A MEASUREMENT OF TOP QUARK PAIR AND PHOTON PRODUCTION CROSS SECTION WITH CMS DETECTOR

by

MIKHAIL MAKOUSKI

M.S., Moscow Institute for Physics and Technologies, 2007

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Physics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2015

Abstract

In this thesis the measurement of production cross section of top-quark pairs in association with a photon in proton-proton collisions at a center of mass energy of 8 TeV is presented. The data was recorded at the CMS experiment at the LHC in 2012. This measurement aims to extend our knowledge of top quark properties and help to test consistency of the Standard Model (SM) of particle physics. Data-driven methods are used to estimate the photon identification efficiency and purity. The measured cross-section agrees with the standard model expectation.

A MEASUREMENT OF TOP QUARK PAIR AND PHOTON PRODUCTION CROSS SECTION WITH CMS DETECTOR

by

MIKHAIL MAKOUSKI

M.S., Moscow Institute for Physics and Technologies, 2007

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Physics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2015

Approved by:

Major Professor
Andrew Ivanov

Copyright

Abstract

In this thesis the measurement of production cross section of top-quark pairs in association with a photon in proton-proton collisions at a center of mass energy of 8 TeV is presented. The data was recorded at the CMS experiment at the LHC in 2012. This measurement aims to extend our knowledge of top quark properties and help to test consistency of the Standard Model (SM) of particle physics. Data-driven methods are used to estimate the photon identification efficiency and purity. The measured cross-section agrees with the standard model expectation.

Table of Contents

List of Figures

List of Tables

Preface

Particle physics is a fascinating subject. It is a world of resonances, flavors, broken symmetries, and forbidden decays that still happen. It is very easy to get lost in the jargon if you are new to it. Here I will outline the basic principles and ideas, and show the motivation and meaning of the work presented in this thesis.

Particle physics (also known as High Energy Physics or HEP) deals with objects small enough that one needs Quantum Mechanics (QM) for the adequate description. But these objects are often moving too fast for this classical theory to work. Description of fast (but not small) moving objects is a job of Special Relativity, but it is known that these two very good theories do not work well together. Moreover, when high energy particles interact, sometimes different particles of new types are created, and this is not described by either theory.

A theory that allows to describe behavior and transformations of small objects was developed in the period from 1950s to 1970s. QM[6,7] is the foundation for Quantum Field Theory (QFT)[8,9,10]. However, it has a very different mathematical structure. QM deals with quantum states represented as wave functions. Physical observables become operators acting on the states (transforming functions into functions). Wave functions have a probabilistic interpretation. QFT makes another big change in the way the world is viewed. A field becomes the main building block, dynamical variable of the theory. Particles are understood as quantum excitations of the field. That is, to have a particle one has to act with a creation operator on the field. Interactions between different particles are described through adding terms to the Lagrangian, which governs the dynamics of the fields. The quantity that can be calculated in QFT is the amplitude of one state being transformed into other. Quantum fields have no obvious connection with the observable world. So essentially QFT defines a set of rules, that help to calculate (using perturbation theory) amplitude of a state to

transform into another state given the fields and Lagrangian. This is useful numerical output of the theory. The question is how do we know what fields are there and what kind of Lagrangian do we use in certain situations? The most important concept – symmetry – helps here. But essentially the process is based on a very educated guessing. This is where input from experimental observations comes into play. The task is to build a model that describes observed results and makes reliable non-trivial predictions. To do this we need enough assorted data from experiments that would allow us to guess the Lagrangian and the properties of the fields involved. When enough parameters were introduced the theory may be in complete agreement with all observable data. Most importantly, even though the number of parameters of theory may be much smaller than amount of input data, the numerical results and predictions may be of extremely high accuracy. The Standard Model (SM) is the most complete framework that includes electromagnetic, weak and strong interactions between the fields that correspond to all known elementary particles. It works remarkably well. For a long time it was ahead of the experimental reach predicting the top quark and scalar boson (Higgs boson) that were necessary parts of its inner workings but were not observed by experiments.

The top quark was discovered in 1995 at Tevatron[11,12], and the Higgs boson in 2013 at Large Hadron Collider[13,14]. It is believed that SM is not the final theory that would explain all physical phenomena. Gravitational interaction could not be added to it. Cosmological observations suggest that Dark Matter can not be explained by any existing SM particle. Neutrinos are considered massless in the SM, however there is strong experimental evidence that it is not the case[15].

How one would improve the model that has demonstrated huge success and is built on a long trusted mathematical foundation? Following the way it was settled, one has to find either an entirely different model that will describe all the existing data (unlikely) or find a way to introduce new fields and terms to the Lagrangian (need some crucial experimental data). Both the top quark and Higgs boson seem to be natural "suspects" to look for new

properties at LHC, simply because they are the least studied parts of the SM. Top quarks are produced at much higher rates than in Tevatron experiments because of higher collision energy and higher collision rate. They are the heaviest elementary particles known today. And since the Higgs boson is responsible for mass generation in the SM there may be some interesting connection between the two. This thesis is based on the experimental study of the top quark. The top quark decays very shortly after it is produced and this makes it a challenging subject. Its mass was measured remarkably well[16]. Production cross section is another very important quantity predicted by SM and measured experimentally[17,18]. This study focuses on electromagnetic coupling of the top quark by measuring the production cross section of top anti-top pair and a photon. The result is then compared with SM prediction.

Chapter 1

Standard Model of Particle Physics

In this chapter the main structural elements of the Standard Model of particle physics will be outlined. A complete description and technical details can be found in[19].

Elementary particles can be divided into different categories based on their properties: spin, mass, charge and other quantum numbers that determine their behavior. It is usually convenient to start classification from spin. There are two groups of particles: with half integer spin (fermions) and integer spin (bosons).

There are 12 particles with spin 1/2 that are grouped into three generations, similar to chemical elements in the periodic table. Figure 1.1 shows all the particles of the SM in one picture. Each generation consists of 2 quarks and 2 leptons. The differences between quarks and leptons are the following. Quarks have color charge (which allows them to take part in strong interaction) and electric charge (in units of absolute value of electron charge) +2/3 for up-type quarks and -1/3 for down-type quarks. Leptons have no color charge and integer electric charge: 0 for neutrinos and 1 for charged leptons (electron, muon, tau-lepton for the three generations). Masses of particles generally go up with number of generation. For every particle there exists an antiparticle with same mass, spin and opposite electric and color charge. All visible matter around us consist of quarks of the first generation (up and down) and electrons. Protons and neutrons are formed by 3 quarks (uud and udd) and with

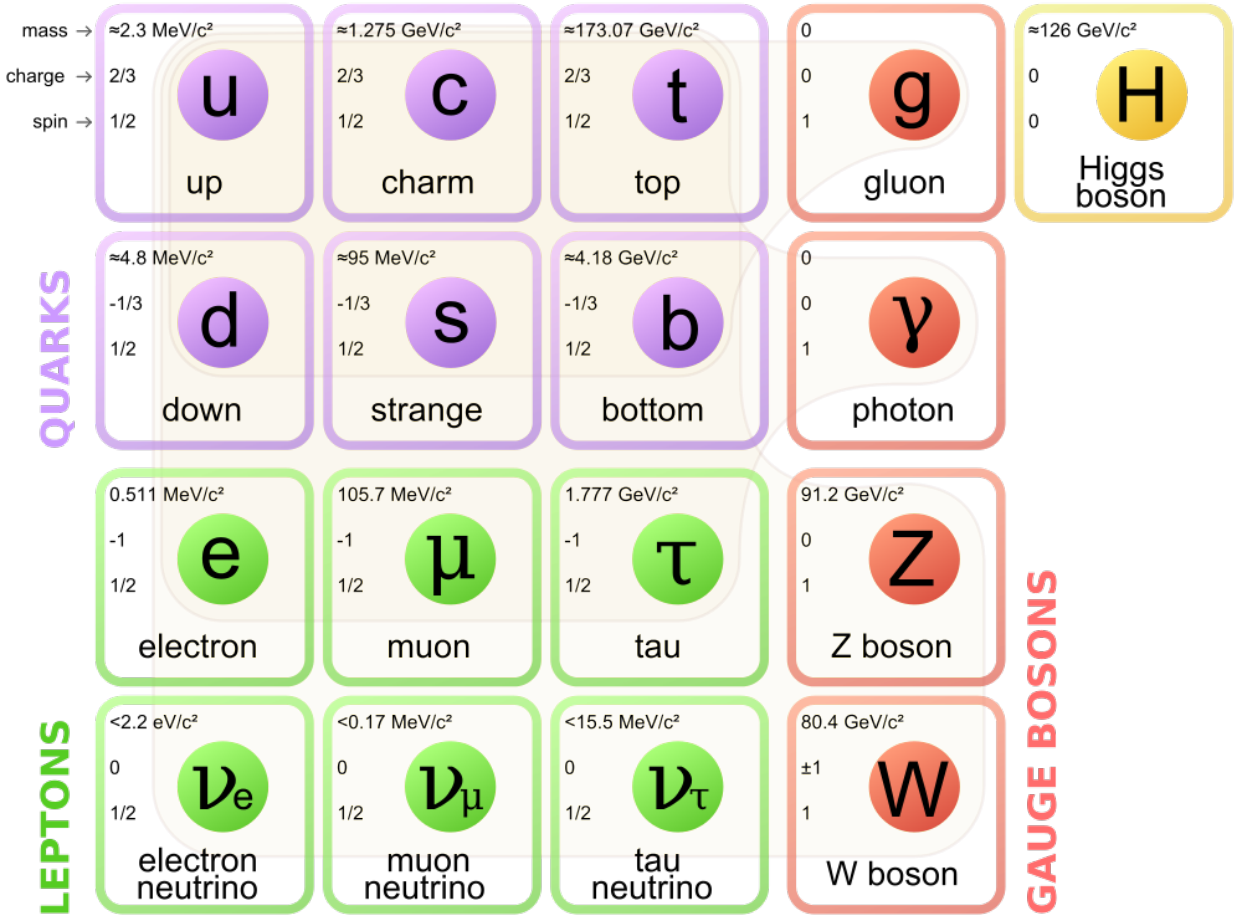

Figure 1.1: *Particles of the standard model, from*[1]

electron they make atoms. Electron neutrino takes part in beta-decay. Particles with spin 1 are photon, Z, W+, W- and gluons (8 different kinds). Photons, Z bosons and gluons are electrically neutral, W bosons have electric charge equal to charge of the electron. They all do not have antiparticles in a usual sense, unlike quarks and leptons, rather they are considered to be their own antiparticles. These particles are called mediators of interactions because quarks and leptons in the Standard Model are interacting between themselves only by exchange of bosons. Different kinds of interactions are described below.

The electromagnetic interaction happens between particles with non-zero electric charge and it is mediated by the exchange of photon. The photon is massless and has no electric charge.

All fermions: quarks and leptons are subject to the weak interaction. It is mediated by the exchange of electrically neutral Z bosons and charged W+ and W- bosons. Z and W can also couple to themselves. Unlike the photon, Z and W are massive ($m_Z \approx 91$ GeV, $m_W \approx 80$ GeV). This fact makes the weak interaction weak at low energies. It is less probable to create (emit) a massive particle if the total energy of the system is less than its mass. Another very peculiar property of weak interaction is the dependence on chirality of the interacting particles. That is, particles that have their spin projection along their momentum will interact differently then particles with spin projection opposite to their momentum. The W boson couples only to left-chiral particles and right-chiral antiparticles. The Z boson couples with different strength to right-chiral and left-chiral particles. This property violates parity: the processes happening as they are and as if they were mirror reflected would have different weak forces acting. Electromagnetic and weak interactions are unified into electroweak interaction in the SM.

The strong interaction occurs between quarks – particles that carry color charge. There are several aspects of strong interaction that make it different to electroweak interaction. Gluons are massless, they carry color charge (unlike photons). This property makes it possible for a gluon to emit another gluon. The coupling constant of the strong interaction is much higher than the weak coupling constant at low energies. At short distances quarks and gluons behave like free particles. As the distance between quarks increases, potential energy increases until the creation of quark-antiquark pair becomes possible. This property makes color charged states experimentally unobservable. Color neutral states can have a quark-antiquark (color – anti-color) structure and are called mesons, or three-quarks (3 colors canceling when added together) and called baryons.

Mathematically the Standard Model is based on gauge quantum field theory[10,9,8]. The main focus is on the symmetry of Lagrangian under continuous group of local transformations. From quantum mechanics we know that the Hamiltonian is the operator that defines time evolution of the system and Dirac equation closely resembles the Schrodinger equation

of the free particle. However, it is more convenient to start with the minimum action principle and consider Lagrangian as a primary point. The symmetry of the Lagrangian reflects the structure of the interactions described by it.

One of the most non-trivial parts of the SM is the Higgs mechanism of electroweak symmetry breaking[20,21]. It allows to assignment of masses to the W and Z bosons without breaking the mathematical structure of the theory.

Despite the overall success of the Standard Model there are some indications that it is not the final "theory of everything". There is no explanation of why three generations of almost identical fermions exist with the masses being the only difference. One can expect to have some mechanism responsible for this, similar to explanations of groups of chemical elements or mesons and baryons. There is no suitable dark matter candidate in the SM however it is known that dark matter is prevailing in the overall matter content of the Universe[22].

There are theories proposed to extend the SM and solve its problems, however they have not had experimental confirmation so far. The most notable are Supersymmetry (SUSY)[23] and extra dimensions[24]. SUSY is a generalization of symmetries of quantum field theory, models involving extra dimensions postulate existence of hidden dimensions of space accessible by gravity but not by other forces.

The goal of this work is to make a cross section measurement of a rare SM process. A deviation from the SM theoretical prediction could be an evidence of physics beyond the Standard Model.

Chapter 2

Top Quark

The top quark is the heaviest of known elementary particles and the most interesting one in the context of searching for new particles beyond the standard model. Due to its high mass it may play an important role in the electroweak symmetry breaking.

Theoretical motivation for the third generation of quarks was present since 1973[25]. The existence of the lightest quark of the third generation (bottom quark) was confirmed by observing Upsilon resonance in 1977[26], which is interpreted as a bound state of the bottom quark-antiquark pair. The existence of the second quark of the third generation (top quark) took 20 years to be experimentally confirmed. Its mass was unknown (even today there is no model explaining the wide range of all quark masses) and different experiments tried to find it using existing particle accelerators. Only in 1995 Tevatron experiments CDF and D0[11,12] observed direct production of top anti-top quark pairs. It is the only quark that does not form bound states (none has been observed so far).

LHC is the only machine capable of producing top quarks right now. Because of its higher center of mass energy compared to the Tevatron, top quarks are produced in abundance and its properties has been studied in more details than before[27].

Top quark production in proton-proton collisions and its decay modes will be briefly described in the following two sections.

2.1 Production mechanism

Top quarks can be produced in pairs via the strong interaction. Production of a single top quark is possible via Wtb vertex, but since it involves creation of a W boson and b quark, the probability of this process is smaller than pair production.

Pair production in the leading order processes can happen in two different ways: quark-antiquark annihilation and gluon-gluon fusion. The former was the dominant mode for the Tevatron collider since protons were collided with antiprotons having both quarks and antiquarks. At the LHC in proton-proton collisions there are no valence antiquarks, and the gluon fusion mechanism is dominant. There are ways to produce top quark pair from quark and gluon, but they include extra interactions and production of additional quarks or gluons.

