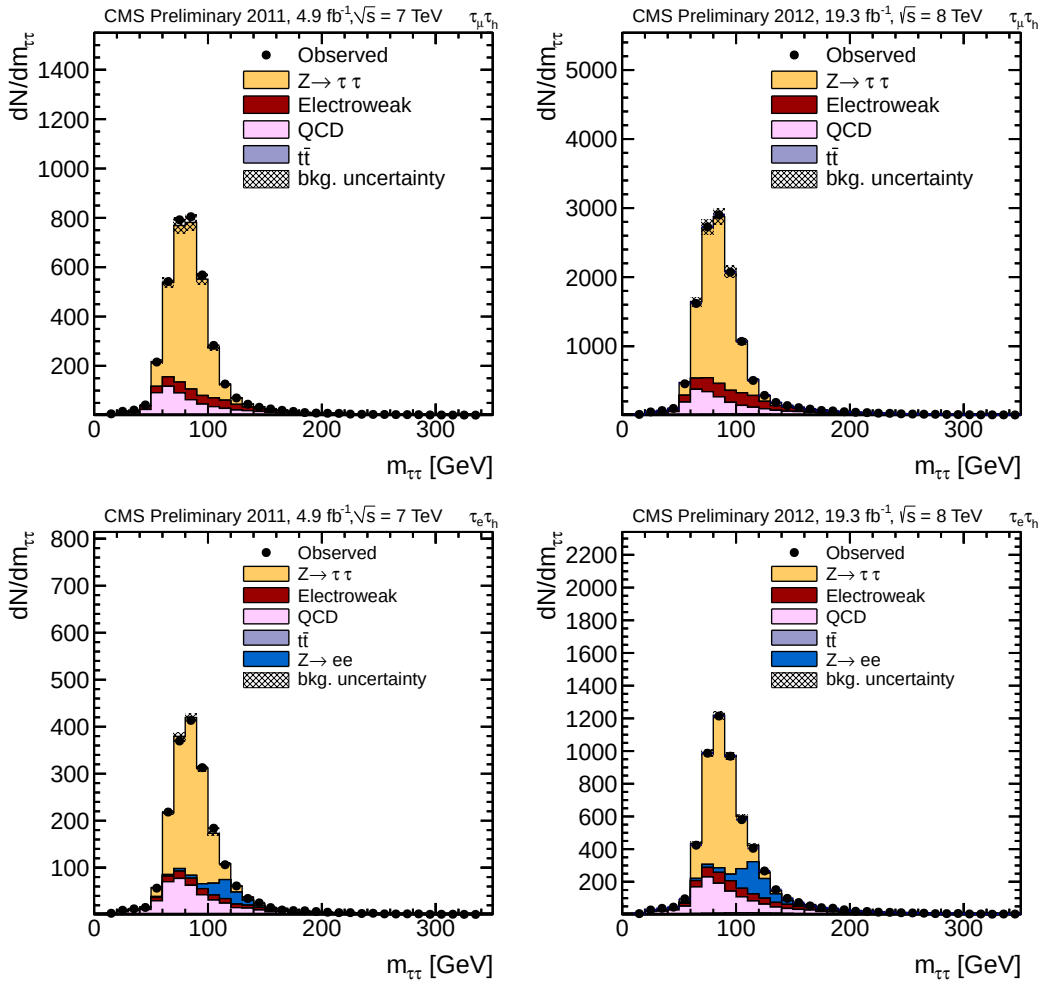


Figure 9.9: Distribution of the SVFit Mass for inclusive selection.

2.6% for 2012 data) [97], jet energy scale (2–5% depending on η and p_T)[98], background normalization (Tables 9.1– 9.2), Z boson production cross section (2.5%) [96], lepton identification and isolation efficiency (1.0%), and trigger efficiency (1.0%). The tau-identification efficiency uncertainty is estimated to be 7% from an independent study done using a tag-and-probe technique [96] including the uncertainty of the trigger efficiency. The lepton identification and isolation efficiencies are stable as a function of the number of additional interactions in the bunch crossing in data and in Monte Carlo simulation. The b-tagging efficiency has an uncertainty of 10%, and

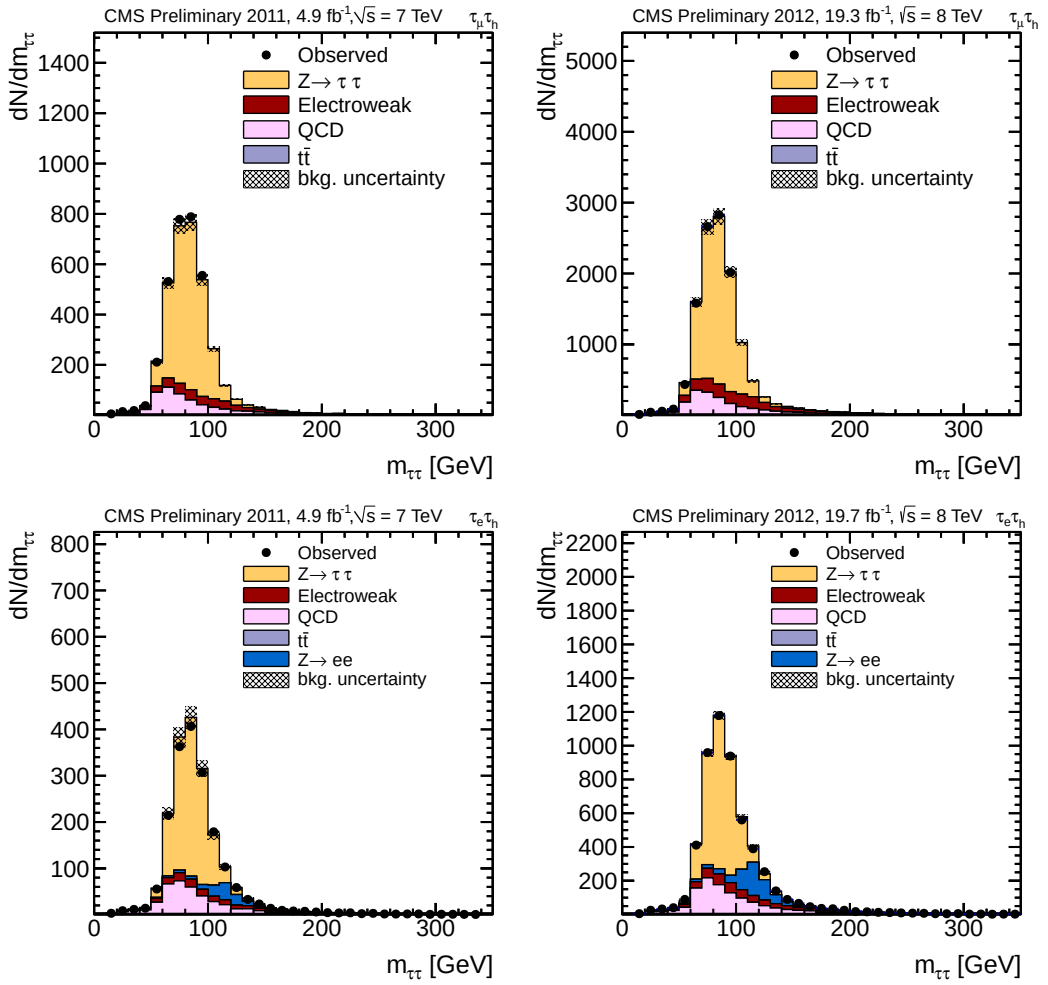


Figure 9.10: Distribution of the SVFit Mass for the no b-Tag selection.

the b-mistag rate is accurate to 30% [91]. Uncertainties that contribute to mass spectrum shape variations include the tau (3%), muon (1%), and electron (1.5%) energy scales[99, 74]. The effect of the uncertainty on the $E_{\mathrm{T}}^{\mathrm{miss}}$ scale, mainly due to pile-up effects, is incorporated by varying the mass spectrum shape as described in the next section. The neutral MSSM Higgs production cross sections and the corresponding uncertainties are provided by the LHC Higgs Cross Section Group [14]. The cross sections have been obtained from the GGH@NNLO [100, 101, 102, 103, 104] and HIGLU [105, 106] programs for the gluon-fusion process. For the $b\bar{b} \to \Phi$ process,

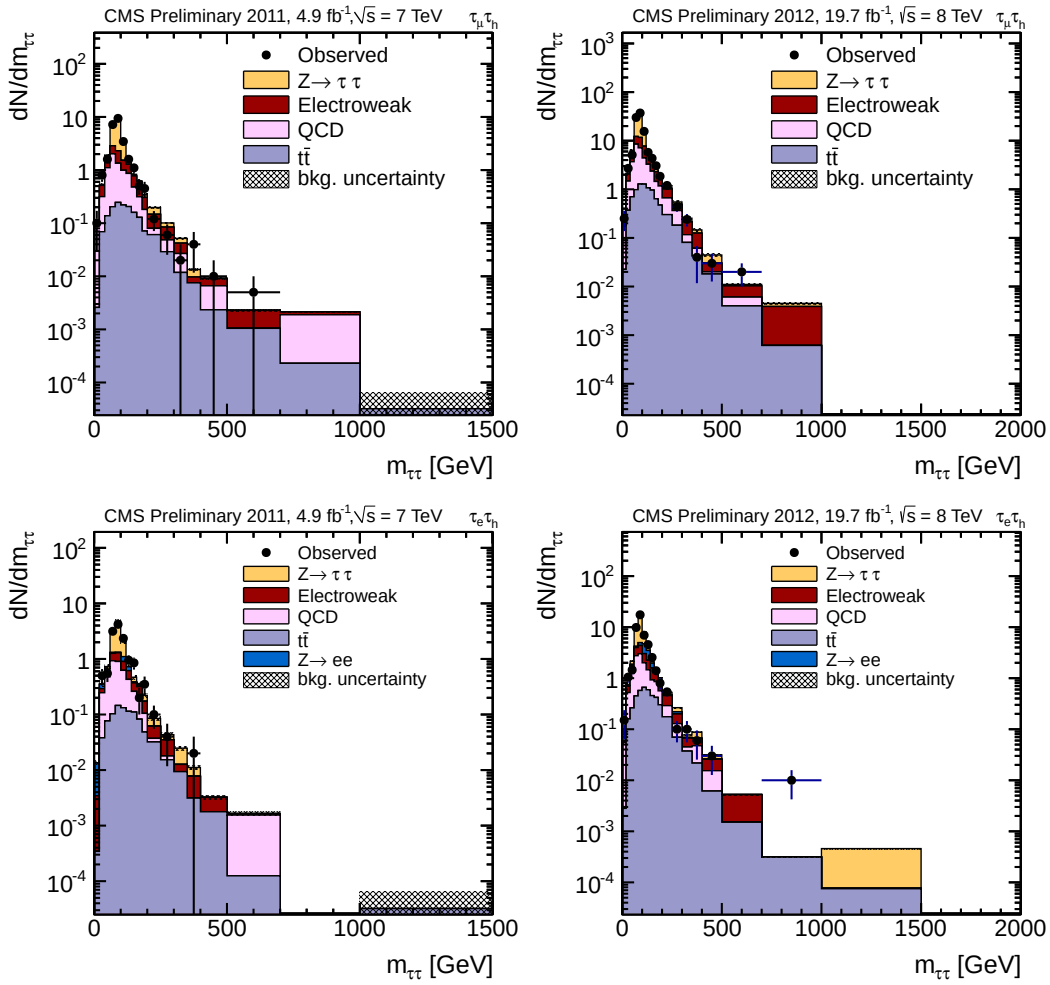


Figure 9.11: Distribution of the SVFit Mass for the b-Tag selection.

the four-flavor calculation [107, 108] and the five-flavor calculation as implemented in BBH@NNLO [109] have been combined using the Santander matching scheme [110]. Yukawa couplings for the m_h^{max} MSSM benchmark scenario have been calculated by FeynHiggs [111, 112, 113]. The uncertainties for the MSSM signal depends on $\tan\beta$ and M_A and can amount up to 25%. The MSTW2008 proton distribution function is used, and the associated uncertainties range from 2-10%. The renormalization and factorization scale uncertainties amount to 5-25% in the gluon-fusion process and 8-15% in the associated-b process.

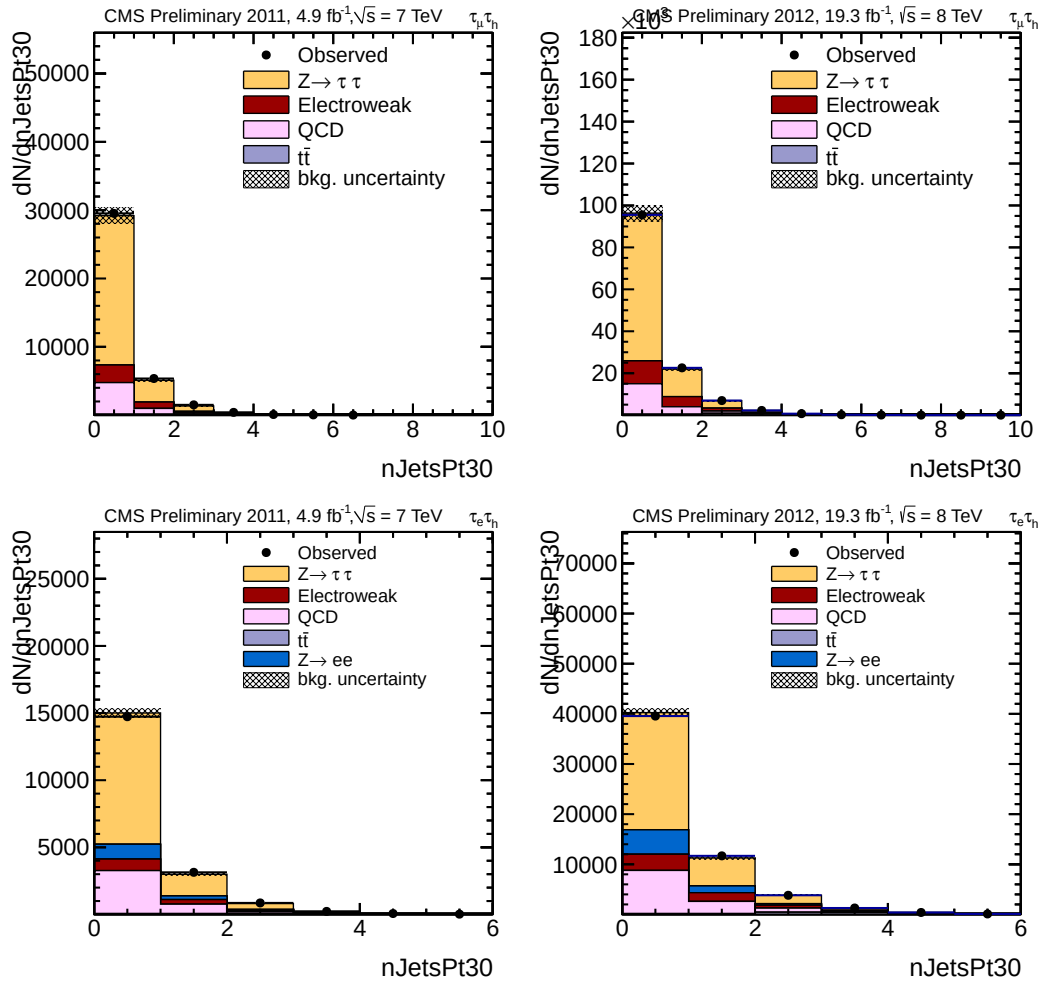


Figure 9.12: Distribution of the number of identified jets in the sample of selected events.

9.5 Results

To search for the presence of a Higgs boson signal in the selected events, a binned maximum likelihood fit is performed. The tau-pair invariant-mass spectrum is used as input for the fit in the eτ_h, $\mu\tau_h$, eμ and $\tau_h\tau_h$ final states. The eμ and $\tau_h\tau_h$ final states are detailed elsewhere [74]. The fit is performed simultaneously for the five final states eτ_h, $\mu\tau_h$, eμ, $\mu\mu$ and $\tau_h\tau_h$ and the two event categories B-tag and no-B-tag. Systematic uncertainties are represented by nuisance parameters in the fitting process. Log-

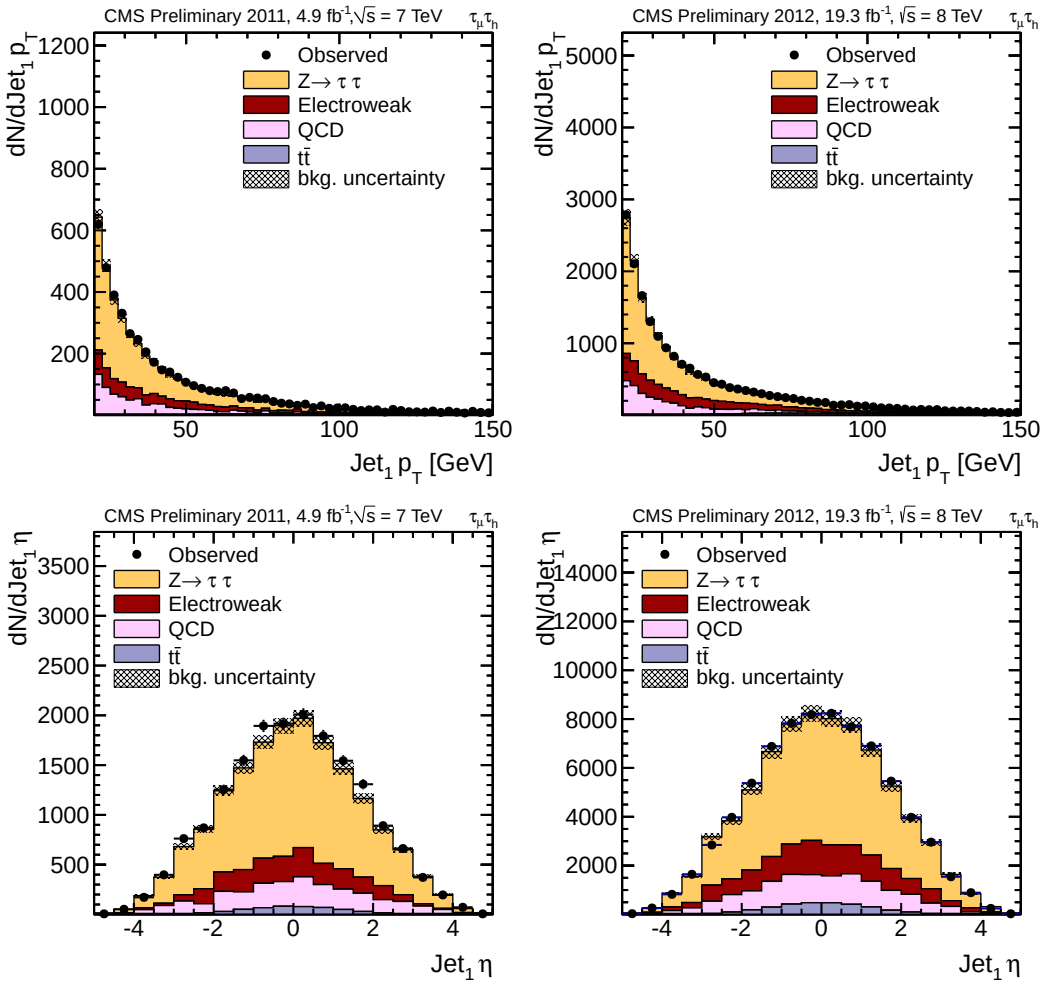


Figure 9.13: Distribution of the leading jet transverse momentum and psuedorapidity.

normal priors are assumed for the normalization parameters, and Gaussian priors for mass-spectrum shape uncertainties. The uncertainties that affect the shape of the mass spectrum, mainly those corresponding to the energy scales, are represented by nuisance parameters whose variation results in a continuous perturbation of the spectrum shape. In the regions of $M_{\tau\tau} > 150$ GeV, where the event statistic of the background templates is reduced, a fit of the form $f = \exp\left(-\frac{M_{\tau\tau}}{c_0+c_1\cdot M_{\tau\tau}}\right)$ is performed, and the uncertainties on the fit parameters c_0 and c_1 are propagated.