One of the most valuable parameters that can be theoretically calculated and measured experimentally is inclusive top pair production cross section $\sigma_{t\bar{t}}$, that is proportional to probability of creation of a pair of top quarks by colliding two protons at given energy. In proton-proton collision calculation is complicated by the fact that protons are composite objects. The problem is solved by parametrization of probability of given parton (indivisible component of a proton) to carry certain fraction of proton's momentum. These probability distributions are called Parton Distribution Functions (PDF). They carry important information about the structure of a proton, and were deduced mainly by analyzing proton-electron and proton-positron collisions.

2.2 Decay modes

Top quark decays by weak interaction almost exclusively into W boson and bottom quark. Different top pair final states are characterized by the decay of two W bosons. Each W boson can decay into lepton and neutrino (leptonic decay) or into quarks (hadronic decay). There are three possible leptonic final states and two different hadronic final states possible

(if we neglect flavor mixing). However since quarks have color there are more options for the unique final states in hadronic W decay. According to[28] the W boson decays to quarks (hadrons) with probability 0.676. This can help get a feeling of different top pair decay modes:

- All hadronic (both W decay into quarks) is $0.676 \cdot 0.676 = 0.457$, a little less than half of the time
- Semileptonic (one W decays into quarks, another into leptons) $2 \cdot 0.676 \cdot (1 - 0.676) = 0.438$, this includes all lepton generations, but we are primarily focused on one of them (electron and electron neutrino)
- Dileptonic (both W decay into leptons) $(1 - 0.676) \cdot (1 - 0.676) = 0.105$, this includes all permutations of 3 lepton generations

Decay modes of a pair of top quarks can be graphically represented by Figure 2.1.

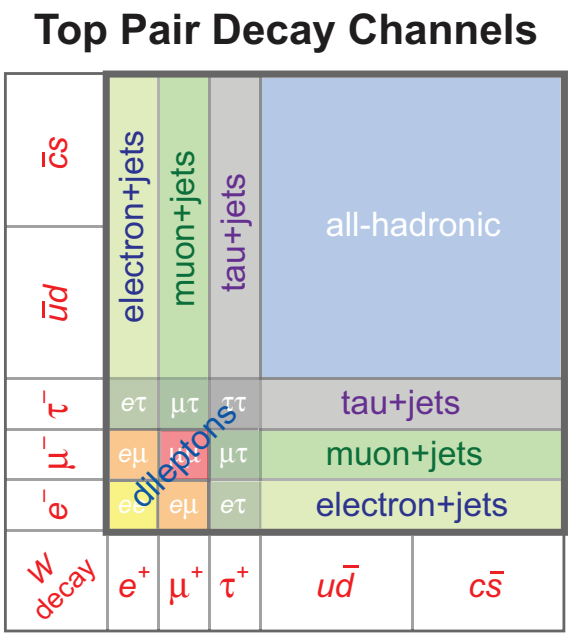

Figure 2.1: *Top pair decay channels, from*[2]

2.3 Top pair and photon

Our long term goal is to study electromagnetic interactions of the top quark. It is equivalent to finding a coupling constant of the vertex with top quark radiating photon (top quark charge) and an anomalous magnetic moment. It is experimentally challenging to select events with hard photons radiated directly by a top quark, as it quickly decays into a W boson and a b quark subsequently producing several jets and leptons. There are methods to separate decay products of each top quark from other objects in the event, an example is the kinematic fitter[29]. The idea of this method is to use a hypothesis of the event topology (top quark pair decay) and by using kinematic constraints (W invariant mass, top quark mass) detector resolution can be improved by varying the reconstructed particle's momenta to better match the known masses. However, presence of a hard photon that may be radiated by initial state particles, top quarks or final state particles, makes the full kinematic reconstruction much more difficult. In this analysis we focus on events containing a pair of top quarks and a photon. With this selection we indirectly assess electromagnetic interaction of top quark by performing inclusive $t\bar{t}+\gamma$ cross section measurement[30,31,32,33].

Top quark pair associated with photon production has some additional peculiarities[30,34]. The photon can be emitted by one of the initial state particles (radiative $t\bar{t}$ production) or top quark and its decay products (radiative $t\bar{t}$ decay). This analysis does not have sensitivity to disentangle these production modes.

In this work we concentrate on the semileptonic top pair decay with an electron in the final state, however, other decay modes can also pass the event selection. The choice of an electron over a muon creates some complications, but resolves other. The main factors are: lepton identification and reconstruction efficiency (which is reasonably higher for muons) and photon purity and misidentification rate. A good description of the latter is one of the most challenging parts of this analysis. By choosing the photon+electron+jets final state one can use a data-driven method to estimate the number of elections misidentified as photons, more details are given in Chapter6.5.

Chapter 3

Experimental Setup

3.1 Large Hadron Collider

The Large Hadron Collider (LHC) is an accelerator located at the European Center for Nuclear Research (CERN) near Geneva, Switzerland. It is designed and built to collide proton beams head on to discover new fundamental particles. The first proton proton collisions happened in 2008. The general layout of CERN accelerator complex is shown on the Figure 3.1 It has a sequential structure. In the first step the hydrogen ions (protons) are injected in the linear accelerator (LINAC), then the beam goes through several synchrotron rings until it reaches the LHC – the biggest and most powerful accelerator currently built.

There are several experimental setups located around the LHC ring. Two of them are made for particular subsets of physics measurements: ALICE (A Large Ion Collider Experiment)[35] and LHCb (Large Hadron Collider beauty experiment)[36]. Two more are called general purpose experiments – to cover a variety of topics: ATLAS (A Toroidal LHC ApparatuS)[37] and CMS (Compact Muon Solenoid)[38].

The data used in this study was collected in 2012 when the accelerator operated at the energy of 4TeV per beam, making energy in the center of mass to be 8TeV.

From the point of view of experiment a single most important quantity describing the

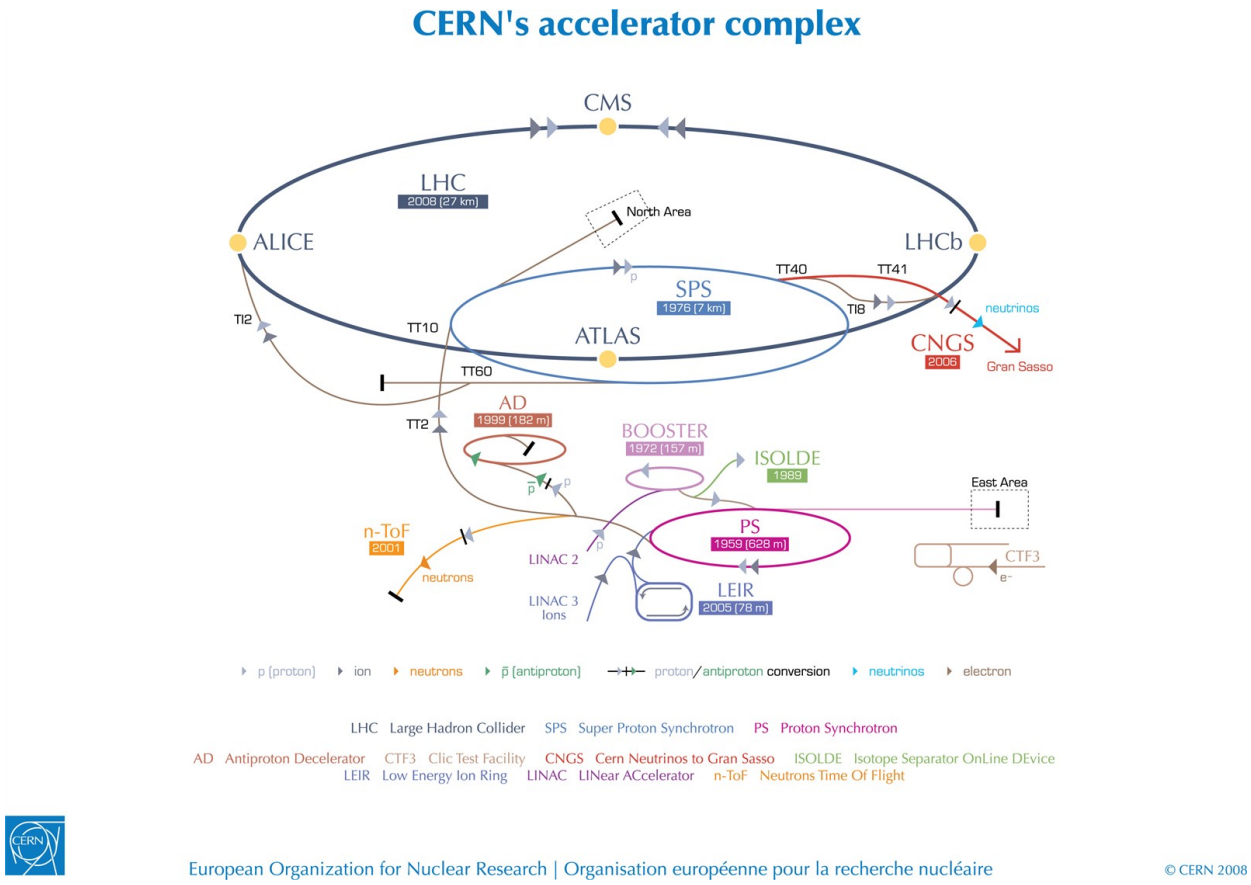

Figure 3.1: *CERN accelerator complex, from*[3]

accelerator performance (other then collision energy) is instantaneous luminosity $\mathcal{L}$. It is proportional to the number of collisions happening in the interaction point per unit of time. It depends on the number of circulating bunches, the number of protons per bunch and the way two beams are focused. There are many other important quantities related to the accelerator operation. Instantaneous luminosity is measured in $cm^{-2}s^{-1}$ (inverse area per unit time). The total amount of data collected is called integrated luminosity (integral of instantaneous luminosity over time) and is usually measured in inverse femtobarns (fb^{-1}).

3.2 Compact Muon Solenoid detector

The Compact Muon Solenoid (CMS) is one of the four major experimental setups to record the results of the proton collisions at the LHC.

It is important to understand the design goals that determined the structure and properties of the CMS detector[38]. It was designed as a general purpose detector, that is, not for detection of only a specific signature or type of events. Since LHC is more of a "discovery machine", the prime goal of the detectors is to discover and not to miss new physics that may be produced in proton-proton collisions. This may seem like a contradiction: the detector should be designed to record a new kind of events, that have unknown properties by definition. But there were several theoretically motivated types of events that, if existed, should be recorded by all means. The Higgs boson, supersymmetric particles, new massive vector bosons, extra dimensions, and of course Standard Model processes were the types of events used to test the possible designs of the experiment.

The Standard model Higgs boson was used as a benchmark to test the performance of the proposed designs. Its mass was constrained by direct searches at LEP and theoretically, but was unknown at that time. Different mass ranges would result in different decay modes and hence different background situations. Decays into hadrons would be nearly impossible to separate from large multijet background, that is why the detection of isolated leptons and photons was preferred. Among leptons, muon has very convenient properties: it is heavy enough to loose very little energy while interacting with matter, but not as heavy as tau lepton to decay quickly. Muons are relatively easy to detect and the CMS has a strong focus on measuring them very well (hence "Muon" in the name of the detector).

Decays of Higgs boson into a pair of photons, pair of Z bosons were among the most interesting signatures. Good energy reconstruction is essential to infer the transverse momentum of escaping neutrino, which is important if W bosons are involved in the search. Identification of b quarks by the displaced secondary vertex is important not only for Higgs searches, but also for Standard Model measurements.

Supersymmetry and extra dimension models would manifest themselves through high energy quarks and leptons and heavy invisible particles escaping the detector. This implies that good energy measurements are also necessary.

The Standard Model offers a vast amount of processes that can be calculated theoretically and measured experimentally. One of them is the subject of this thesis.

Overall, the design requirements for the CMS detector are the following[4]:

- Good muon identification and momentum resolution
- Good charged particle momentum resolution and reconstruction efficiency, efficient triggering and offline tagging of tau leptons and b-jets, requiring a pixel detector close to the interaction point
- Good electromagnetic energy resolution, good diphoton and dielectron mass resolution, π^0 rejection and efficient photon and lepton isolation at high luminosities
- Good missing transverse momentum and dijet mass resolution, requiring calorimeters to be finely segmented

One of the most influential design decisions is the magnetic field configuration. In order to precisely measure momentum of energetic charged particles large bending power is needed. Magnetic field is needed inside the tracker to measure momenta of all charged particles and outside – to measure momenta of muons. The choice was made to use one superconducting solenoid to supply the magnetic field, unlike the ATLAS detector[39] where a superconducting solenoid is located inside hadronic calorimeter, and system of coils produces toroidal magnetic field outside exclusively for the muon system.

The layout of CMS detector is shown in Figure 3.2. The detector has cylindrical structure: it consists of central "barrel" part and two "endcaps". All detectors in the barrel, except for the muon chambers, are housed inside the superconducting solenoid. Muon detectors are sandwiched between iron structures that serve as a return yoke for the solenoid magnetic field.

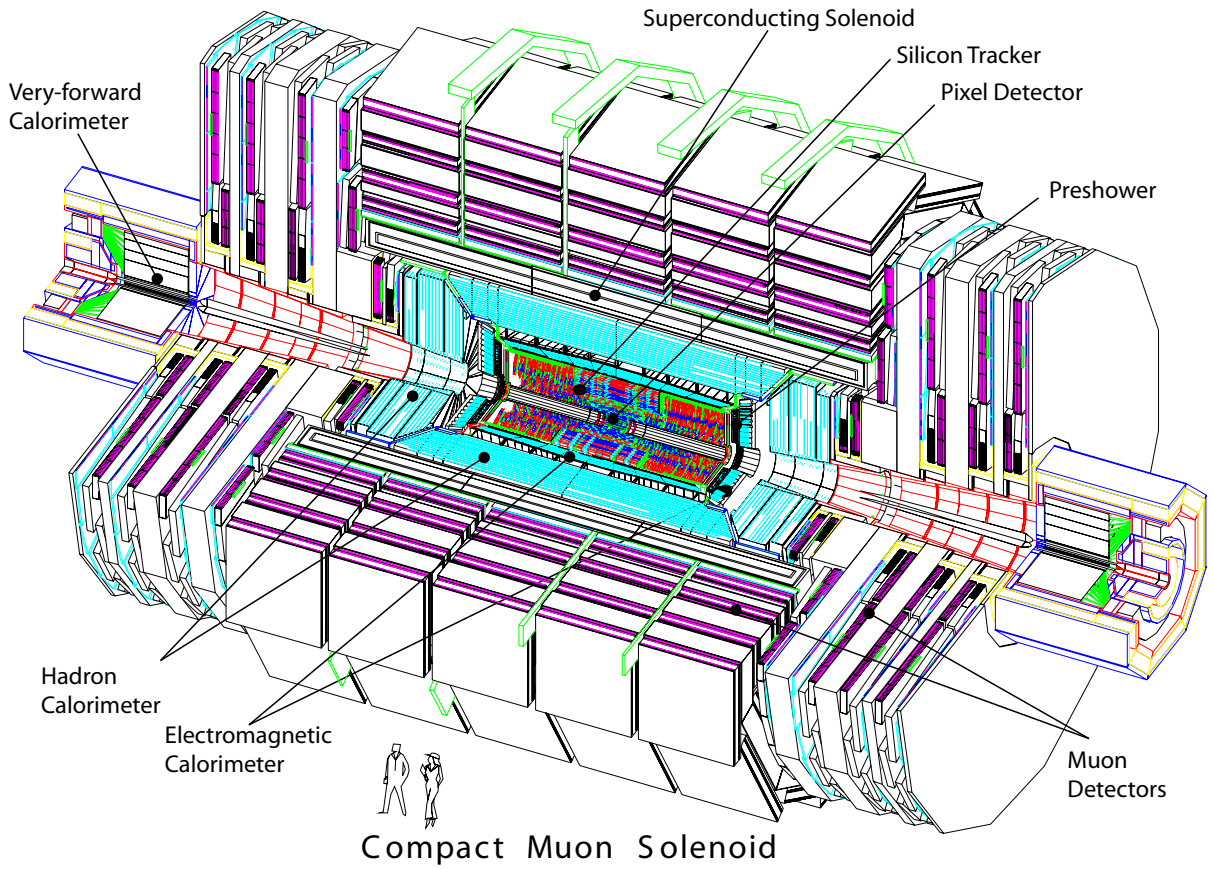

Figure 3.2: *Overview of CMS detector, from*[4]

Inside the magnet there are several layers of detectors that perform different functions.