The signal expectation is determined in each point of the parameter space as

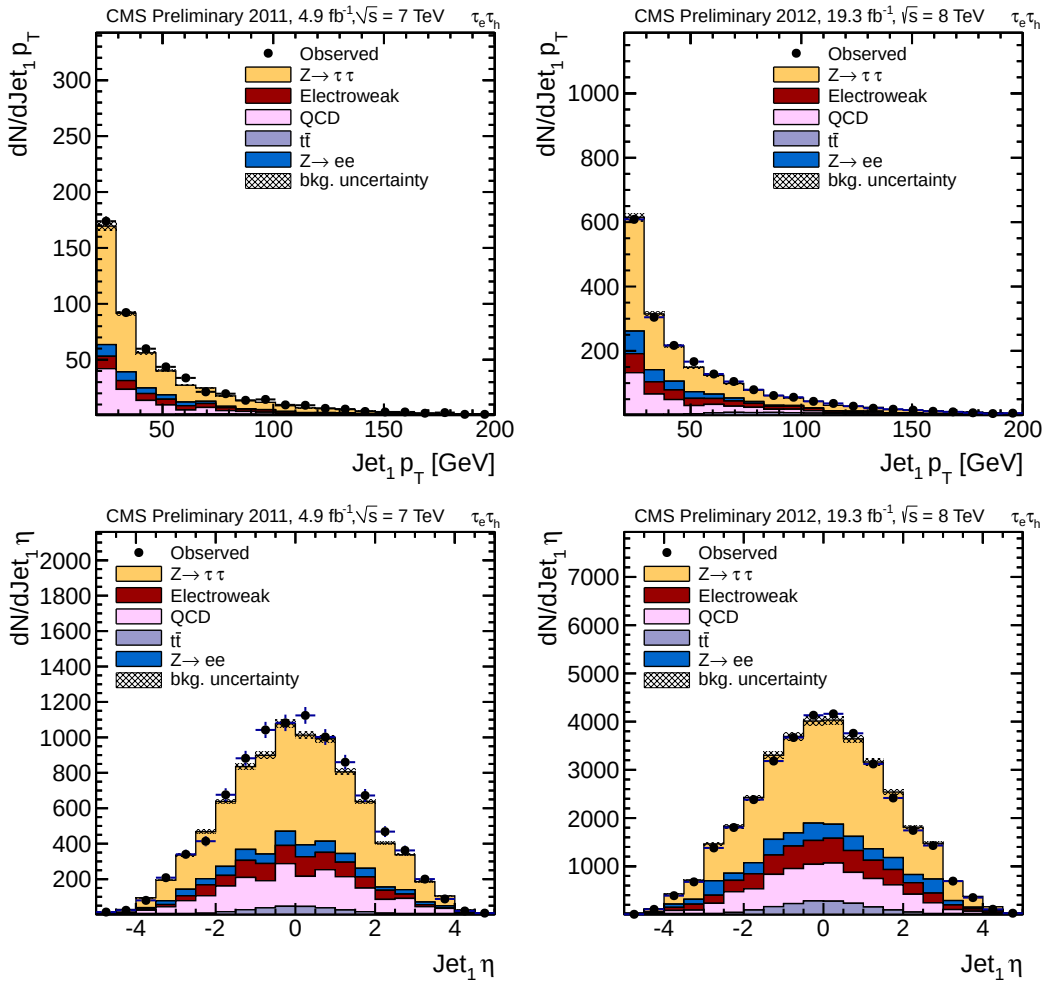


Figure 9.14: Distribution of the leading jet transverse momentum and psuedorapidity.

follows:

- At each point of M_A and tanβ the mass, the gluon-fusion and associated-b production cross sections and the branching ratio to $\tau\tau$ are determined for h, H and A.

- For each neutral Higgs boson the expected reconstructed di-τ mass is obtained via "horizontal template morphing" [114], using as input the $M_{\tau\tau}$ shape templates of the nearest lower and upper mass–points for which Monte Carlo sam-

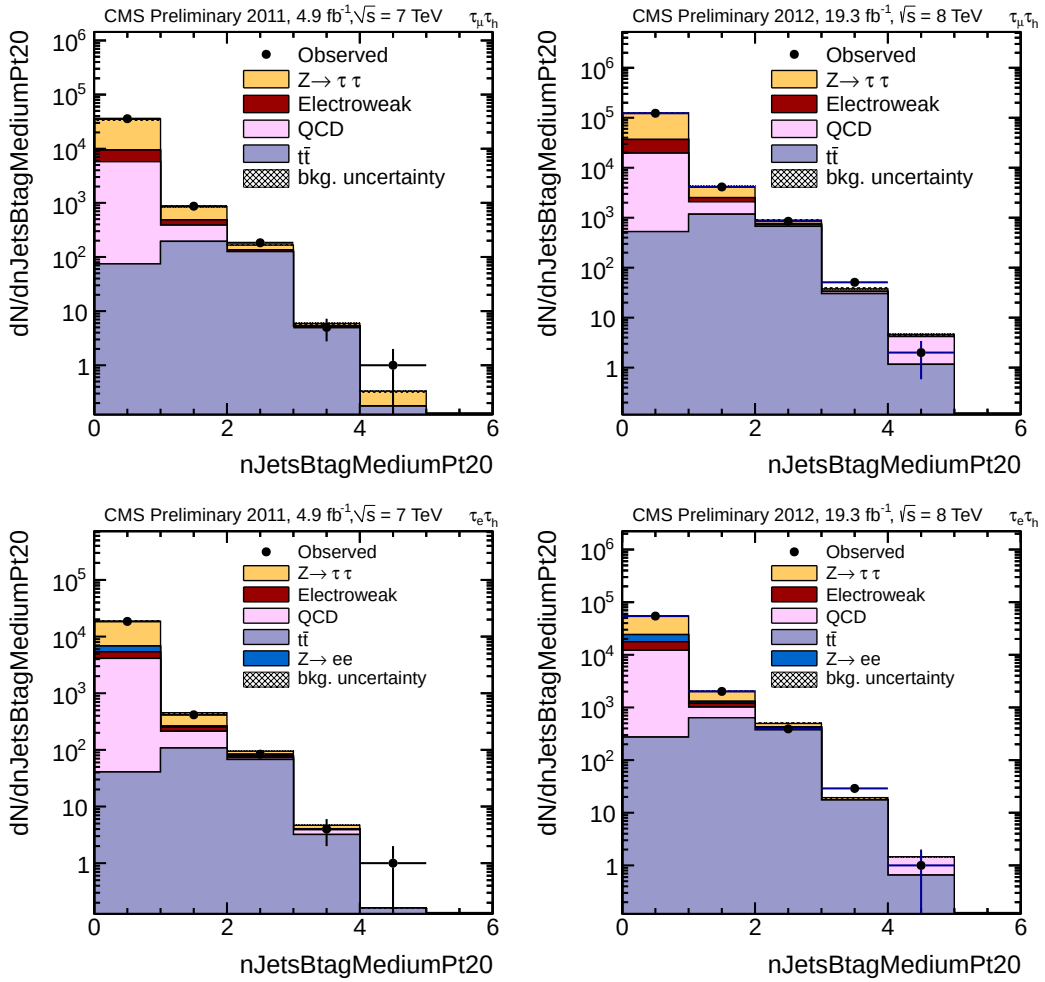


Figure 9.15: Distribution of the number of identified b-quark jets in the sample of selected events from $\tau_\mu\tau_h$ (top) and $\tau_e\tau_h$ (bottom) channels. The background components are scaled relative to the results of a fit to the inclusive mass distribution. The uncertainty on the background is represented in the shaded region and comes from the fitted nuisances.

Table 9.1: Number of expected events in the two event categories in the $\mu\tau_{\mathrm{h}}$ channel, where the combined statistical and systematic uncertainty is shown. The signal yields for the sum of all three neutral MSSM Higgs bosons, A+H+h, expected for M_A= 160 GeV and $\tan\beta$= 8 are given for comparison. Also given are the products of signal efficiency times acceptance for M_Φ= 160 GeV.

$\mu\tau_{had}$ channel				
	$\sqrt{s}$ = 7 TeVdata		$\sqrt{s}$ = 8 TeVdata	
	no–B–tag	B–tag	no–B–tag	B–tag
$gg \to \Phi + gg \to \Phi b$	228 ± 17	17 ± 2	946 ± 59	68 ± 6
$Z/\gamma^* \to \tau\tau$	26424 ± 2319	278 ± 26	85600 ± 7364	1103 ± 100
$Z/\gamma^* \to \ell\ell$ ($\ell = e, \mu$)	736 ± 129	11 ± 2	3496 ± 625	52 ± 10
W + jets	2882 ± 288	51 ± 15	12670 ± 1267	240 ± 72
QCD	4950 ± 495	134 ± 27	17913 ± 1791	551 ± 110
$t\bar{t}$	86 ± 12	38 ± 8	592 ± 84	184 ± 35
Single top + di–boson	107 ± 20	14 ± 3	516 ± 89	59 ± 10
Total background	35185 ± 2392	526 ± 41	120787 ± 7710	2189 ± 170
Data	36055	542	123239	2219

Signal Eff.

	B–tag	no–B–tag
gg$\to \Phi$	$2.36 \cdot 10^{-4}$	$1.91 \cdot 10^{-2}$
bb$\to \Phi$	$2.82 \cdot 10^{-3}$	$1.63 \cdot 10^{-2}$

ples have been produced.

- The contributions of all three neutral Higgs boson are added using the corresponding cross sections times branching fraction.

Fig. 9.16 shows the distribution of the tau-pair mass for the five final states in the no-B-Tag category, which is more sensitive to the gluon-fusion production mechanism, compared with the background prediction. Fig. 9.17 shows the distribution of the tau-pair mass for the five final states in the B-Tag category, which enhances the sensitivity to the bb$\to \Phi$ production mechanism.

The invariant mass spectra show no clear evidence for the presence of a MSSM Higgs boson signal, therefore 95% CL upper bounds on tanβ as a function of the

Table 9.2: Number of expected events in the two event categories in the $e\tau_{\mathrm{h}}$ channel, where the combined statistical and systematic uncertainty is shown. The signal yields for the sum of all three neutral MSSM Higgs bosons, A+H+h, expected for M_A= 160 GeV and $\tan\beta$= 8 are given for comparison. Also given are the products of signal efficiency times acceptance for M_Φ= 160 GeV.

$e\tau_{had}$ channel				
	$\sqrt{s}$ = 7 TeVdata		$\sqrt{s}$ = 8 TeVdata	
	no–B–tag	B–tag	no–B–tag	B–tag
$gg \to \Phi + gg \to \Phi b$	129 ± 9	10 ± 1	476 ± 30	38 ± 3
$Z/\gamma^* \to \tau\tau$	11712 ± 1028	135 ± 13	29637 ± 2601	445 ± 41
$Z/\gamma^* \to \ell\ell$ ($\ell = e, \mu$)	1494 ± 252	9 ± 1	6541 ± 1080	86 ± 14
W + jets	1384 ± 138	30 ± 9	6150 ± 615	127 ± 38
QCD	3789 ± 379	76 ± 15	11028 ± 1103	192 ± 38
$t\bar{t}$	46 ± 6	20 ± 4	307 ± 44	96 ± 20
Single top + di–boson	53 ± 10	8 ± 2	230 ± 40	29 ± 5
Total background	18478 ± 1133	278 ± 22	53893 ± 3087	975 ± 72
Data	18785	274	54547	975

Signal Eff.

	B–tag	no–B–tag
gg$\to \Phi$	$1.16 \cdot 10^{-4}$	$1.04 \cdot 10^{-2}$
bb$\to \Phi$	$1.61 \cdot 10^{-3}$	$8.52 \cdot 10^{-3}$

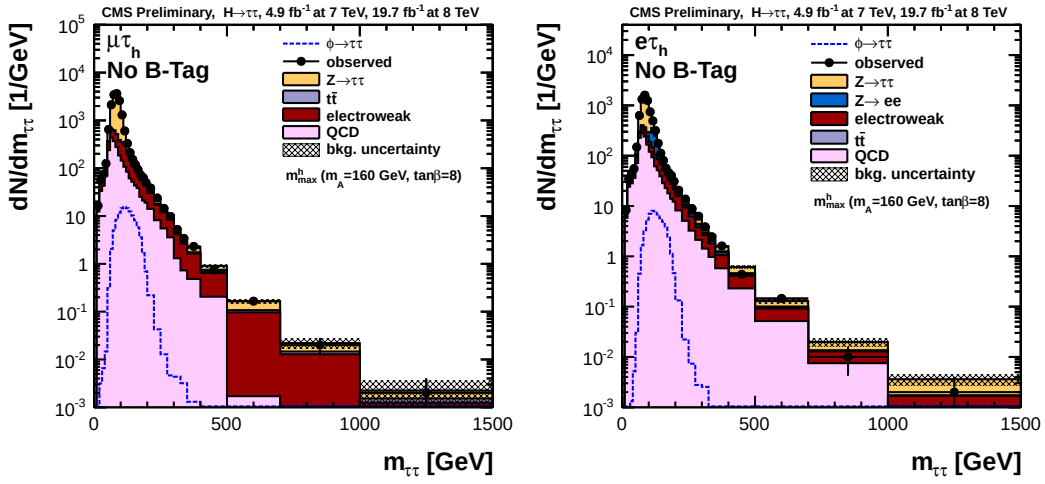


Figure 9.16: Reconstructed di-τ mass in the no-b-tag category for the $\mu\tau_h$ and eτ_h channels.

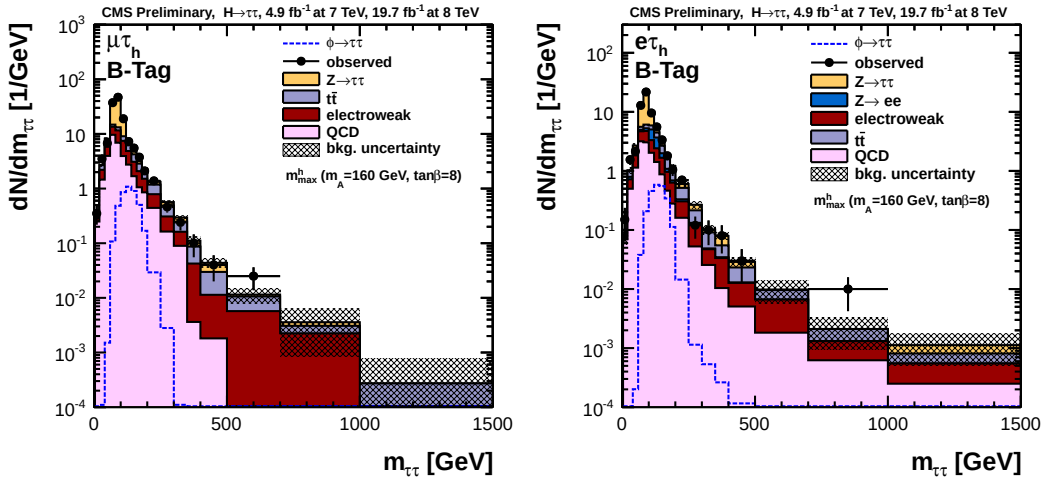


Figure 9.17: Reconstructed di-τ mass in the b-tag category for the $\mu\tau_h$ and eτ_h channels.

pseudoscalar Higgs boson mass M_A are set. The limits are computed using the modified Frequentist method [115]. Figure 9.18 shows the 95% CL exclusion in the tanβ-M_A parameter space for the MSSM m_h^{max} scenario. The exclusion limit set by the LEP experiments [79] is also shown. Numerical values for the expected and observed exclusion limits are given in Tab. 9.3. The expected limit has been computed for the case that no Higgs signal, neither of SM nor of MSSM type, is present in the data. The limit expected in case a SM Higgs boson is present in the data is computed separately and differs by 1-2 units in tanβ at low M_A. At high M_A there is also some effect as the limit is mainly driven by the light scalar Higgs h, which has the largest expected cross section.

Model independent limits on $\sigma\cdot$BR($\Phi \rightarrow \tau\tau$) for gluon-fusion and b-associated Higgs production as a function of the Higgs mass have been determined. The results for 8 TeVcenter-of-mass energy are shown in Fig. 9.19. In this case, a single resonance search (for a resonance of mass m_Φ) is performed. The results are also shown in Tables 9.4 and 9.5.