The innermost part closest to the interaction point (geometrical center of the detector) is silicon pixel tracker which has 3 layers, see Figure 3.3. The silicon micro-strip tracker is outside and has 10 layers of detectors, see Figure 3.4.

These detectors are measuring the positions of charged particles that go through them. Charged particles create electron–hole pairs in sensitive semiconductor regions, electrical signal is then amplified and recorded. Timing is a very important part of the data.

Tracking information is used to measure particle impact parameters and find secondary vertices. The curvature of the tracks gives information about the momenta of charged particles. This detector should be made as thin as possible as not to affect the particles trajectories and avoid interactions like photon conversions. Tracking information is essential

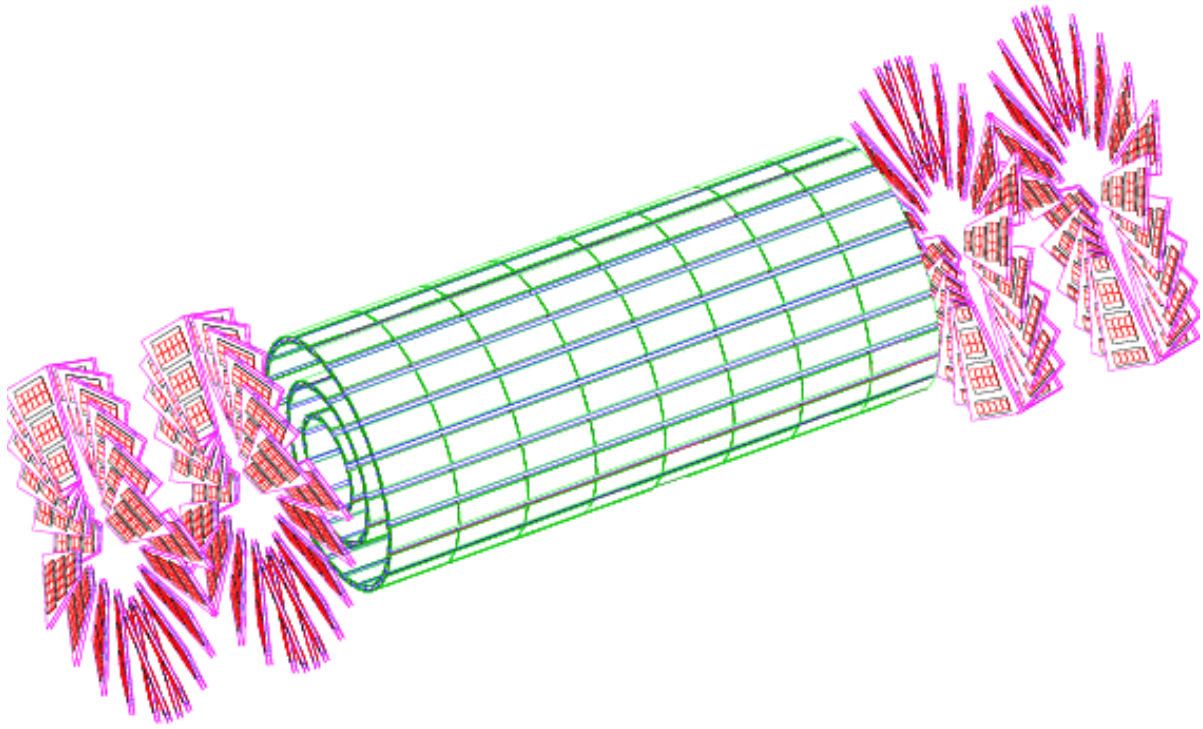

Figure 3.3: *Layout of pixel tracker, from*[4]

and serves as a starting point for particle reconstruction.

Outside of the tracker is the electromagnetic calorimeter (ECAL), Figure 3.5. It uses lead tungstate (PbWO4) crystals to absorb energy of electrons and photons by creating electromagnetic showers. Energy is inferred by measuring scintillation light. The length and density of the crystals allows electrons and photons to loose all energy. By using a finely segmented structure, the position of the electromagnetic shower can be reconstructed with a high precision.

The hadronic calorimeter (HCAL) surrounds the ECAL. It is brass/scintillator sampling calorimeter that measures energy of hadrons that have passed through the ECAL by creating hadronic showers and measuring the light produced in the scintillators by charged particles. Unlike the ECAL, where the absorber and scintillator is the same medium, the HCAL layers of brass serve as an absorber (where most of hard interactions take place) and the scintillating plastic sandwiched in-between produces light when particles go through. The

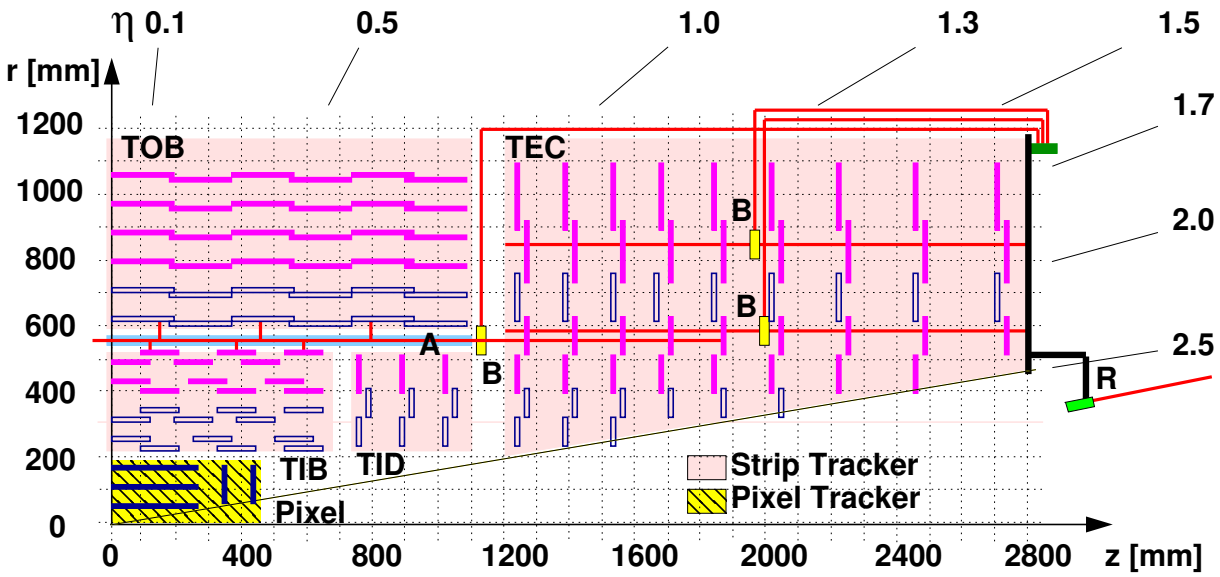

Figure 3.4: *Schematic view of silicon tracker cross section together with laser calibration system, from*[5]

light is transferred by optical fibers to photodetectors. This method has lower spacial and energy resolution, but there is no better alternative considering the space constraint and possible energies of the incoming hadrons.

Outside of the magnet 4 layers of muon station are located, see Figure 3.6. Muons are the only known charged particles that can not be slowed down by calorimeters. Muon stations are tracking detectors, although with much coarser granularity than the inner tracker. Different types of detectors are used in different regions depending on the expected charged particle flux. Drift tubes (DT) are used in the barrel region, resistive plate chambers (RPC) are used in barrel and endcap regions, and cathode strip chambers (CSC) are used in the endcap and forward regions.

The endcaps have ECAL, HCAL and muon system parts that extend the coverage closer to the beam pipe.

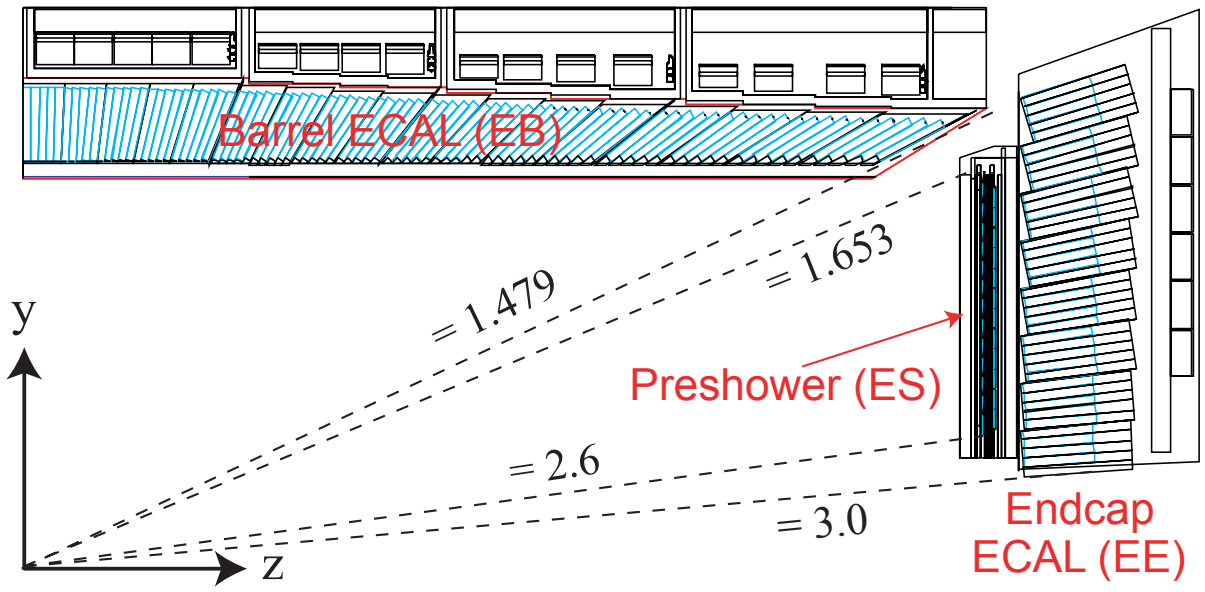

Figure 3.5: *Cross section view of electromagnetic calorimeter, from*[4]

3.3 Trigger System and Particle Identification

During normal operation the LHC delivers bunches of protons crossing at the rate of 40 MHz, which is about 10^9 interactions per second. Considering the amount of raw data collected for one event and the rate at which data can be written to archival media, it is impossible to record all events. In fact, it is necessary to achieve a reduction factor on the order of 10^6. This can be compared to the task of finding a needle in a truck full of hay as it drives by.

The decision of whether this particular event is worth saving must be made very quickly – other potentially interesting events are appearing in the pipeline. It is also critical not to make a mistake in the algorithm: if some class of events is not saved it will never be analyzed (and discovered). Hence trigger system has a very strong influence on what kind of physics can be studied.

CMS trigger system has two main parts: Level-1 trigger (L1) and High Level Trigger (HLT). Events are first very quickly and superficially analyzed by L1 trigger and decision is made to keep or discard them. The event rate is then reduced by an order of 10^3. Then

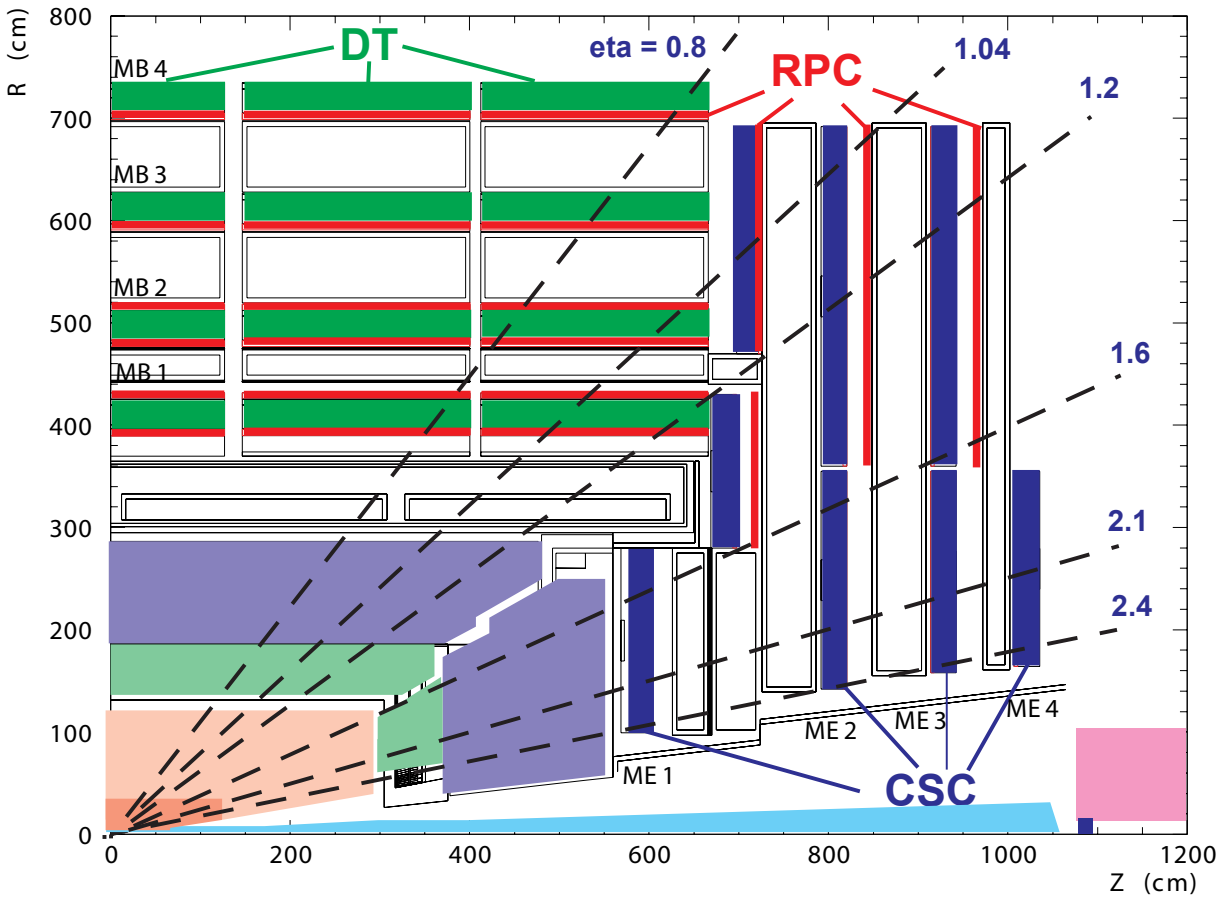

Figure 3.6: *Cross section view of CMS detector with emphasis on muon system, from*[4]*. Detectors chematically shown: pixel tracker in brown, strip tracker in light brown, ECAL in light green, HCAL in purple.*

events that passed the L1 are fully reconstructed and sorted by the trigger paths – particular type of events of interest. Examples may be: events with at least one isolated muon, events with three jets and missing transverse momentum, etc. Events that do not fall into any trigger path are discarded.