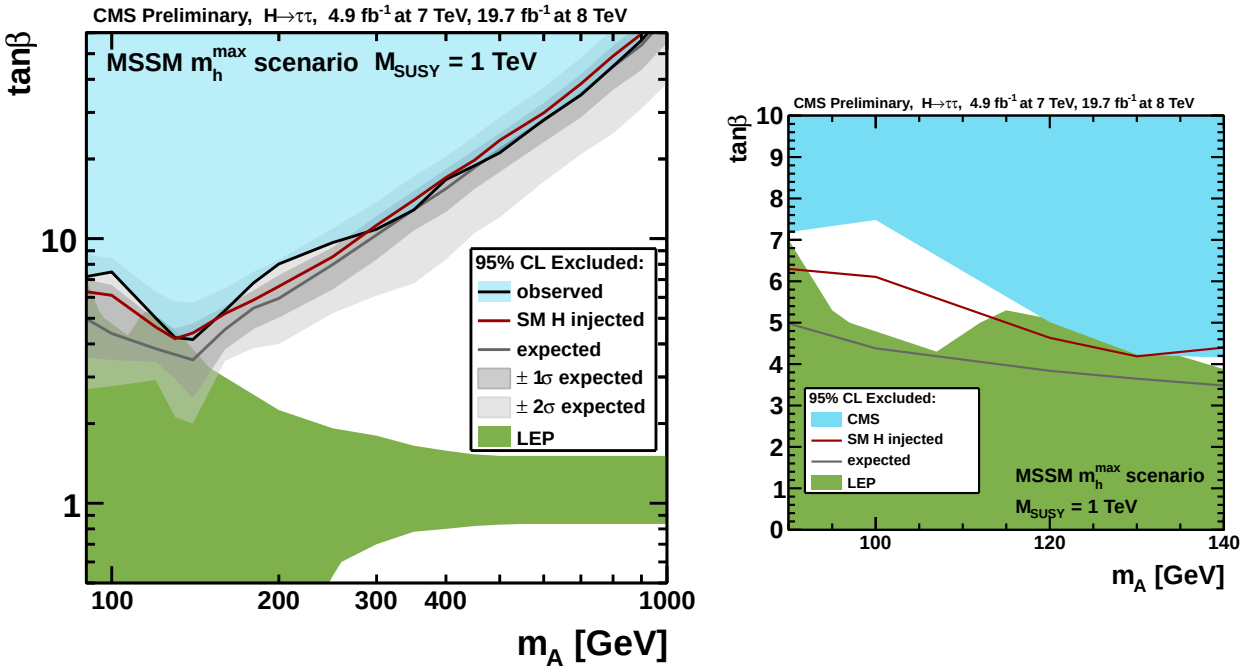


Figure 9.18: Left: Exclusion at 95% CL in the tanβ-M_A parameter space for the MSSM m_h^{max} scenario. The exclusion limits from the LEP experiments are also shown. Expected limits are computed for two cases: for the assumption that there is no Higgs $\rightarrow \tau\tau$ signal (neither MSSM nor SM) present in the data (dark grey line) and assuming that there is no MSSM, but a SM Higgs of mass 125–126 GeV present (red line). Right: 95% CL exclusion limit in the low M_A region.

Table 9.3: Expected range and observed 95% CL upper limits for $\tan\beta$ as a function of M_A, for the MSSM search.

MSSM Higgs	Expected $\tan\beta$ limit					Observed
m_{A} [GeV]	-2σ	-1σ	Median	$+1\sigma$	$+2\sigma$	$\tan\beta$ limit
90 GeV	2.70	3.56	4.98	7.02	8.71	7.19
100 GeV	2.77	3.48	4.38	6.68	8.44	7.48
120 GeV	2.92	3.42	3.84	4.94	6.30	5.01
130 GeV	2.12	2.95	3.65	4.57	5.79	4.23
140 GeV	2.00	2.50	3.48	4.78	5.74	4.16
160 GeV	3.45	3.81	4.54	5.56	6.49	5.35
180 GeV	3.86	4.55	5.46	6.36	7.49	6.80
200 GeV	3.99	5.03	5.95	7.28	8.32	8.02
250 GeV	5.22	6.43	7.98	9.27	10.9	9.66
300 GeV	6.09	8.34	10.3	12.2	13.8	10.9
350 GeV	6.77	10.7	12.9	15.1	17.2	12.8
400 GeV	8.33	12.6	15.4	18.3	20.4	16.7
450 GeV	10.5	15.4	18.5	21.6	24.4	18.8
500 GeV	12.0	17.9	21.7	25.0	28.8	21.1
600 GeV	16.4	23.1	27.9	32.6	37.2	28.1
700 GeV	20.8	28.7	35.2	41.8	48.2	34.9
800 GeV	25.0	36.5	44.7	53.6	> 60[1]	44.8
900 GeV	31.0	43.4	53.9	> 60[1]	> 60[1]	> 60[1]
1000 GeV	38.5	54.9	> 60[1]	> 60[1]	> 60[1]	> 60[1]

[1] We do not quote limits above $\tan\beta = 60$ as the theoretical calculations which provide the relation between cross section and $\tan\beta$ become unreliable at high $\tan\beta$.

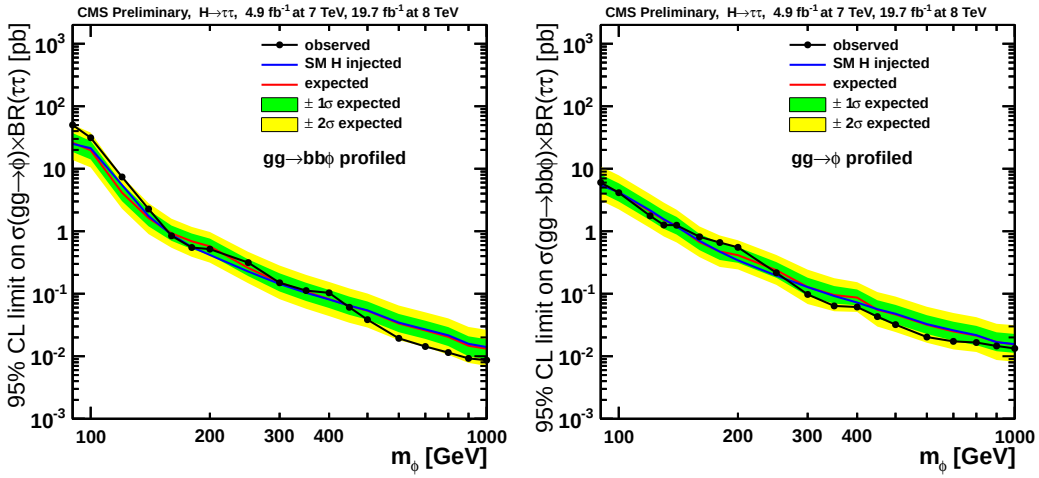


Figure 9.19: 95% CL upper limit on $\sigma \cdot \mathrm{BR}(\Phi \rightarrow \tau\tau)$ for gluon-fusion (left) and b-associated (right) production at 8 TeVcenter-of-mass energy as a function of M_Φ. Expected limits are computed for two cases: for the assumption that there is no Higgs $\rightarrow \tau\tau$ signal (neither MSSM nor SM) present in the data (red line) and assuming that there is no MSSM, but a SM Higgs of mass 125–126 GeV present (blue line).

Table 9.4: 95% CL upper limits for $\sigma\cdot$BR(ggΦ) (pb) as a function of M_Φ.

MSSM Higgs	Expected $\sigma\cdot$BR (ggΦ) limit					Observed
m_Φ [GeV]	-2σ	-1σ	Median	$+1\sigma$	$+2\sigma$	$\sigma\cdot$BR (ggΦ) limit
90 GeV	13.8	18.6	26.1	37.2	50.9	50.2
100 GeV	10.5	14.1	19.8	27.9	37.8	31.3
120 GeV	2.29	3.03	4.14	5.71	7.56	7.38
140 GeV	$8.99 \cdot 10^{-1}$	1.20	1.63	2.18	2.82	2.27
160 GeV	$5.44 \cdot 10^{-1}$	$7.02 \cdot 10^{-1}$	$9.26 \cdot 10^{-1}$	1.23	1.58	$8.45 \cdot 10^{-1}$
180 GeV	$3.90 \cdot 10^{-1}$	$5.19 \cdot 10^{-1}$	$6.91 \cdot 10^{-1}$	$9.19 \cdot 10^{-1}$	1.17	$5.49 \cdot 10^{-1}$
200 GeV	$3.14 \cdot 10^{-1}$	$4.17 \cdot 10^{-1}$	$5.73 \cdot 10^{-1}$	$7.62 \cdot 10^{-1}$	$9.73 \cdot 10^{-1}$	$5.17 \cdot 10^{-1}$
250 GeV	$1.47 \cdot 10^{-1}$	$1.91 \cdot 10^{-1}$	$2.59 \cdot 10^{-1}$	$3.54 \cdot 10^{-1}$	$4.67 \cdot 10^{-1}$	$3.15 \cdot 10^{-1}$
300 GeV	$8.18 \cdot 10^{-2}$	$1.09 \cdot 10^{-1}$	$1.51 \cdot 10^{-1}$	$2.12 \cdot 10^{-1}$	$2.83 \cdot 10^{-1}$	$1.50 \cdot 10^{-1}$
350 GeV	$5.74 \cdot 10^{-2}$	$7.65 \cdot 10^{-2}$	$1.07 \cdot 10^{-1}$	$1.50 \cdot 10^{-1}$	$2.00 \cdot 10^{-1}$	$1.12 \cdot 10^{-1}$
400 GeV	$4.39 \cdot 10^{-2}$	$5.91 \cdot 10^{-2}$	$8.20 \cdot 10^{-2}$	$1.15 \cdot 10^{-1}$	$1.54 \cdot 10^{-1}$	$1.03 \cdot 10^{-1}$
450 GeV	$3.43 \cdot 10^{-2}$	$4.59 \cdot 10^{-2}$	$6.41 \cdot 10^{-2}$	$8.99 \cdot 10^{-2}$	$1.21 \cdot 10^{-1}$	$6.07 \cdot 10^{-2}$
500 GeV	$2.88 \cdot 10^{-2}$	$3.84 \cdot 10^{-2}$	$5.34 \cdot 10^{-2}$	$7.52 \cdot 10^{-2}$	$1.01 \cdot 10^{-1}$	$3.85 \cdot 10^{-2}$
600 GeV	$1.80 \cdot 10^{-2}$	$2.42 \cdot 10^{-2}$	$3.34 \cdot 10^{-2}$	$4.75 \cdot 10^{-2}$	$6.48 \cdot 10^{-2}$	$1.93 \cdot 10^{-2}$
700 GeV	$1.42 \cdot 10^{-2}$	$1.89 \cdot 10^{-2}$	$2.59 \cdot 10^{-2}$	$3.70 \cdot 10^{-2}$	$5.04 \cdot 10^{-2}$	$1.43 \cdot 10^{-2}$
800 GeV	$1.05 \cdot 10^{-2}$	$1.46 \cdot 10^{-2}$	$2.04 \cdot 10^{-2}$	$2.92 \cdot 10^{-2}$	$4.03 \cdot 10^{-2}$	$1.15 \cdot 10^{-2}$
900 GeV	$7.82 \cdot 10^{-3}$	$9.98 \cdot 10^{-3}$	$1.47 \cdot 10^{-2}$	$2.08 \cdot 10^{-2}$	$2.94 \cdot 10^{-2}$	$9.23 \cdot 10^{-3}$
1000 GeV	$7.02 \cdot 10^{-3}$	$8.96 \cdot 10^{-3}$	$1.32 \cdot 10^{-2}$	$1.87 \cdot 10^{-2}$	$2.64 \cdot 10^{-2}$	$8.65 \cdot 10^{-3}$

Table 9.5: Expected range and observed 95% CL upper limits for σ·BR(bbΦ) (pb) at 8 TeVcenter-of-mass energy as a function of M_Φ.