Level-1 trigger receives information from calorimeters and the muon system. The decision is based on the trigger primitive objects: photons, electrons, muons, jets (all above certain transverse momentum thresholds). These objects are reconstructed with reduced resolution data. Logic is made with custom Application Specific Integrated Circuits (ASIC), Field Programmable Gate Arrays (FPGA) and such for maximum speed. During the time the

decision is being made, all of the data from detector must be kept in pipelined memory.

When L1 the trigger gives positive decision all the data related to particular event is transferred to one of the HLT machines that run full reconstruction software. This approach gives much more flexibility and configurability (it is easier to make changes to the reconstruction software than to ASIC), but it requires a large computing farm. (It would be impossible for L1 trigger to operate in this manner)

Full event reconstruction is a complicated task involving taking all information from the detector and identifying and reconstructing each individual particle that was created in proton-proton collision. The Particle Flow (PF) algorithm[40] is currently used for event reconstruction.

Tracking information is a key element in PF reconstruction. The momentum of charged hadrons is measured by a tracker much better that by calorimeters (in the range below few hundred GeV). The tracker also provides very precise information on the location of the interaction point, and helps to distinguish particles from different collisions that happened simultaneously (10-20 interaction happen in each bunch crossing). Tracks are matched with clusters of energy in the calorimeters and hits in the muon system to identify the type of particle that made each track. All reconstructed particles are categorized by their type:

- PF muon
- PF electron
- PF charged hadron
- PF photon
- PF neutral hadron

The last two categories are for clusters of energy in the calorimeters which do not have tracks pointing to them. Jets (collimated bunches of lower energy particles produced by a

high energy quark or gluon) are then reconstructed by merging reconstructed particles, not just clusters of energy in calorimeters.

The number of protons per bunch is big enough so that several proton-proton interactions happen during the bunch crossing. They are spatially separated and tracking information is enough to find the vertices – points in space where tracks originate. The vertex with the highest total transverse momentum of the tracks originated there is called the primary vertex. It is the point where the most energetic collision occurred.

For any analysis, the first step is to select interesting events. Selection is usually done by requiring certain objects passing quality cuts. There is no one single recipe of selection that satisfies all needs, as background contribution may be different in each case. Usually events are selected by the presence of isolated leptons (electrons or muons) or jets above certain energy threshold.

3.4 Coordinate System and Variables

The CMS uses a right-handed coordinate system with Y axis pointing up, X axis towards the center of LHC and Z axis along counter-clockwise direction around the ring, see Figure 3.7. A polar angle θ is measured from the positive Z axis direction, and azimuthal angle ϕ in XY plane is measured as shown. Among other useful variables is pseudorapidity

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

For an ultra-relativistic particle pseudorapidity is equal to rapidity, which is linear with respect to Lorentz boosts. A transformation made by two consecutive boosts along the Z axis will have rapidity equal to a simple sum of rapidities of these two separate transformations. This makes pseudorapidity a convenient choice of "distance" between two particles in the θ direction, as they may be moving along the Z axis together. When the ϕ distance is added,

one gets a standard ΔR measure of distance between two particles:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$$

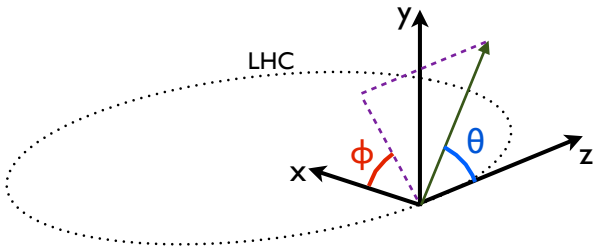

Figure 3.7: *CMS coordinate system*

There are several variables that are extensively used later in the text.

- Transverse momentum: the projection of a particle's momentum onto the X-Y plane (transverse to the beam line). It is a vector quantity.
- Isolation of a particle: the sum of absolute values of transverse momenta of other particles surrounding it within a circle of a certain radius in $\eta - \phi$ coordinates. See Figure 3.8 for clarification. It can be calculated separately for different types of reconstructed particles, and three types of isolation are used: charged hadron isolation, neutral hadron isolation, photon isolation.
- Relative isolation: isolation divided by a particle's transverse momentum.
- Detector based isolation: calculated by using spatial cluster separation in the calorimeters. Does not require particle flow machinery, and is used for triggering.
- Missing Transverse Energy (MET): negative vector sum of transverse momenta of all objects reconstructed in the event. It should be called missing transverse momentum,

(which it is) but in the old days this quantity was calculated using only energy deposits in the calorimeters, hence its name.

- $\sigma_{i\eta i\eta}$ photon's or electron's electromagnetic shower core elongation in η direction, calculated for 5x5 EM calorimeter crystal array with the most energetic crystal in the center.

$$\sigma^2_{i\eta i\eta} = \frac{\sum_i (\eta_i - \bar{\eta})^2 w_i}{\sum_i w_i}; \bar{\eta} = \frac{\sum_i \eta_i w_i}{\sum_i w_i}; w_i = max(0, 4.7 + log(E_i/E_{5\times5}))$$

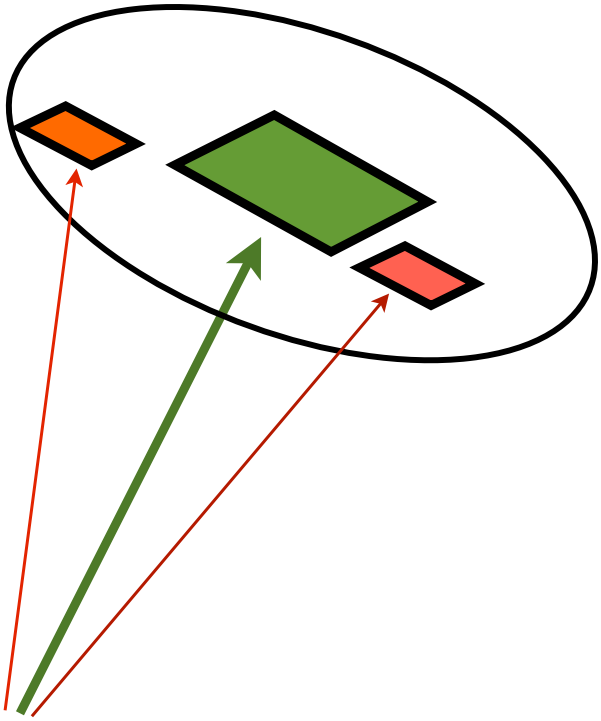

Figure 3.8: *Graphical representation of particle isolation. Green: energy cluster of the particle, Red: energy deposits that fall into isolation cone.*

In cases when there are more than one hard interaction in the event (which happens almost all the time) the isolation of photons and electrons will be "spoiled" by particles unrelated to the primary interaction. The amount of extra particles may be different from event to event. To correct for the presence of this random energy deposits the algorithm

called "ρ correction" was developed. It estimates the level of "noise" in the event, which can be subtracted from particle isolation using a correction factor.

3.5 Simulation of Events and Detector Response

The complexity of the detector and reconstruction algorithms requires some way to check that chosen detector design suits the purpose and the reconstruction performs well. The software package GEANT4[41] is capable modeling all necessary physics processes happening when charged or neutral particles or quanta of light interact with matter in the detector. A very detailed model of the CMS detector was created first to test the expected performance before the detector was built and continues to be extensively used and updated.

The GEANT4 model can simulate energy deposits in active detector material (simulation). Detector electronics response is then simulated (digitization) and it can be fed into standard reconstruction algorithms (reconstruction). Results of the reconstruction can be compared with simulated particles to estimate detector performance.

This leaves the initial step of generation of the particles that reach the detector. Event generators[42,43] are computer programs that use the existing knowledge of proton structure and Standard Model Lagrangian to randomly sample the possible final states that can occur after high energy particle collision. The calculation is done in several steps. First, hard scattering is modeled, where two incoming particles generate several energetic remnants, such as photons, W or Z bosons, gluons, quarks and leptons. In this process quarks and gluons are treated as free particles. Then, in order to satisfy the condition that all final state particles must be color neutral, the showering of the quarks and gluons is done. The algorithms use some heuristics to make color neutral final particles by adding quark-antiquark pairs to the event, making multiple mesons and baryons (thus making "jets" of particles). Some of the mesons and baryons can be unstable and decay, these processes are taken into account. In the end, the generator provides the list of particles with their momenta as input for the

Table 3.1: *Simulated samples used in the analysis.*

Process	Description	cross section(pb)
TTGamma	top quark pair and photon	2.9 (estimate)
TTJets1l	top quark pair and extra jets, semileptonic decay	104.7
TTJets2l	top quark pair and extra jets, dileptonic decay	25.09
TTJets0l	top quark pair and extra jets, all hadronic	109.2
W3Jets	W boson and exactly 3 extra jets	519
W4Jets	W boson and 4 or more extra jets	214
ZJets	Z boson and extra jets	3350
Wgamma	W boson and photon	553.9
Zgamma	Z boson and photon	159.1
SingleTop	single top or anti-top quark	1.8 - 56.4

GEANT4 detector model.

It is convenient to separate simulated collision events by the type of intermediate and final states and interactions involved. In this way events with different probabilities (cross sections, or weights) are kept in separate "datasets". One should be aware that this distinction is not absolute, and overlap between datasets is possible. This is done for convenience of bookkeeping and analysis. In real collisions all processes occur with their natural probabilities and only by selection one can enrich the sample with events of certain type. Table 3.1 summarizes the datasets used in this analysis listing their cross sections, which are proportional to the probability to see this event in proton-proton collision (more details in Section 6).

There are several ways to produce a single top quark (by different ways to look at Wtb vertex).

Most of the simulated datasets are generated using MadGraph 5[42] with parton distribution function set CTEQ6L1. Single top quark events are generated by POWHEG[44]. Decay of tau lepton is modeled with TAUOLA[45]. Signal TTGamma sample was simulated with MadGraph. Hadronization and showering is done by PYTHIA 6.4[43].

Chapter 4

Analysis Outline

In this chapter the analysis methodology is presented.

To perform the $t\bar{t} + \gamma$ cross section measurement, events with a top pair and a hard photon are selected. Among all of the possible top pair decays, we are interested in a semileptonic signature with one electron and jets in the final state.

In order to reduce systematic uncertainties associated with top pair production cross section $t\bar{t}$ and integrated luminosity we measure the cross section ratio $\sigma_{t\bar{t}+\gamma}/\sigma_{t\bar{t}}$. Then it can be converted to $\sigma_{t\bar{t}+\gamma}$ by multiplying by a recent theoretical or experimental $t\bar{t}$ cross section.

The analysis can be divided in two steps:

1. *Top Selection* is performed to get the $t\bar{t}$ enriched data sample, that is used for normalization of the $t\bar{t}$ simulation and subsequent selection.

2. *Photon Selection* is applied on events passing top selection and aimed to select events containing a genuine photon.

In both steps of the analysis there are events that pass the selection but do not have the desired property:

- $t\bar{t}$ background contains events that pass top selection but do not have a top quark

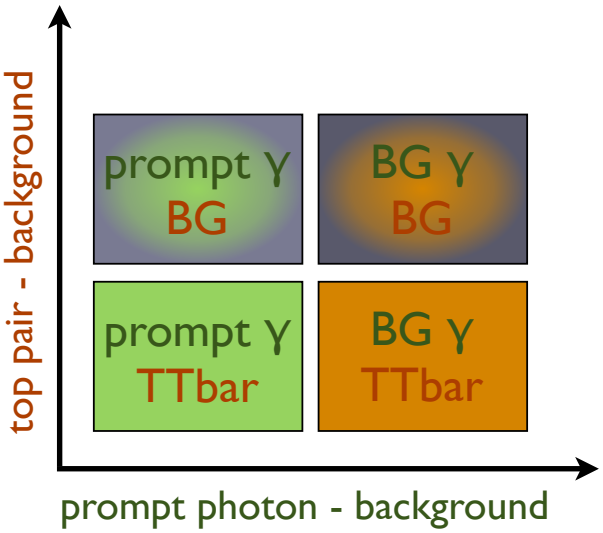

Figure 4.1: *Cartoon representation of signal and background for the two steps of the analysis.*

pair. They may come from W or Z boson production associated with jets and photons, multijet (or QCD) events, single top quark production.

- Photon selection is aimed at isolated photons from hard scattering. Non isolated photons from jet fragmentation and jets with small hadronic component or electrons may be misidentified as isolated photons.

These two properties of event are independent, as schematically shown on Figure 4.1. The details of event selection are described in Chapter 5.

This analysis employs template fits to find the purity of $t\bar{t}$ and purity of genuine photons passing photon selection. Finally, the number of signal events is inferred using these parameters and simulated datasets. See Chapter 6 for details.

4.1 TTGamma Signal Simulation

To generate TTGamma events it is necessary to define the final state with some cuts on final state particles' energies and spacial separation, especially photon energy and separation of

photon from other charged final state particles. The reason for it lies in the method of cross section calculation. It happens that a cross section grows when photon energy becomes smaller, and goes to infinity when photon energy goes to zero (infrared divergency). A similar problem exists when considering photons emitted very close to the charged particle (collinear divergency). Hence, to get a meaningful result from the calculation the cuts on photon energy and distance to other particles have to be applied.

The phase space for generation was chosen in the following way:

- $p_T(\gamma) > 13GeV$
- $|\eta(\gamma)| < 3.0$
- $\Delta R(\gamma, all) > 0.3$
- $p_T(jet) > 15GeV$
- $p_T(b) > 20GeV$
- $|\eta(b)| < 5.0$
- $|\eta(jet)| < 5.0$
- $|\eta(lepton)| < 3.0$
- $\Delta R(jet, jet) > 0.5$
- $\Delta R(jet, lepton) > 0.5$

There is no cut on the lepton transverse momentum, but there are cuts on the quarks' (jets) momenta. This makes the ratio of hadronic and leptonic W decays generated with these cuts differ from the usual number.

4.2 Overlap Removal

A signal sample (TTGamma) will overlap with TTJets samples in cases when a hard photon is radiated by initial state quarks, top quarks, b quarks, W and its decay products (light quarks or leptons) during the showering step of simulation. The difference between these photons and signal photons is in the generation step they are produced. Signal photons from TTGamma sample are produced in the very first step – they are in the matrix element. Showering photons are subject to higher modeling uncertainty and should be avoided if possible.

To evade double counting of events we apply an overlap removal procedure to remove such events from TTJets samples. This procedure is applied only on TTJets events. In order to be considered overlapping with TTGamma, the event has to have at least one generator level photon with the following properties:

- $p_T(\gamma) > 13GeV$
- $|\eta(\gamma)| < 3.0$
- Only quarks, gluons, bosons or leptons are in the parents list. This ensures that photons from π_0 decays are not considered as signal
- $\Delta R(\gamma, other) > 0.2$ where other particles include leptons, b quarks and final state particles (hadrons, charged leptons, photons) with transverse momenta above 5GeV

The last cut is used to suppress photons from showers. In such cases, parental information will show that photon is radiated by electron, but it may be collinear to it. Figure 4.2 shows the distribution of the ΔR distance between generator photon and nearest generator particle for signal (TTGamma) and TTJets2l where considerable fraction of reconstructed photons comes from electrons radiating photons.

Similarly, there is an overlap between ZJets and Zgamma, and WJets and Wgamma samples. Overlap removal procedure is applied on ZJets and WJets samples. In this case

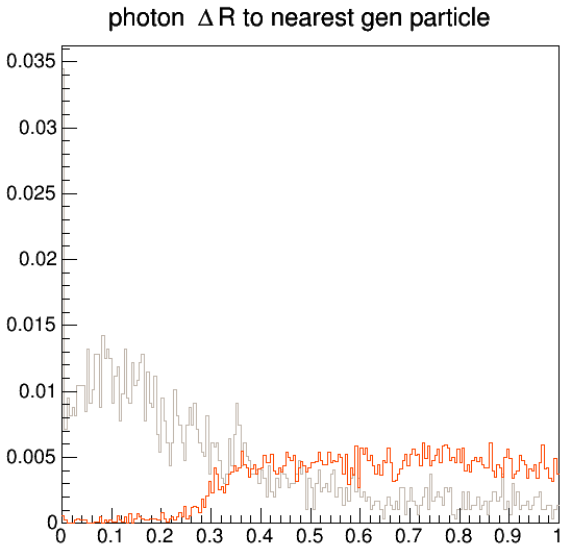

Figure 4.2: *Distance from generator photon (matched with reconstructed photon) to nearest generator particle for TTGamma (Red) and TTJets2l (Grey). Parentage cut is applied. Normalized to match total area.*

events with a signal photon on the generator level are removed if they are initial state radiation (emitted off the colliding partons) or final state radiation of leptons from Z or W (will be included in Zgamma or Wgamma).

Chapter 5

Event Selection

Data recorded with the CMS detector in the year 2012 is used in this analysis. It corresponds to an integrated luminosity of 19.6 fb^{-1}.

An event should contain at least one primary vertex with the following properties:

- Number of degrees of freedom in the vertex fit should be greater or equal to 4
- Z coordinate of the vertex should be within 24cm from the geometric center of the detector
- Distance from vertex to the z axis should be less or equal to 2cm.

5.1 Trigger selection

Events that passed "single isolated electron" triggers are used in this analysis. Below are some details on how selection at high level trigger is done. Trigger selection starts with L1 single egamma (EG) object with p_T above 20 or 22 GeV, for different data-taking periods. Then the high level trigger selection requires: p_T greater than 27 GeV, shower shape cut on $\sigma_{i\eta i\eta}$, relative detector based Ecal and Hcal isolation cuts (no ρ corrections applied) and the relative track isolation cut. In the next step track matching is done. This allows a selection

of isolated electrons on the trigger level. Detector based isolations are not used later in the analysis. ρ correction is always used later in object selection. This makes the comparison of trigger cuts and analysis cuts rather difficult.

Electron selection on the analysis level applies several additional cuts to identify isolated electron and veto events with more than one isolated electron.

5.2 Physics objects selection

The next part of the event selection where individual reconstructed objects are selected is done in several steps. There are selection cuts applied on an individual object basis. Electrons, loose electrons, jets, loose muons and photons are selected. Then an additional set of requirements is applied based on the relative positions of the objects (ΔR cuts). After that, a final decision is made if the event is to be considered in the further analysis.

The isolation of all objects is ρ corrected to reduce pileup dependence.

Electrons are required to have:

- Transverse momentum p_T greater than 35 GeV
- Absolute value of pseudorapidity less than 2.5 excluding the gap 1.4442 to 1.566
- Combined relative PF isolation in cone 0.3 to be less than 0.1
- Trigger version of electron multivariate discriminator (`MVAtrig`) greater than 0.9
- Conversion rejection: there should be no extra tracks pointing same direction
- Distance in the plane transverse to the beam between electron track and primary vertex should be less than 0.02cm

Loose electrons are selected from electron candidates that failed electron selection and have:

- Transverse momentum greater than 20 GeV
- Combined relative isolation in cone 0.3 to be less than 0.2
- `MVAtrig` greater than 0.0

The motivation for having loose electrons in the selection originally was to reject events containing Z decaying to two electrons. The selection of exactly one electron rejects these events well, so the number of loose electrons is not used in the event selection.

Jets are reconstructed by the PF AK5 algorithm, and the following selections are made (PF loose Jet ID). Before applying any selection, the following corrections are made to account for imperfect jet energy measurement: Jet Energy Scale correction, and Jet Energy Resolution smearing. See more details about systematic uncertainties in Section 6.8. Jets are selected by applying the following requirements:

- Transverse momentum greater than 30 GeV
- Absolute value of pseudorapidity less than 2.4
- Number of constituents greater than 1
- Charge multiplicity greater than 0
- Neutral hadron fraction of energy less than 0.99
- Neutral electromagnetic energy fraction less than 0.99
- Charged EM energy fraction less than 0.99
- Charged hadron energy fraction greater than 0

These cuts help to avoid picking detector noise and Ecal spikes as jets.

B-tagged jets are identified with Combined Secondary Vertex b-tagging algorithm in its medium working point (CSVM) Event re-weighting is applied to correct for the difference in b-tagging efficiency in data and simulation as explained later in Section 5.5.

Loose Photon Candidates are defined as *Photons* without $\sigma_{i\eta i\eta}$ and charged hadron isolation cuts.

Photons are selected by applying modified EGamma Loose photon ID. The definition of isolation was modified in order to make data-driven estimate of selection purity more robust. Photons are required to have:

- Transverse momentum p_T greater than 25 GeV
- Absolute value of pseudorapidity less than 1.4442 (only Ecal Barrel region is considered)
- No pixel seeds on the way from primary vertex to the photon cluster in the calorimeter
- No matched prompt electron should be found (the last two cuts significantly reduce fraction of electrons misidentified as photons)
- Hadronic over electromagnetic energy fraction should be less than 0.05
- Electromagnetic shower core elongation in η direction $\sigma_{i\eta i\eta}$ should be less than 0.012
- Particle flow neutral hadron isolation less than $3.5\ GeV + p_{T\gamma} \cdot 0.04$
- Particle flow photon isolation less than $1.3\ GeV + p_{T\gamma} \cdot 0.005$
- Super Cluster Footprint Removed (SCFR) charged hadron isolation less than 5 GeV

Standard particle flow charged hadron isolation cut is not used in photon selection.

Muons are selected by requiring:

- Transverse momentum greater than 10 GeV
- Absolute value of pseudorapidity less than 2.5
- Combined relative isolation in cone 0.4 to be less than 0.2.

5.3 Footprint removal for photon isolation

Charged hadron isolation is used for the photon purity calculation, and this part of photon selection requires special attention. We use a method called supercluster footprint removal[46] to find one of the components of photon isolation.

The idea of this method is to calculate isolation in such a way that super cluster leakage into the isolation cone is minimized.

Isolation of a photon is defined as sum of transverse momenta of all particles falling in the isolation cone. The standard cone size is $\Delta R < 0.3$ which means $\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.3$. Differences in ϕ and η are taken from photon to other particle. The photon momentum should not contribute to the isolation sum and, hence, particles found close to the photon are not summed (veto cone).

In the standard particle flow algorithm isolation is calculated separately for different types of particles. This is done for charged hadrons, neutral hadrons and photons. This procedure has some potential pitfalls.

- If some energy deposit in the calorimeter is not associated with any reconstructed particle it is not counted in the isolation sum and thus lost.

- If photon energy is spread widely in the detector it can be reconstructed as several particles and they will affect the isolation of the photon.

These drawbacks are not too serious when dealing with standard cut-based particle ID, but if the isolation profile shape is of interest some improvements can be made.

Super cluster footprint removed isolation in contrast with PF isolation deals with particle flow *candidates*. The fact that particle flow candidates, but not identified particles are used makes effect of reconstruction inefficiency on isolation less pronounced. This helps to avoid the first pitfall.

The second problem is about leakage of photon energy into the isolation cone. A different method of defining veto area around the photon is used. The isolation is calculated by

summing particle flow candidates that fall into the isolation cone, but not close to the photon super cluster in the electromagnetic calorimeter. η and ϕ dimensions of the super cluster are calculated by checking positions of hits that contribute to it. If a particle flow candidate extrapolated to the calorimeter surface falls in the rectangle of the super cluster size enlarged by 25%, then it is considered close to the photon and is not added to the isolation sum.

The standard ρ correction is used to account for pileup.

5.4 Event selection

In this section we explain how information about identified objects is combined to make a decision if the event passes the selection.

The isolation cut for an electron will reject objects that are close to hadronic jets. To make this rejection even stronger we apply ΔR cuts on jets and electrons in the following order. First, jets that have $\Delta R < 0.1$ to any electron or loose electron are removed from the jet list. Second, electrons that have a reconstructed jet with transverse momentum above 20 GeV within $0.1 < \Delta R < 0.5$ are removed from the electron list. It is important to take into account not only jets, that passed jet ID and transverse momentum cut on the analysis level.

The following cuts are made on the event level:

- Number of electrons is exactly 1
- No muons are present
- Number of jets is greater or equal to 3
- Number of b-tagged jets is greater or equal to 1
- Missing transverse momentum (negative vector sum of all particles transverse momenta) greater than 20 GeV

Particle flow corrected MET (corrections type 0 and type 1) is used in the selection cut. This selection and lepton veto enhances purity of semileptonic top pair decays since they have a genuine neutrino and only one isolated lepton. It happens due to reduction of multijet contribution, which usually do not have high missing energy.

The cuts described above are *Top Selection* cuts, and all events should pass them to be considered further. Events passing this selection are used to find number of $t\bar{t}$ events.

Next, another set of ΔR cuts is applied. Photons that lie between 0.1 and 0.7 to any jet with transverse momentum above 20 GeV are removed from the photon list. Also, photons that are closer than 0.7 to an electron are removed. This cut helps to avoid photons that come from final state radiation of quarks and electrons.

Events passing *Photon Selection* should have at least one photon.

We employ data-driven method for photon purity estimation as explained later. This method requires additional information about photon identification background. We keep events passing *Top Selection* and having at least one loose photon candidate for this purpose.

5.5 Corrections of simulated events

Simulation does not describe all the processes happening in proton-proton collisions and in the detector perfectly. There are several reasons why we want to correct simulation to improve description of data. The simplest way to do this is to find a variable that characterize the difference between data and simulation and re-weight simulated events in such a way that chosen distribution matches data. The variable should be chosen carefully so it will not be directly related to the quantity to be measured, otherwise such re-weighting can bias the result. Here, we list the variables used for simulated events re-weighting. The total event weight is a product of all of them.

Pileup re-weighting. Simulated events are produced with some reasonable number of pileup interactions, but they have to be weighted in order to match distribution of pileup

interactions measured in data.

B-taggging re-weighting. B-tagging is known to have slightly different efficiencies in data and simulation. The re-weighting is done according to the recommendations of the BTV group. The scale-factors (SF) for probability of tagging jet in data if it was tagged in simulation are provided. To calculate the weight for the event with 1 or more b-tag, we find probability of having exactly zero b-tags by $\prod_i(1 - SF_i)$ iterating over jets that were b-tagged in simulation. The weight for one or more b-tag is $1 - \prod_i(1 - SF_i)$.

Electron trigger, reconstruction and identification efficiencies in data and simulation were measured by tag and probe method in bins of p_T and η. Scale factors are calculated as ratio of data to simulation efficiencies in each bin.

Top p_T re-weighting. It was found that transverse momentum of top quark in simulation does not match one on data. The event weighting procedure was developed to fix it. For every simulated $t\bar{t}$ event event the weight is calculated to bring top transverse momentum closer to one observed in data.

Chapter 6

Cross Section Measurement

6.1 Ratio measurement, description of steps

The usual form of "counting experiment" cross section measurement is

$$\sigma = \frac{N_{signal}}{\mathcal{L}\epsilon A},$$

where N_{signal} is number of signal events in data passing selection, $\mathcal{L}$ is integrated luminosity, A is signal acceptance – fraction of signal events that fall into the region of phase-space chosen for event selection, and ϵ is selection efficiency – fraction of signal events that pass event selection after acceptance cuts. In the presence of background, we will find N_{signal} as $N_{observed} - N_{background}$ where number of background events should be estimated separately.

If event selection is done in two steps sequentially, the expression for cross section will change. In our notation selection step one is *top selection* and selection step two is *photon selection*, which is applied after top selection, as described above.

$$\sigma_{t\bar{t}+\gamma} = \frac{N_{signal}}{\mathcal{L}\epsilon_{top}A_{top}\epsilon_{\gamma}A_{\gamma}}$$

where ϵ_{top} and A_{top} are efficiency and acceptance of top selection, for top pair plus photon

signal. Then, ϵ_γ and A_γ are the efficiency and acceptance of the photon selection, applied after top selection. The idea of this calculation can be directly seen from a conditional probability expression $P(A \cdot B) = P(A) \cdot P(B|A)$. Here, we consider the signal as a top pair and photon. N_{signal} is number of events in data after top and photon selection with genuine photon and top pair.

In this way we can find the desired top pair plus photon cross section provided we can find all parameters involved.

The next step is to normalize the result to inclusive top pair production cross section to simplify treatment of systematic uncertainties and cancel extra terms:

$$R = \frac{\sigma_{t\bar{t}+\gamma}}{\sigma_{t\bar{t}}} = \frac{N_{signal}}{\epsilon_{top}^{t\bar{t}\gamma} A_{top}^{t\bar{t}\gamma} \epsilon_\gamma^{t\bar{t}\gamma} A_\gamma^{t\bar{t}\gamma}} \cdot \frac{\epsilon_{top}^{t\bar{t}} A_{top}^{t\bar{t}}}{N^{t\bar{t}}}.$$

Here, the second fraction comes from $t\bar{t}$ inclusive cross section: $\epsilon_{top}^{t\bar{t}}$ and $A_{top}^{t\bar{t}}$ are the efficiency and acceptance of $t\bar{t}$ process, and $N^{t\bar{t}}$ is the number of top pair events passing top selection. The efficiency for top pair selection is very similar for $t\bar{t}$ and $t\bar{t}\gamma$ samples, as selection of lepton, jets, b-jet, MET are the same. Acceptance depends on the generator cuts and event selection, so it should be calculated for both $t\bar{t}$ and $t\bar{t}\gamma$ separately. Integrated luminosity has been canceled out.