MSSM Higgs	Expected σ·BR (bbΦ) limit					Observed
m_Φ [GeV]	-2σ	-1σ	Median	$+1\sigma$	$+2\sigma$	σ·BR (bbΦ) limit
90 GeV	3.11	4.15	5.79	8.07	10.9	6.03
100 GeV	2.24	2.99	4.17	5.85	7.85	4.14
120 GeV	1.13	1.50	2.09	2.93	3.93	1.76
140 GeV	$6.61 \cdot 10^{-1}$	$8.85 \cdot 10^{-1}$	1.22	1.70	2.21	1.25
160 GeV	$3.85 \cdot 10^{-1}$	$4.98 \cdot 10^{-1}$	$6.68 \cdot 10^{-1}$	$9.05 \cdot 10^{-1}$	1.19	$8.14 \cdot 10^{-1}$
180 GeV	$2.68 \cdot 10^{-1}$	$3.47 \cdot 10^{-1}$	$4.73 \cdot 10^{-1}$	$6.49 \cdot 10^{-1}$	$8.56 \cdot 10^{-1}$	$6.59 \cdot 10^{-1}$
200 GeV	$2.45 \cdot 10^{-1}$	$3.12 \cdot 10^{-1}$	$4.14 \cdot 10^{-1}$	$5.54 \cdot 10^{-1}$	$7.17 \cdot 10^{-1}$	$5.53 \cdot 10^{-1}$
250 GeV	$1.39 \cdot 10^{-1}$	$1.80 \cdot 10^{-1}$	$2.38 \cdot 10^{-1}$	$3.21 \cdot 10^{-1}$	$4.17 \cdot 10^{-1}$	$2.17 \cdot 10^{-1}$
300 GeV	$6.84 \cdot 10^{-2}$	$9.20 \cdot 10^{-2}$	$1.28 \cdot 10^{-1}$	$1.80 \cdot 10^{-1}$	$2.43 \cdot 10^{-1}$	$9.75 \cdot 10^{-2}$
350 GeV	$5.24 \cdot 10^{-2}$	$6.94 \cdot 10^{-2}$	$9.52 \cdot 10^{-2}$	$1.33 \cdot 10^{-1}$	$1.77 \cdot 10^{-1}$	$6.38 \cdot 10^{-2}$
400 GeV	$5.12 \cdot 10^{-2}$	$6.53 \cdot 10^{-2}$	$8.67 \cdot 10^{-2}$	$1.17 \cdot 10^{-1}$	$1.53 \cdot 10^{-1}$	$6.13 \cdot 10^{-2}$
450 GeV	$2.98 \cdot 10^{-2}$	$3.98 \cdot 10^{-2}$	$5.53 \cdot 10^{-2}$	$7.87 \cdot 10^{-2}$	$1.06 \cdot 10^{-1}$	$4.31 \cdot 10^{-2}$
500 GeV	$2.44 \cdot 10^{-2}$	$3.30 \cdot 10^{-2}$	$4.66 \cdot 10^{-2}$	$6.62 \cdot 10^{-2}$	$9.01 \cdot 10^{-2}$	$3.20 \cdot 10^{-2}$
600 GeV	$1.64 \cdot 10^{-2}$	$2.27 \cdot 10^{-2}$	$3.19 \cdot 10^{-2}$	$4.53 \cdot 10^{-2}$	$6.19 \cdot 10^{-2}$	$2.03 \cdot 10^{-2}$
700 GeV	$1.29 \cdot 10^{-2}$	$1.78 \cdot 10^{-2}$	$2.49 \cdot 10^{-2}$	$3.57 \cdot 10^{-2}$	$4.92 \cdot 10^{-2}$	$1.73 \cdot 10^{-2}$
800 GeV	$1.17 \cdot 10^{-2}$	$1.50 \cdot 10^{-2}$	$2.14 \cdot 10^{-2}$	$3.06 \cdot 10^{-2}$	$4.21 \cdot 10^{-2}$	$1.66 \cdot 10^{-2}$
900 GeV	$8.70 \cdot 10^{-3}$	$1.20 \cdot 10^{-2}$	$1.69 \cdot 10^{-2}$	$2.41 \cdot 10^{-2}$	$3.31 \cdot 10^{-2}$	$1.46 \cdot 10^{-2}$
1000 GeV	$8.17 \cdot 10^{-3}$	$1.13 \cdot 10^{-2}$	$1.54 \cdot 10^{-2}$	$2.27 \cdot 10^{-2}$	$3.14 \cdot 10^{-2}$	$1.33 \cdot 10^{-2}$

Chapter 10

Conclusions

In summary, this thesis has presented two results: a standard model cross section measurement which is the first of its kind at the LHC, the only ever performed using p-p collisions at 7 TeV corresponding to an integrated luminosity of 4.9 fb^{-1}, and a search for a neutral MSSM Higgs boson decaying to tau pairs using p-p collisions with a dataset corresponding to an integrated luminosity of 24.6 fb^{-1}, with 4.9 fb^{-1} at 7 TeV and 19.7 fb^{-1} at 8 TeV.

The W+$\text{b}\bar{\text{b}}$ events were selected in the $\text{W} \to \mu\nu$ decay mode with a muon of $\text{p}_\text{T} > 25\text{GeV}$ and $|\eta| < 2.1$, and two b jets with $\text{p}_\text{T} > 25\text{GeV}$ and $|\eta| < 2.4$. The data sample corresponds to an integrated luminosity of $5.0 fb^{-1}$. To extract the total number of $\text{W} + \text{b}\bar{\text{b}}$ events a maximum likelihood fit was performed using a fit The final number of $\text{W} + \text{b}\bar{\text{b}}$ events were extracted via a binned maximum likelihood fit. To constrain the most prominent backgrounds and reduce the final systematic uncertainty the fit is performed simultaneously on the p_T of the leading jet (J_1) in the signal region after all selection requirements have been applied, and on the $m_{\text{J}_3\text{J}_4}$ distribution obtained from the $t\bar{t}$ control region. The J_1 p_T is chosen as the final fit variable due to its discrimination power against top-related backgrounds. The

measured cross section $\sigma(pp \to \mathrm{W} + \mathrm{b\bar{b}}, p_T^{\mathrm{b}} > 25\ \mathrm{GeV}, |\eta^{\mathrm{b}}| < 2.4) \times \mathcal{B}(\mathrm{W} \to \mu\nu, p_T^{\mu} > 25\ \mathrm{GeV}, |\eta^{\mu}| < 2.1) = 0.53 \pm 0.05\,(\mathrm{stat.}) \pm 0.09\,(\mathrm{syst.}) \pm 0.06\,(\mathrm{theo.}) \pm 0.01\,(\mathrm{lum.})\ pb.$ for production of a W boson in association with two b jets is in agreement with the SM predictions. This result is approaching the precision of theoretical predictions at NNLO, allowing a sensitive test of perturbative calculations in the SM. This result compliments previous W+b and W+bb results which have shown varying levels of agreement with standard model predictions [50, 51, 52]. This measurement of $\mathrm{W} + \mathrm{b\bar{b}}$ also serves as an important benchmark at the LHC especially to searches which include a single, isolated lepton and one or more b jets in the final state, (for example, the search for a neutral MSSM higgs boson also presented in this thesis) as $\mathrm{W} + \mathrm{b\bar{b}}$ is an irreducible background.

In the search for neutral MSSM Higgs bosons decaying to tau pairs, five different $\tau\tau$ final states are studied: $\mathrm{e}\tau_h, \mu\tau_h, \mathrm{e}\mu,\ \mu\mu$ and $\tau_h\tau_h$. To enhance the sensitivity to neutral MSSM Higgs bosons from the minimal supersymmetric extension of the standard model (MSSM), events containing zero and events containing one b-tagged jet are analyzed in separate categories. No excess is observed in the tau-pair invariant-mass spectrum. Exclusion limits in the MSSM parameter space have been obtained for the m_h^{max} scenario[111, 112, 113]. This search extends previous results to larger values of M_A and excluded values of $\tan\beta$ as low as 4.2 at $M_A = 140$ GeV. In addition, model independent upper limits on the Higgs boson production cross section times branching fraction for gluon-gluon fusion and b-associated production are given.