The main difficulty in this analysis is to find the number of signal events.

Objects that pass photon ID are not guaranteed to be real photons. Jets or electrons can be misidentified as photons. These two cases are distinct.

A jet misidentified as a photon does not have genuine photons in them but happen to pass photon selection. Electron misidentified as photon has all the properties of a genuine photon (isolated electromagnetic shower, etc), but in addition to that has a charged track pointing to the energy deposit. These two sources of background will be treated differently.

First, the purity of isolated electromagnetic objects (photons and electrons) passing photon selection $\pi_{e\gamma}$ will be calculated using an isolation profile fit. This method, proven to work on simulated samples, will be used to find photon purity in data.

Then, by using a kinematic variable, M3, the fraction of events having a top pair will be estimated.

These two parameters and total number of observed events in data will be used to estimate the number of signal events in data. The procedure will rely on simulation corrected to describe the important properties of data.

6.2 Multijet estimation

Multijet (QCD) processes give noticeable background to electron + jets top selection channel and are hard to simulate because of their large cross section and small probability of passing the selection.

One of the jets can be misidentified as an electron and trigger the event. Other event selection requirements may also be fulfilled. One needs to simulate a very large number of events to get enough statistics for shape description. Another approach is to get necessary sample from data. We will use the fact that multijet events can fire single electron trigger, and that these misidentified electrons will not be isolated and will not pass MVA ID cut.

We select a sample of multijet events by inverting the isolation and MVA cuts:

- Electron relative isolation from 0.25 to 1.0
- Electron `MVAtrig` from -1 to -0.1

All other electron cuts remain untouched, as well as other top selection cuts. For the multijet sample we do not require a photon.

TTJets and WJets events with genuine electrons can also pass the non-isolated electron selection. Some important distributions may be affected by this contamination. We keep all other event selection cuts so that data-driven multijet events will be by all other means alike normal events passing selection. This selection leaves noticeable amount of TTJets events. See Figure 6.1. We estimate their contribution using simulation and subtract it from multijet sample selected in data.

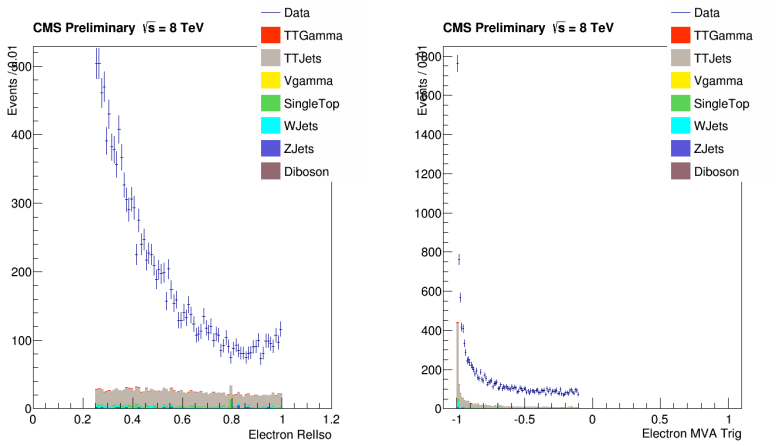

Figure 6.1: *Result of multijet selection. Data and simulation.*

At this point the shape of QCD is known but we have to determine correct normalization to properly mix it with other backgrounds. We will use a MET distribution since it is sensitive discriminator for this type of events. MET in QCD originates from jet energy mis-measurement and has a different shape than in TTJets or WJets where MET is created by high energy neutrino.

To have better shape information we remove the cut on the missing transverse energy. This is for fitting purposes only, and once the normalization is found we apply the MET cut as described in event selection.

The MET shape fit is performed with data after top selection and two templates: QCD shape and all other MC samples added and scaled according to their cross sections. The fit uses only shape information, so that all distributions are normalized to unity. In this way uncertainties of non-QCD normalization will not affect the result. The only parameter of the fit is the relative contribution of the QCD and non-QCD shapes in the data distribution. QCD normalization is determined form scaling integral of data by fraction found from the fit.

The result of the fit is shown on Figure 6.2. After taking into account the number of events in data and fit result, the data-driven multijet histograms have to be scaled. After

the photon selection is applied this gives 28 expected background events. It has a minor contribution to the final result.

Uncertainty of the multijet estimation comes from the fit uncertainty and statistical fluctuation of the templates, but most importantly from the selected procedure. To be sure that multijet uncertainty is not under-estimated we apply the following procedure. The nominal shapes after MET fit is scaled by factor of 2 up or down and the whole cross section ratio is repeated. The difference in the answer is taken as systematic error.

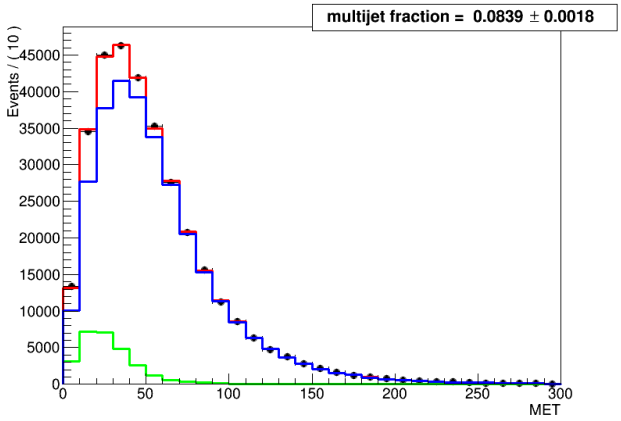

Figure 6.2: *Fit of missing transverse momentum shape to find QCD normalization. Black: data, Green: QCD shape, blue: all MC, red: their sum.*

6.3 Number of top pair events

This part of the analysis is very similar to measuring inclusive top pair production cross section. There are several good methods that had been used for this. We employ the "M3" method. This method uses an M3 variable as a discriminator between events with a top pair and background. It is not very sensitive to jet energy scale (JES) uncertainty.

We perform a fit of the M3 variable to extract $t\bar{t}$ normalization from data. This is done after top selection.

M3 is defined as invariant mass of three jets, selected among all jets in the event, that

have the highest transverse momentum of the 3-body system constructed of these jets. By definition it can not be calculated for events with less than 3 jets. The idea of the M3 variable is the following. In semi-leptonic top pair decay there will be 3 jets from one of the top quarks hadronic decay. By choosing a combination with highest total p_T we have a good chance to select these 3 jets from top quark decay if it was boosted. Hence, the M3 distribution from the top decay will have a peak near 175 GeV, and for all other processes not having hadronic top, it will be a smooth combinatorial background.

The fit procedure is similar to the MET fit described above. We use the signal M3 distribution from $t\bar{t}$ simulation (all decay channels added together), and the M3 distribution from WJets sample as a background. Since WJets normalization is known to be relatively poorly modeled we let it float in the fit. There are other background processes that are estimated from simulation and normalized according their theoretical cross sections. These processes include: single top, Z+jets, Z+gamma and W+gamma. Their normalization is not varied in the fit and their contribution is relatively small. The multijet component of the background is also fixed. These background processed are subtracted from the data histogram before the fit.

The number of top pair events after top selection, $N^{t\bar{t}}$, is calculated by taking the integral of data distribution, subtracting integrals of fixed backgrounds and multiplying the result by a signal fraction from the fit. Both $t\bar{t}$ and WJets simulated distributions are scaled to conform with the fit result and number of events in data. The scale factor for $t\bar{t}$ is measured to be $0.993 \pm 0.008(stat)$, while the WJets scale factor is $2.56 \pm 0.05(stat)$. This shows good agreement of the fit result with top pair production cross section, that was used to normalize the simulated samples. The big scale factor for WJets stems from the fact that only exclusive W+3jets and W+4jets samples were used in the fit. In this analysis WJets is a background process and all we want is to model its contribution correctly. Figure 6.3 shows the fit and the result of scale factor application.

This gives us a number of $t\bar{t}$ events estimate in data after top selection $162560 \pm$

1240($stat$).

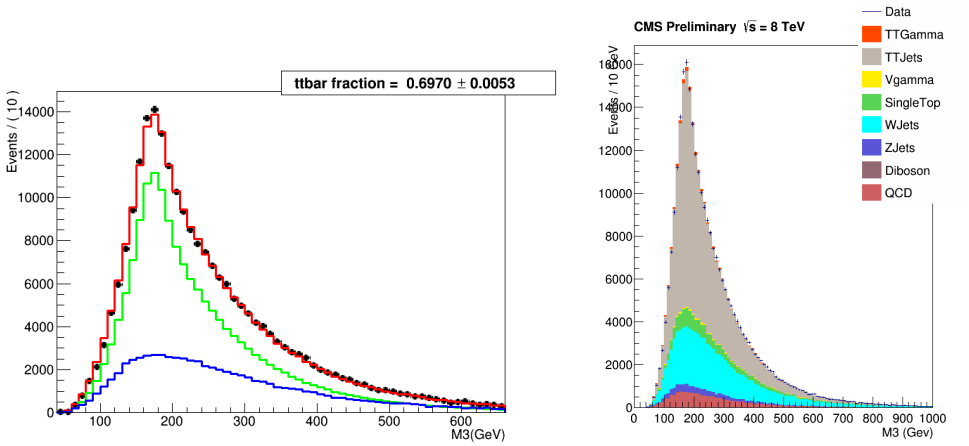

Figure 6.3: *M3 distribution fit and the result of applying scale factors on simulation.*

A similar procedure is done after photon selection to find the total fraction of events having a top pair in the selected sample. In this case, again, there are two templates (top and background). One is taken from the sum of $t\bar{t}$ and $t\bar{t}\gamma$ samples. Here, normalization of $t\bar{t}\gamma$ is taken from a theoretical NLO calculation. This does not add bias to the measurement because M3 shapes for both samples are nearly indistinguishable. The fit was also done with the shape of $t\bar{t}$ before photon selection and gave the same result.

The background shape is constructed from all non-ttbar processes where the shape is taken from the top selection (higher statistics) and rescaled according to the expected event yield after photon selection.

The fit gives a fraction of ttbar events after photon selection as shown on the Figure 6.4.

6.4 Photon purity estimation

The photon purity is a fraction of selected photons that are prompt (signal) photons. A prompt photon originates in a hard interaction and is emitted by hard scattering particles. Non-prompt photons appear during quark hadronization and interaction of charged particles

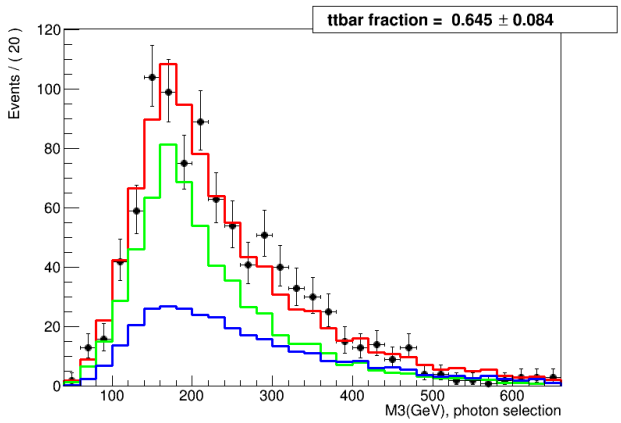

Figure 6.4: *M3 distribution fit after photon selection.*

with detector material. They are not isolated and generally have lower energy.

One way to find photon purity is simulation. In simulated events we have information about all particles involved. To find whether a given reconstructed photon is prompt (signal) we are looking for the generator level photon that can be matched with it. Matching is done by comparing η and ϕ coordinates and transverse momenta of reconstructed photon and generator level photon. We iterate over generated particles collection in the order they were created by the generator. If ΔR between reconstructed photon and generated photon is less than 0.2 and $|p_T^{reco} - p_T^{gen}|/p_T^{gen}$ is less than 1.0 we consider it a match and stop iterating. Then we inspect the parentage of this matched photon looking where it came from. If there are only quarks, gluons, charged leptons or bosons among its ancestors we conclude that photon is prompt. However having the photon originating in the hard interaction does not necessarily mean that it is isolated. Hadronization and showering may introduce nearby activity and the photon originally generated as signal will not be isolated. These photons may pass event selection, but for the purpose of the method explained below it is necessary to add some selection cuts. To differentiate between prompt signal photons and non-isolated photons the following set of additional selections is applied on reconstructed photon and matched generator photon in addition to described above.

- $(p_T^{reco} - p_T^{gen})/p_T^{gen} < 0.1$
- $\Delta R(\gamma gen, other) > 0.2$ where other particles include leptons, photons and final state particles (hadrons, leptons, photons) with transverse momenta above 5 GeV
- $|\Delta\eta(\gamma reco, \gamma gen)| < 0.005$
- $\Delta R(\gamma reco, \gamma gen) < 0.01$

These cuts help to select photons that are well measured and have no undesired activity around them. Otherwise reconstructed photon considered non-prompt or mis-identified (background).

Electrons give a trace in the electromagnetic calorimeter very similar to photons. In addition electron usually has a charged track pointing to the energy deposit. Electrons can be misidentified as photons, although special selection is made to minimize the contamination. Nevertheless there is noticeable amount of electrons misidentified as photons. We need a criteria to find generator level electrons that are isolated. The selection is very similar to generator photon selection. We consider electrons from W and Z decays only.

- $(p_T^{reco} - p_T^{gen})/p_T^{gen} < 0.1$
- $\Delta R(e, other) > 0.2$ where other particles include leptons, photons and final state particles (hadrons, leptons, photons) with transverse momenta above 5 GeV
- $|\Delta\eta(\gamma reco, e)| < 0.005$
- $\Delta R(\gamma reco, e) < 0.04$

These cuts were checked with TTJets2l sample which has significant amount of electrons identified as photons.

We want to have independent method of photon purity estimation since it directly affects the cross section ratio.

The photon purity can be found using the fact that signal photons are isolated, while misidentified and non-prompt photons appear near or within a jet. We use super cluster removed charged hadron isolation as discriminating variable. Then the template shape fit is performed to find the fraction of prompt photons.

Template shapes for signal and background are determined from data using a method similar to one described in[46].

Signal shape is found by random cone isolation method where the same events that passed photon selection are used to find sum of PF candidates transverse energy in the same isolation cone with same η but random ϕ direction away from jets and leptons. This gives isolation of "completely isolated" photon. See data – simulation comparison of the photon variables on Figure 6.5.

Background shape is found by taking side-band in photon $\sigma_{i\eta i\eta}$ in the range 0.012 to 0.016. These events have non-prompt and misidentified photons.