10.1 Future Outlook

In 2015 the CMS experiment will have completed the necessary upgrades and maintenance during long shutdown 1 (LS1) and will begin taking data produced in p-p

collisions at 13TeV. The run in 2015 is planned to collide protons at a bunch crossing separation of 25 ns with a peak luminosity of $2{\times}10^{34}$ $\mathrm{cm}^{-2}\mathrm{s}^{-1}$. It is likely that CMS will collect over 100 fb^{-1} in 2015-2017. Thus far, the standard model of particle physics has been shown to perform well in describing experimental observations at energies around the electroweak scale of $O(246\mathrm{GeV})$. The recent discovery of a standard model-like higgs boson brought with it even more confidence in the standard model. However, questions arise when this model is seen as a part of a grand unified theory. What happens when probing regions between the electroweak and planck scales $O(1.22 \times 10^{19}\mathrm{GeV})$ where the standard model encounters renormalization issues? Supersymmetry has been proposed as a reliable method to cancel out anomalies that arise at these high energies. Therefore, exciting times lie ahead as the search for physics beyond the standard model continues.

Bibliography

[1] J.J. Thomson.

[2] David J Griffiths. Introduction to Quantum Mechanics; 2nd ed. 2005.

[3] Francis Halzen and Alan Martin. *Quarks & Leptons: An introductory course in modern particle physics*. John Wiley & Sons, New York, USA, 1984.

[4] A. Pais. Inward bound: of matter and forces in the physical world. 1988.

[5] D. P. et al. Barber. Discovery of three-jet events and a test of quantum chromodynamics at petra. *Phys. Rev. Lett.*, 43:830–833, Sep 1979.

[6] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys.Lett.*, B716:30–61, 2012.

[7] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys.Lett.*, B716:1–29, 2012.

[8] John M. Campbell and R.K. Ellis. Mcfm for the tevatron and the lhc. *Nucl. Phys. Proc. Suppl.*, 205:10, 2010.

[9] Simon Badger, John M. Campbell, and R.K. Ellis. Qcd corrections to the hadronic production of a heavy quark pair and a w-boson including decay correlations. *JHEP*, 03:027, 2011.

[10] Stephen P. Martin. A Supersymmetry primer. 1997.

[11] M. S. Carena et al. MSSM Higgs boson searches at the Tevatron and the LHC: Impact of different benchmark scenarios. *Eur. Phys. J. C*, 45:797, 2006.

[12] M. S. Carena et al. Suggestions for benchmark scenarios for MSSM Higgs boson searches at hadron colliders. *Eur. Phys. J. C*, 26:601, 2003.

[13] CMS Collaboration. Performance of tau-lepton reconstruction and identification in cms. *JINST*, 7:P01001, 2012.

[14] LHC Higgs Cross Section Working Group. Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables. CERN Report CERN-2011-002, 2011.

[15] M. Wilson. Superconducting magnets for accelerators.

[16] The CMS Collaboration. Precise mapping of the magnetic field in the cms barrel yoke using cosmic rays. *JINST*, 5(03):T03021, 2010.

[17] The CMS Collaboration. Description and performance of the cms track and primary vertex reconstruction. *JINST*. TRK-11-001.

[18] The CMS Collaboration. Cms physics technical design report, volume ii: Physics performance. *Journal of Physics G: Nuclear and Particle Physics*, 34(6):995, 2007.

[19] R. Fruhwirth. Application of kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth. A*, 262:444, 1987.

[20] K. Rose. Deterministic annealing for clustering, compression, classification, regression and related optimisation problems. In *Proceedings of the IEEE Vol. 86*, volume 86, 1998.

[21] W. Adam et. al., *Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC*, **CMS Note 2005/001**.

[22] S. Chatrchyan et al. Search for neutral MSSM Higgs bosons decaying to tau pairs in pp collisions at $\sqrt{s} = 7$ TeV. *Phys. Rev. Lett.*, 106:231801, 2011.

[23] CMS Collaboration. Particle flow event reconstruction in cms and performance for jets, taus, and met. Cms physics analysis summary, 2009. CMS PAS-MSSM/PLOTS/-001.

[24] CMS Collaboration. Commissioning of the particle-flow event reconstruction with leptons from j/psi and w decays at 7 tev. Cms physics analysis summary, 2010. CMS-PAS-PFT-10-003.

[25] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_{T} jet clustering algorithm. *JHEP*, 04:063, 2008.

[26] Gavin P. Salam and Grgory Soyez. A practical seedless infrared-safe cone jet algorithm. *Journal of High Energy Physics*, 2007(05):086, 2007.

[27] GavinP. Salam. Towards jetography. *The European Physical Journal C*, 67(3-4):637–686, 2010.

[28] M. H. Seymour S. Catani, Y. L. Dokshitzer and B. R. Webber. *Nucl. Phys. B*, 406:187, 1993.

[29] S. Moretti Y. L. Dokshitzer, G. D. Leder and B. R. Webber. *JHEP*, 9708(001), 1997.

[30] G. P. Salam M. Cacciari and G. Soyez. *JHEP*, 0804(005), 2008.

[31] CMS Collaboration. Jet Energy Corrections determination at $\sqrt{s} = 7$ TeV. *CMS PAS JME*, 10-010, 2010.

[32] The CMS collaboration. Determination of jet energy calibration and transverse momentum resolution in cms. *Journal of Instrumentation*, 6(11):P11002, 2011.

[33] Serguei Chatrchyan et al. Identification and Filtering of Uncharacteristic Noise in the CMS Hadron Calorimeter. *JINST*, 5:T03014, 2010.

[34] Christian Weiser. A combined secondary vertex based b-tagging algorithm in cms. (CMS-NOTE-2006-014), Jan 2006.

[35] Byron P. Roe, Hai-Jun Yang, Ji Zhu, Yong Liu, Ion Stancu, and Gordon McGregor. Boosted decision trees as an alternative to artificial neural networks for particle identification. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 543:577 – 584, 2005.

[36] Andy Buckley, Jonathan Butterworth, Stefan Gieseke, David Grellscheid, Stefan Hoche, et al. General-purpose event generators for LHC physics. *Phys.Rept.*, 504:145–233, 2011.

[37] Fabio Maltoni, Thomas McElmurry, Robert Putman, and Scott Willenbrock. Choosing the Factorization Scale in Perturbative QCD. 2007.

[38] E.Laenen K.Tollefson M. Dobbs, S.Frixione. Les houches guidebook to monte carlo generators for hadron collider physics. 2004.

[39] John M. Campbell, J.W. Huston, and W.J. Stirling. Hard Interactions of Quarks and Gluons: A Primer for LHC Physics. *Rept.Prog.Phys.*, 70:89, 2007.

[40] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J. C*, 63:189, 2009.

[41] Serguei Chatrchyan et al. Study of double parton scattering using W + 2-jet events in proton-proton collisions at $\sqrt{s} = 7$ TeV. 2013.

[42] Johan Alwall et al. MadGraph/MadEvent v4: the new web generation. *JHEP*, 09:028, 2007.

[43] Z. Wąs. TAUOLA the library for tau lepton decay, and KKMC/KORALB/KORALZ... status report. *Nucl. Phys. B, Proc. Suppl.*, 98:96, 2001.

[44] T. Sjöstrand, L. Lönnblad and S. Mrenna, "PYTHIA 6.2: Physics and manual", arXiv:hep-ph/0108264.

[45] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *JHEP*, 11:070, 2007.

[46] Carlo Oleari and Laura Reina. $W^{\pm}b\bar{b}$ production in POWHEG. *JHEP*, 08:061, 2011.

[47] J. Campbell and R.K. Ellis, "Monte Carlo for FeMtobarn processes", http://mcfm.fnal.gov/.

[48] S. Agostinelli , Nucl. Instr. and Methods A 506 (2003) 250-303.

[49] J. et al. Beringer. Review of particle physics. *Phys. Rev. D*, 86:010001, Jul 2012.

[50] First measurement of the b-jet cross section in events with a w boson in $\mathrm{p\overline{p}}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev. Lett.*, 104:131801, 2010.

[51] Measurement of the $\mathrm{p\bar{p}} \to w + b + x$ production cross section at $\sqrt{s} = 1.96$ tev. *Phys. Lett. B*, 718:044, 2013.

[52] Measurement of the cross section for the production of a w boson in association with b jets in pp collisions at $\sqrt{s} = 7$ tev with the atlas detector. *JHEP*, 1306:084, 2013.

[53] Measurement of the cross section and angular correlations for associated production of a Z boson with b hadrons in pp collisions at $\sqrt{s} = 7$ TeV. *Submitted to JHEP*, page , 2013.

[54] Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks. Submitted to *Phys. Rev. D*, 2013.

[55] Georges Aad et al. Measurement of the cross-section for W boson production in association with b-jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *JHEP*, 1306:084, 2013.

[56] Rikkert Frederix, Stefano Frixione, Valentin Hirschi, Fabio Maltoni, Roberto Pittau, et al. W and $Z/\gamma*$ boson production in association with a bottom-antibottom pair. *JHEP*, 09:061, 2011.

[57] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer. Madgraph 5: going beyond. *JHEP*, 06:128, 2011.

[58] Fabio Maltoni, Giovanni Ridolfi, and Maria Ubiali. b-initiated processes at the LHC: a reappraisal. *JHEP*, 1207:022, 2012.

[59] Torbjorn Sjöstrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 physics and manual. *JHEP*, 05:026, 2006.