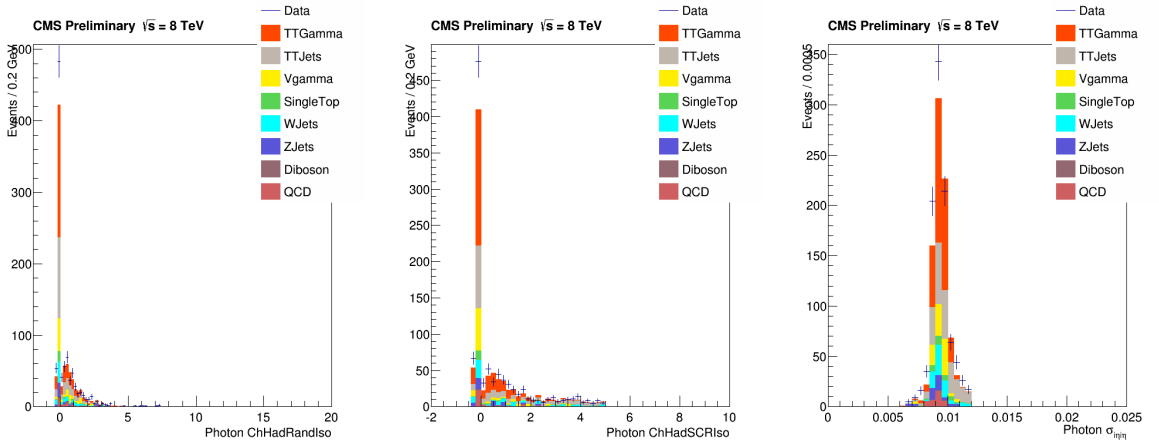

Figure 6.5: *Data and simulation comparison for photon variables: random cone isolation, SCFR charged hadron isolation, $\sigma_{i\eta i\eta}$.*

The fit is done in the range of SCFR charged hadron isolation up to 20 GeV/c, to ensure good description of background. We use loose photon candidates for fitting, applying nominal $\sigma_{i\eta i\eta}$ cut. In this way they have same properties as photons except the isolation

distribution extends farther. After fit is done the result can be corrected by applying nominal isolation cut on the scaled templates and finding fraction of photons and electrons.

As a cross-check of this method we applied it to the mixture of simulated samples mixed according to theoretical cross sections and integrated luminosity of the data used in the analysis. The $t\bar{t}$ has photons coming from jet misidentification, and $t\bar{t}\gamma$ has majority of prompt photons. Other simulated samples are added to have more statistics and better mixture of different kinds of events. We found that side-band region has only about 1% of prompt photons. This simulation sample is also used for a cross-check of the method in general. We perform data-driven technique of photon purity estimation on it and also cross-check if the signal photon purity agrees with generator information. Signal and background isolation templates can also be compared: Figure 6.6 and Figure 6.7

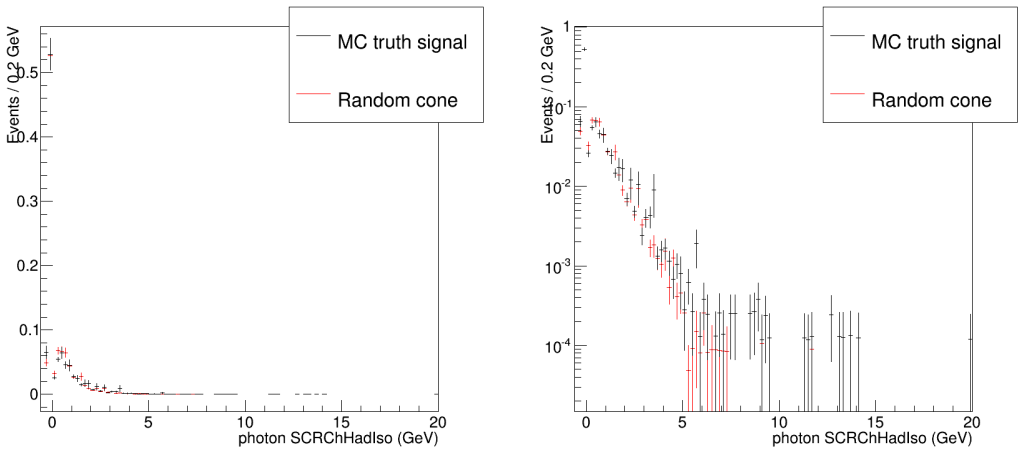

Figure 6.6: *Signal shape from random cone isolation and prompt photons by generator particle matching, all from simulation. Linear scale (left) and log scale (right)*

After the fit is done (Figure 6.8 we get photon purity equal to 0.571 ± 0.031 (includes statistical uncertainty for pseudo data only). By using generator level information the prompt photon and isolated electron fraction in this test sample is 0.585.

This is a very important result. It shows that selection of signal photons and electrons misidentified as photons on generator level is consistent with estimation using data-driven

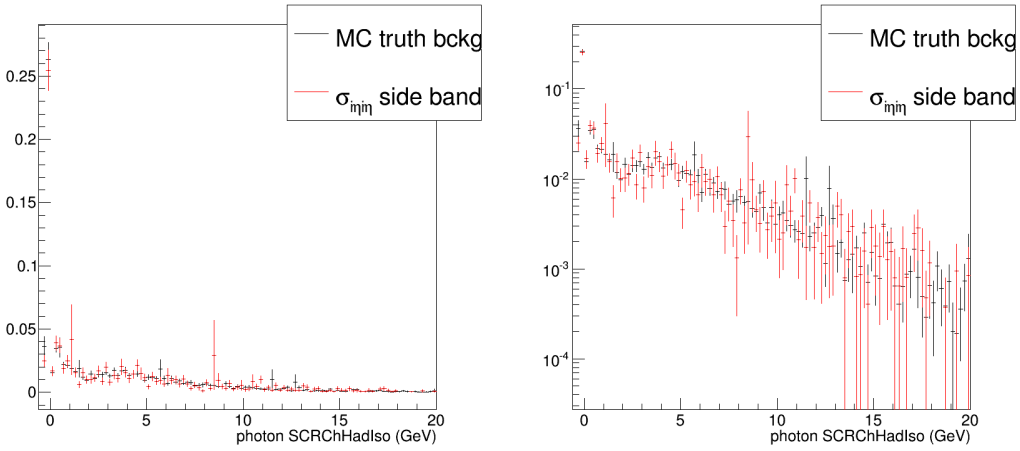

Figure 6.7: *Comparison of isolation profiles of sideband region and isolation of non-prompt and mis-identified photons by generator particle matching. Linear scale (left) and log scale (right).*

isolation shapes.

We have to consider several sides of the problem here. The idea is to make a signal photon selection on generator level that will:

- Select as much signal photons from TTGamma sample as possible
- Make isolation profile of selected signal photons agree with random cone isolation (data-driven signal shape) by filtering out non-isolated or badly reconstructed photons
- Do not reject isolated photons since they would contaminate background isolation profile (whatever is not selected as signal goes to background)

On the other hand the data-driven background sample ($\sigma_{i\eta i\eta}$ sideband region) should contain as few signal photons as possible and still preserve the shape of background photon isolation of the nominal $\sigma_{i\eta i\eta}$ region.

On the event selection level it is important to avoid contamination from nearby jets. They may be not a part of the photon and have no relation to it but contaminate isolation and make it differ from random cone isolation shape.

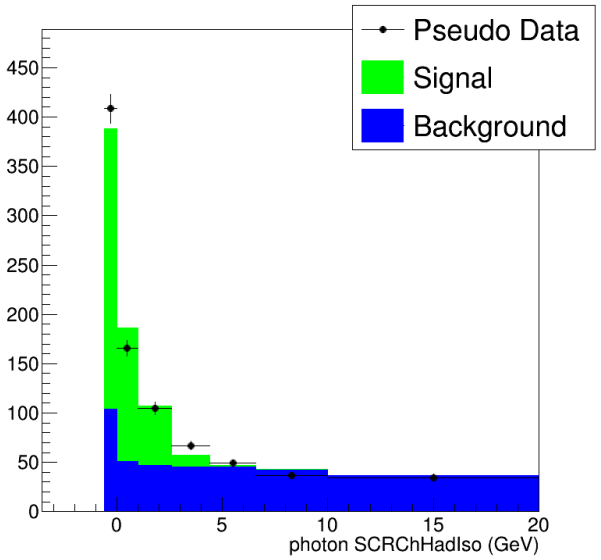

Figure 6.8: *Fit of charged hadron isolation profile in pseudo data for closure test.*

After we successfully tested this method on the simulation we apply it on data, see Figure 6.9. Statistical uncertainty given by RooFit does not include variation of the templates but only uncertainty of data. However since templates are data-driven and suffer from low statistics we have to take this uncertainty into account. We perform pseudo experiments where for data, signal and background templates in each bin the value is randomly chosen by Poisson distribution with mean of the bin content in the original histogram. Width of the fitted purity distribution obtained in this way is taken as statistical error. This gives photon purity in data $0.489 \pm 0.06(stat)$ (with isolation less than 20 GeV). The uncertainty is of a statistical nature but will be considered as systematic as it is related to measurement procedure.

To get maximum information from the signal and background shapes we fit isolation profile in the range up to 20 GeV. Photons used in the cross section calculation have charged hadron isolation cut at 5 GeV, so this cut is applied to the templates after the fit is done. Photon purity in the sample with 5 GeV cut on isolation is $0.564 \pm 0.063(stat)$. Next step is to estimate contamination from electrons.

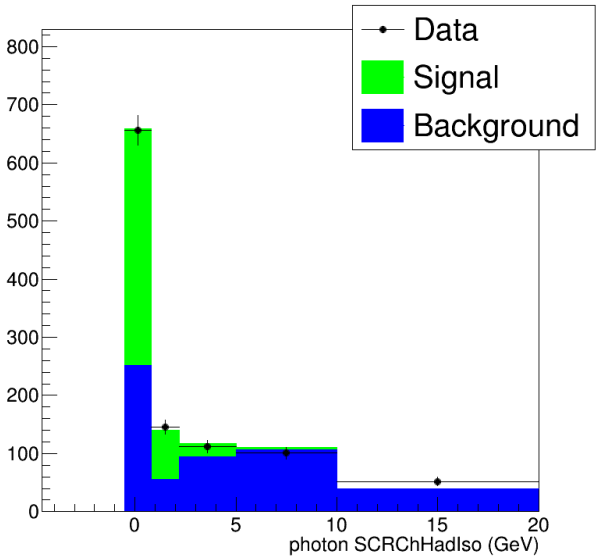

Figure 6.9: *Fit of charged hadron isolation profile in data.*

6.5 Estimation of electrons misidentified as photons

To estimate contamination from electrons misidentified as photons we will use simulation with correction derived from $Z \to ee$ process in data.

In simulated events we use matching procedure to identify electrons that are reconstructed as photons. As described in the Section 6.4 matching of reconstructed photons to generated ones is done by ΔR and $|p_T^{reco} - p_T^{gen}|/p_T^{gen}$. If no such photons are found the same procedure with the same cuts is performed to find generated electron matching the reconstructed photon.

After the matching simulated events can be categorized by the reconstructed photon origin:

- Signal: matched to genuine photon
- Misidentified electron: matched to electron
- Misidentified jet: not matched to genuine photon or electron

After the procedure of isolation template fit we have estimation of sum of the first two compared to the third.

In order to cross-check the results given by simulation we can use the "standard candle" for electrons: $Z \to ee$ process. To use it we will plot invariant mass of electron and photon. If the photon was misidentified electron from Z decay we will see a peak near 90 GeV.

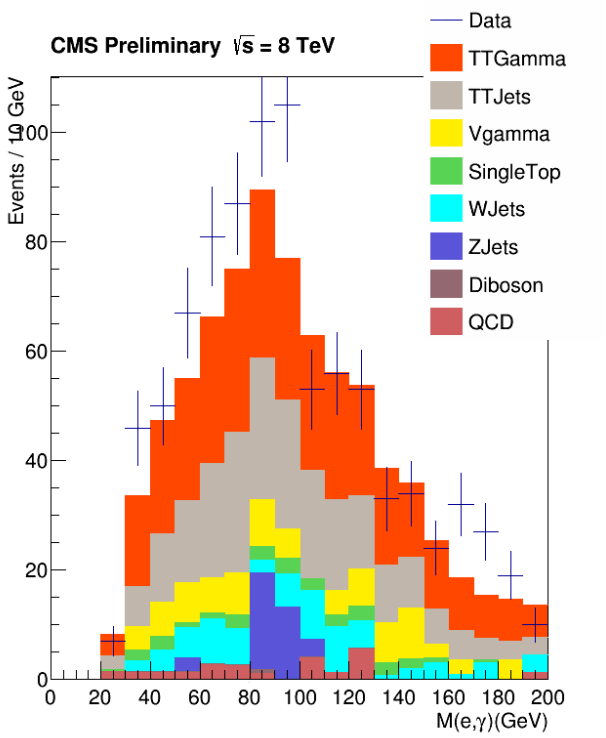

Figure 6.10: *Invariant mass of electron and photon, nominal selection.*

As can be seen on Figure 6.10 Contribution from $Z \to ee$ (ZJets process) is highly suppressed by event selection and does not give enough statistics to make any conclusion. However we see no drastic disagreement between data and simulation.

In order to improve the situation event selection should be changed to allow more Z events to pass. However as chances of misidentifying electron as photon may depend on environment (number of tracks in the event for example) it is better to make this change as slight as possible. By relaxing the requirement of having a b-tagged jet in the event and

keeping all other cuts the same $Z \to ee$ contribution is enhanced, see Figure 6.11. All steps for multijet estimation and M3 fit are repeated for this new selection.

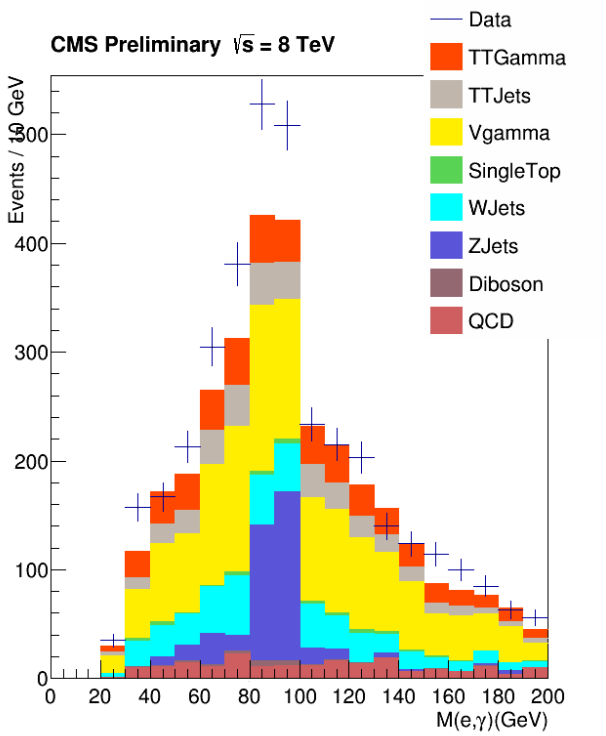

Figure 6.11: *Invariant mass of electron and photon, selection without b-tag requirement.*

Since Z peak is much more pronounced the template fit can be performed to estimate number of electrons from Z decay passing this modified event selection and compare its number to expectation from simulated events.

Two templates will be used for the fit, both of them derived from simulation.

- Electron template: combination of ZJets and Vgamma events where reconstructed photon is matched to generated electron. Makes narrow resonance shape of electron and photon invariant mass.

- Background template: all simulated and data-driven multijet samples excluding the previous category.