[60] S. Alioli, P. Nason, C. Oleari, and E. Re. NLO vector-boson production matched with shower in POWHEG. *JHEP*, 07:060, 2008.

[61] Paolo Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004.

[62] Hung-Liang Lai, J. Huston, Z. Li, P. Nadolsky, J. Pumplin, D. Stump, and C.-P. Yuan. Uncertainty induced by QCD coupling in the CTEQ global analysis of parton distributions. *Phys. Rev. D*, 82:054021, 2010.

[63] S. Agostinelli et al. Geant4: a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003.

[64] R. Field. Early lhc underlying event data - findings and surprises. page , 2010.

[65] Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s} =$ 7 TeV. *JHEP*, 01:080, 2011.

[66] Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Phys. Lett. B*, 659:119–126, 2008.

[67] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Catchment Area of Jets. *JHEP*, 04:005, 2008.

[68] Serguei Chatrchyan et al. Determination of jet energy calibration and transverse momentum resolution in CMS. *JINST*, 6:P11002, 2011.

[69] Identification of b-quark jets with the CMS experiment. *JINST*, 8:P04013, 2013.

[70] CMS Collaboration. Absolute Calibration of the Luminosity Measurement at CMS: Winter 2012 Update. *CMS Physics Analysis Summary*, 2012.

[71] Iain W. Stewart and Frank J. Tackmann. Theory uncertainties for Higgs and other searches using jet bins. *Phys. Rev. D*, 85:034011, 2012.

[72] Study of observables sensitive to double parton scattering in w + 2 jets process in p-p collisions at $\sqrt{s} = 7$ tev. *CMS Physics Analysis Summary*,

http://cds.cern.ch/record/1542311CMS-PAS-FSQ-12-028, 2013. To be submitted to JHEP.

[73] Measurement of hard double-parton interactions in $w(\rightarrow l\nu)+$ 2 jet events at $\sqrt{s} = 7$ tev with the atlas detector. *New J. Phys.*, 15:033038, 2013.

[74] Higgs to tau tau (MSSM). Technical Report CMS-PAS-HIG-13-021, CERN, Geneva, 2013.

[75] ATLAS Collaboration. Search for neutral MSSM Higgs bosons decaying to $\tau^+\tau^-$ pairs in proton-proton collisions at $\sqrt{s} = 7$ Tev with the ATLAS detector. *Phys. Lett. B*, 705:174, 2011.

[76] CDF and D0 Collaborations. Combined CDF and D0 upper limits on MSSM Higgs boson production in tau-tau final states with up to 2.2 fb^{-1}. 2010.

[77] V. M. Abazov. Search for Higgs bosons decaying to $\tau^+\tau^-$ pairs in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Lett. B*, 707:323, 2012.

[78] T. Aaltonen et al. Search for Higgs bosons predicted in two-Higgs-doublet models via decays to tau lepton pairs in 1.96 TeV $p\bar{p}$ collisions. *Phys. Rev. Lett.*, 103:201801, 2009.

[79] S. Schael et al. Search for neutral MSSM Higgs bosons at LEP. *Eur. Phys. J. C*, 47:547, 2006.

[80] CMS Collaboration. Search for the standard–model higgs boson decaying to tau pairs in proton–proton collisions at $\sqrt{s} = 7$ and 8 TeV. *CMS Paper in preparation*, CMS-PAPER-HIG-13-004, 2013.

[81] S. Heinemeyer, O. Stal, and G. Weiglein. Interpreting the LHC Higgs Search Results in the MSSM. *Phys.Lett.*, B710:201–206, 2012.

[82] M. Carena, S. Heinemeyer, O. Stl, C.E.M. Wagner, and G. Weiglein. MSSM Higgs Boson Searches at the LHC: Benchmark Scenarios after the Discovery of a Higgs-like Particle. 2013.

[83] CMS Collaboration. Electron reconstruction and identification at $\sqrt{s} = 7$ TeV. CMS Physics Analysis Summary CMS-PAS-EGM-10-004, 2010.

[84] S. Chatrchyan et al. Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$TeV. *JINST*, 7:P10002, 2012.

[85] S. Chatrchyan et al. Measurement of inclusive Z cross section via decays to tau pairs in pp collisions at $\sqrt{s}$=7 TeV. *JHEP*, 8:117, 2011.

[86] CMS Collaboration. Particle–flow event reconstruction in CMS and performance for jets, taus, and $E_{\mathrm{T}}^{\mathrm{miss}}$. CMS Physics Analysis Summary CMS-PAS-PFT-09-001, 2009.

[87] CMS Collaboration. Commissioning of the particle-flow reconstruction in minimum-bias and jet events from pp collisions at 7 TeV. CMS Physics Analysis Summary CMS-PAS-PFT-10-002, 2010.

[88] CMS Collaboration. Commissioning of the particle-flow event reconstruction with leptons from J/ψ and W decays at 7 TeV. CMS Physics Analysis Summary CMS-PAS-PFT-10-003, 2010.

[89] M. Cacciari, G. P. Salam, and G. Soyez. FastJet user manual.

[90] M. Cacciari and G. P. Salam. Dispelling the N^3 myth for the k_t jet-finder. *Phys. Lett. B*, 641:57, 2006.

[91] CMS Collaboration. b-jet identification in the cms experiment. CMS Physics Analysis Summary CMS-PAS-BTV-11-004, 2012.

[92] EGamma POG. https://twiki.cern.ch/twiki/bin/view/CMS/EgammaCutBasedIdentification

[93] Muon POG. https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonId#Tight_Mu

[94] CMS Collaboration. Performance of missing transverse momentum reconstruction algorithms in proton–proton collisions at $\sqrt{s} = 8$ tev with the cms detector. CMS Physics Analysis Summary CMS-PAS-JME-12-002, 2012.

[95] M. Cacciari and G. P. Salam. Pileup subtraction using jet areas. *Phys. Lett. B*, 659:119, 2008.

[96] V. Khachatryan et al. Measurement of inclusive W and Z cross sections in pp collisions $\sqrt{s}$=7 TeV. *JHEP*, 1110:132, 2011.

[97] CMS Collaboration. Measurement of CMS luminosity. CMS Physics Analysis Summary CMS-PAS-EWK-10-004, 2010.

[98] Determination of the jet energy scale in cms with pp collisions at $\sqrt{s}$ = 7 TeV. *CMS PAS*, http://cms-physics.web.cern.ch/cms-physics/public/JME-10-010-pas.pdfJME-10-010, 2010.

[99] Tau POG. https://twiki.cern.ch/twiki/bin/viewauth/CMS/TauIDRecommendation.

[100] R. V. Harlander and W. B. Kilgore. Next-to-next-to-leading order Higgs production at hadron colliders. *Phys. Rev. Lett.*, 88:201801, 2002.

[101] C. Anastasiou and M. Charalampos. Higgs boson production at hadron colliders in NNLO QCD. *Nucl. Phys. B*, 646:220, 2002.

[102] V. Ravindran, J. Smith, and W. L. van Neerven. NNLO corrections to the total cross section for Higgs boson production in hadron hadron collisions. *Nucl. Phys. B*, 665:325, 2003.

[103] R. V. Harlander and W. B. Kilgore. Production of a pseudo-scalar Higgs boson at hadron colliders at next-to-next-to leading order. *JHEP*, 10:017, 2002.

[104] Charalampos Anastasiou and Kirill Melnikov. Pseudoscalar Higgs boson production at hadron colliders in NNLO QCD. *Phys. Rev. D*, 67:037501, 2003.

[105] M. Spira, A. Djouadi, D. Graudenz, and P. M. Zerwas. Higgs boson production at the LHC. *Nucl. Phys. B*, 453:17, 1995.

[106] M. Spira. HIGLU: A Program for the Calculation of the Total Higgs Production Cross Section at Hadron Colliders via Gluon Fusion including QCD Corrections. 1995.

[107] S. Dittmaier, M. Kramer, and M. Spira. Higgs radiation off bottom quarks at the Tevatron and the LHC. *Phys. Rev. D*, 70:074010, 2004.

[108] S. Dawson, C.B. Jackson, L. Reina, and D. Wackeroth. Exclusive Higgs boson production with bottom quarks at hadron colliders. *Phys. Rev. D*, 69:074027, 2004.

[109] R. V. Harlander and W. B. Kilgore. Higgs boson production in bottom quark fusion at next-to-next-to-leading order. *Phys. Rev. D*, 68:013001, 2003.

[110] R. Harlander, M. Kramer, and M. Schumacher. Bottom-quark associated Higgs-boson production: reconciling the four- and five-flavour scheme approach. CERN-PH-TH/2011-134, FR-PHENO-2011-009, TTK-11-17, WUB/11-04, 2011.

[111] S. Heinemeyer, W. Hollik, and G. Weiglein. FeynHiggs: a program for the calculation of the masses of the neutral CP-even Higgs bosons in the MSSM. *Comput. Phys. Commun.*, 124:76–89, 2000.

[112] S. Heinemeyer, W. Hollik, and G. Weiglein. The Masses of the neutral CP-even Higgs bosons in the MSSM: Accurate analysis at the two loop level. *Eur. Phys. J. C*, 9:343, 1999.

[113] G. Degrassi et al. Towards high-precision predictions for the MSSM Higgs sector. *Eur. Phys. J. C*, 28:133, 2003.

[114] Alexander L. Read. Linear interpolation of histograms. *Nucl.Instrum.Meth.*, A425:357–360, 1999.

[115] A. L. Read. Modified frequentist analysis of search results (the cls method). CERN Report CERN-OPEN-2000-005, 2000.