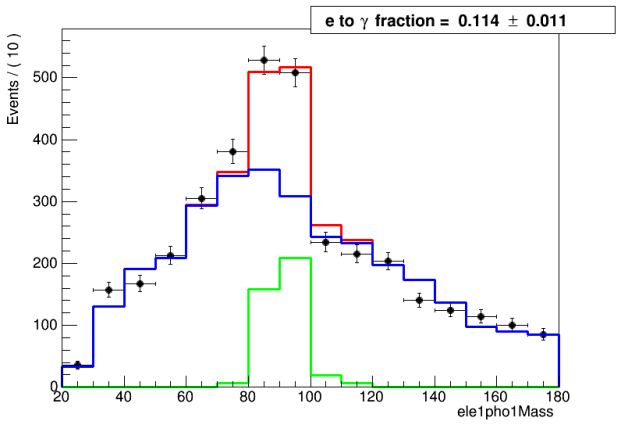

Figure 6.12: *Template fit of the invariant mass of electron and photon, selection without b-tag requirement. Green: Electron template, Blue: Background template, Red: weighted sum, Black: data.*

Result of the template fit is shown on Figure 6.12. Fit gives the fraction of electrons from Z decay reconstructed as photons. To get the expected number of such events integral of invariant mass distribution in data in the fit range should be multiplied by this fraction. Comparing this number with expected number of events in simulation we get the scale factor of $1.46 \pm 0.19(fit + stat)$.

This calculation is based on the assumption that simulated samples of ZJets and Zgamma (minor contribution to Z peak) being normalized to their cross sections and integrated luminosity describe the normalization of such events in data. This is not difficult to cross-check. If we modify the event selection to require two electrons instead of one electron and a photon we will have a lot statistics since in this case we will be looking at real (not misidentified) electrons. All other event selection details remain the same (missing transverse energy, number of jets, etc.) In that case we will be able to repeat the template fit with invariant mass of two electrons to verify the normalization of simulated events. We do it for both cases: with and without b-tag cut. Here we do not have to take into account multijet contribution because $Z \to ee$ cross section is relatively big and events from multijet processes are not expected to have peak in ee invariant mass.

With all samples normalized to cross sections and integrated luminosity, di-electron invariant mass with and without b-tag requirement are shown on Figure 6.13. Special care was taken to account for electron trigger and identification scale factors. Since event with two electrons has higher chances to fire a single electron trigger the trigger scale factor for the event is calculated in this way:

$$SF_{trigger} = 1 - (1 - SF^{ele1}_{trigger})(1 - SF^{ele2}_{trigger})$$

Identification scale factor is taken as a product of the scale factors for both electrons (they should be both independently identified)

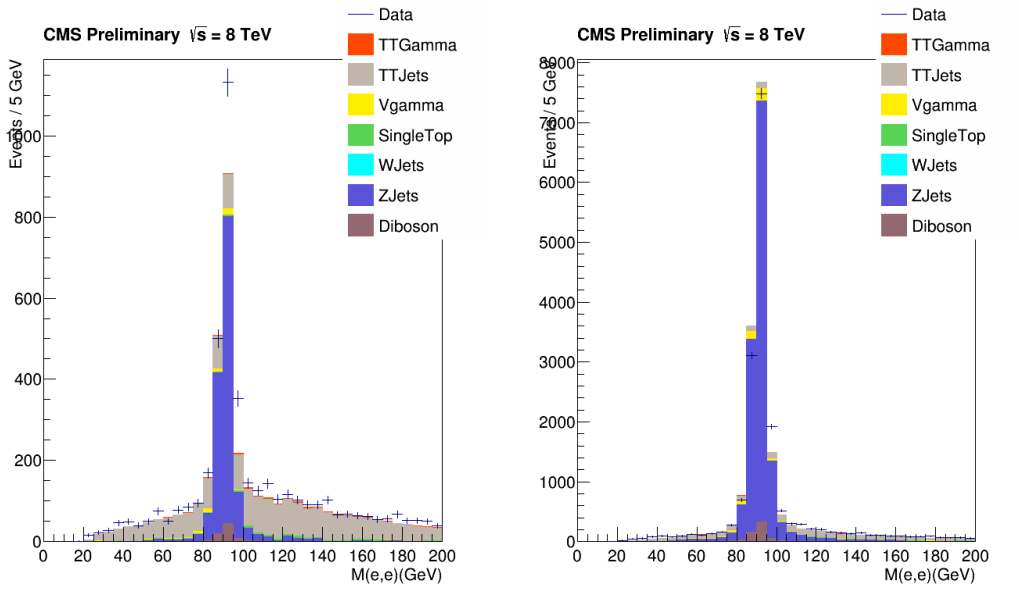

Figure 6.13: *Invariant mass of two electrons, with (left) and without (right) b-tag requirement.*

Templates for the fits are taken from simulation. ZJets and Vgamma are taken as $Z \to ee$ component, and everything else is "background". The small bump in the background shape is coming from di-boson samples. Even without fitting, di-electron mass distribution suggests that ZJets normalization with b-tag requirement is a bit off. The fits are shown on

Figure 6.14.

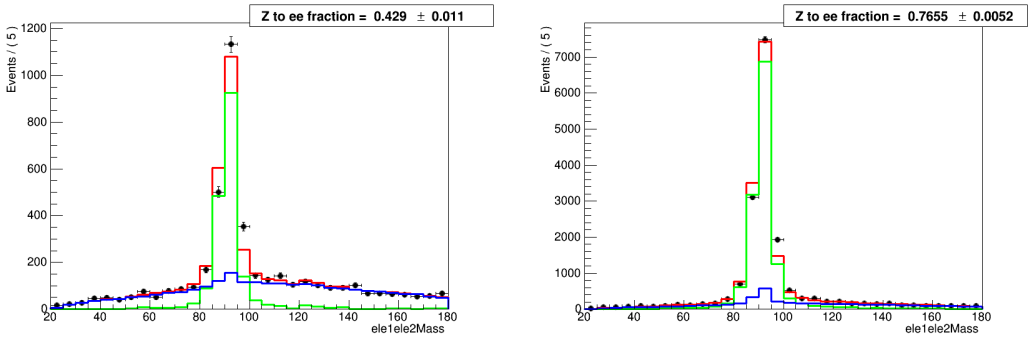

Figure 6.14: *Template fits for Z contribution, with (left) and without (right) b-tag requirement. Green: Z to ee component, Blue: everything else, Red: weighted sum, Black: data.*

Doing similar calculation to estimate correction necessary to apply to simulated Z events normalization we get the following.

- With b-tag requirement: simulated samples should be scaled by $1.20 \pm 0.06(stat + fit)$. This should correct normalization of ZJets sample in nominal selection. ZJets sample was normalized according to the CMS measurement, but presence of b-jet may be the reason for correction.

- Without b-tag requirement: simulated samples should be scaled by $0.95 \pm 0.015(stat + fit)$. This shows that for the particular region of phase space used for electron + photon invariant mass fit we have to apply extra correction to Z normalization.

This means that the estimate of Z events has to be corrected when we calculate electron to photon misidentification rate correction. After taking this into account we have to multiply number of electrons misidentified as photons in simulation by $1.5 \pm 0.2(fit + stat)$ to get their estimate in data. This number obtained by dividing 1.46 by 0.95. This correction is confirmed by other studies.

The following Table 6.1 summarizes simulated samples by origin of the reconstructed photon. Simulated samples are normalized by cross sections with correction from M3 fit

and di-electron fit where applicable. Misidentified Electron counts are not multiplied by the scale factor. Data driven multijet sample is not expected to have signal photons or electrons. TTGamma sample normalized by theoretical estimate of cross section. All errors are statistical. To get estimated number of background events from misidentified electrons the third column must be multiplied by the scale factor obtained as explained above.

Table 6.1: *Simulated samples categorized by reconstructed photon origin, after nominal selection.*

Sample	Total Events	Misidentified Jet	Misidentified Electron	Signal Photon
TTGamma	358 ± 5	10 ± 1	0.2 ± 0.1	348 ± 5
TTJets	234 ± 4	217 ± 4	18 ± 1	0
Vgamma	96 ± 13	3 ± 2	1 ± 0.7	92 ± 13
WJets	75 ± 10	67 ± 10	0	8 ± 3
ZJets	39 ± 10	11 ± 4	28 ± 7	0
Multijet	35 ± 8	35 ± 8	0	0
SingleTop	29 ± 4	16 ± 3	2.1 ± 1	12 ± 2

6.6 Number of signal events in data

To reach the goal of this study we need to answer the question: given the amount of observed events in data, how many of them are signal events? Table 6.1 can help to understand the difficulty. Reconstructed photon can come from signal photon, isolated electron or jet. Different simulated samples contribute to these categories in different ways.

As was shown in Section 6.5 misidentified electron counts should be corrected by applying the scale factor as this misidentification rate is not modeled perfectly.

Similar consideration should be applied to jets misidentified as photons. For this case we should have another (so far unknown) misidentification scale factor.

Normalization of the TTJets, WJets, ZJets and Multijet(QCD) samples is cross-checked and corrected by different template fits to data. This leaves two major contributing sources: TTGamma, which is unknown by the setting, and Vgamma. The latter (combination of

Zgamma and Wgamma) is normalized to NLO theoretical cross section and is known to be potentially mis-modeled to the order of 20%[47]. Vgamma has the highest contribution to signal photon category among non-ttbar samples. That is why it is important to cross-check its normalization before taking into account.

Single top samples are normalized to theoretical cross sections, but do not contribute in a significant way to the signal photon count and known to be relatively well modeled.

This leaves us with 3 unknowns:

- TTGamma scale factor (or normalization, the main unknown)
- Vgamma scale factor
- Jet to photon misidentification scale factor

From the collected events in data we extract the following:

- Photon purity (also includes misidentified electrons) from isolation data-driven template fit
- TTbar purity from M3 fit after photon selection
- Total number of selected events in data

Every quantity is measured with uncertainty, so the problem can't be solved just as 3 equations with 3 unknowns. Moreover, it is necessary to propagate the uncertainties to the unknown quantities.

We can relate knowns to unknowns using the information from simulated samples and build the likelihood function. $\mathcal{L}(t\bar{t}\gamma SF, V\gamma SF, jet \to \gamma SF) = e^{-\chi^2/2}$ where χ^2 is the sum of three terms:

$$\chi^2(tt\gamma SF, V\gamma SF, jet \to \gamma SF) = \frac{(\pi_{e\gamma}^{data} - \pi_{e\gamma}^{MC})^2}{\sigma_{\pi_{e\gamma}}^2} + \frac{(\pi_{t\bar{t}^{data}} - \pi_{t\bar{t}}^{MC})^2}{\sigma_{\pi_{t\bar{t}}}^2} + \frac{(N_{events}^{data} - N_{events}^{MC})^2}{\sigma_{N_{events}}^2}$$

Three unknown scale factors are used correct relevant contributions in the simulated event counts to calculate these quantities: $\pi_{e\gamma}^{MC}$ ratio of events with reconstructed photon matched to isolated electron or photon over all events in simulation, $\pi_{t\bar{t}}^{MC}$ ratio of events with pair of top quarks over all events in simulation, N_{events}^{MC} number of events in simulation.

Then scan over possible values of the parameters is made to find the combination with the best likelihood. For every parameter one-dimensional projection of likelihood is calculated by integration over other two parameters. The ranges of integration are chosen to cover all relevant space where likelihood is non-zero except for $V\gamma SF$. In this case we use prior knowledge that it can't be negative and not expected to be greater than 3.

Figure 6.15 shows three projections of likelihood. The width of the distribution is a conservative estimate of the uncertainty on this parameter.

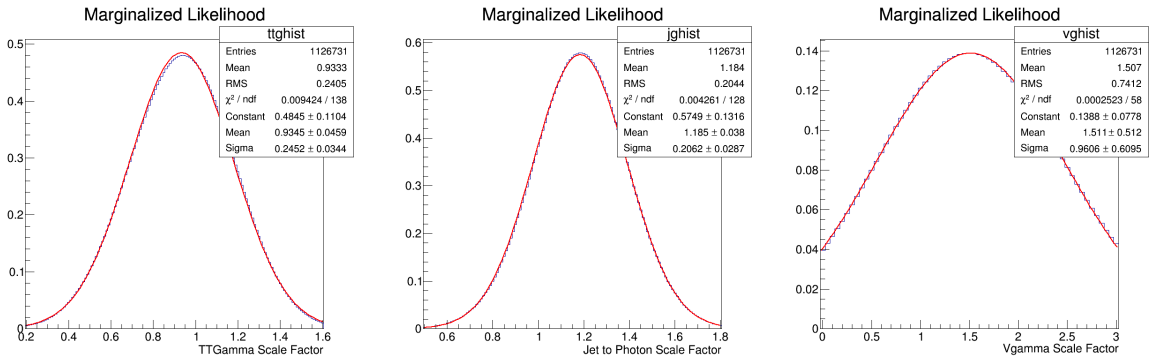

Figure 6.15: *Likelihood distribution for $t\bar{t}\gamma SF$, $jet \to \gamma SF$, $V\gamma SF$.*

The best agreement is achieved for these values of parameters, with errors from likelihood width: $t\bar{t}\gamma SF = 0.95 \pm 0.25$, $V\gamma SF = 1.35 \pm 1$, $jet \to \gamma SF = 1.19 \pm 0.20$. It brings simulated photon purity, top fraction and number of events to excellent agreement with quantities measured in data.

6.7 Signal acceptance calculation

Acceptance calculation for this analysis differs from usual inclusive cross section measurements because we measure ratio of cross sections. The event selection is chosen to make use of this fact: two steps (top selection and photon selection) are done sequentially.

So if we consider inclusive $t\bar{t}$ process we start with number of generated events (within some fiducial phase space) and count how many events are left after top event selection. We have 3 separate datasets with 0, 1, 2 leptons in the final state, they were scaled according their branching fractions. This gives acceptance of inclusive $t\bar{t}$ process: $\epsilon_{top}^{t\bar{t}} \cdot A_{top}^{t\bar{t}} = 0.03404 \pm 0.00002(stat)$ with negligible statistical error.

The same can be done for $t\bar{t}\gamma$ (signal sample). To get acceptance times efficiency we have to divide number of events passing top selection by total number of events considered. However in this case we have to choose what we take as denominator. The signal $t\bar{t}\gamma$ sample is inclusive, but theoretical calculations for cross section are done for final states with 1 and 2 leptons[30]. On Figure 6.16 we can see that 0 lepton (all hadronic channel) is not surviving event selection. To make comparison with theory easier we consider fiducial space for signal when 1 or 2 leptons are present (they will be mostly electrons, but other combinations are possible). Acceptance times efficiency of top selection for signal sample with 1 or 2 leptons: $\epsilon_{top}^{t\bar{t}\gamma} \cdot A_{top}^{t\bar{t}\gamma} = 0.0632 \pm 0.0003(stat)$.

The calculation of signal acceptance explained above is done for the sake of comparison with theoretical prediction. The biggest difference in the generated phase space and analysis selection is transverse energy cut on photon. (13 GeV in generated sample and 25 GeV in analysis). So effectively we measure the tail of this distribution and propagate it to the generated phase space using $\epsilon_{top} \cdot A_{top}$. In order to avoid this propagation of the result into larger phase space we also quote the *visible cross section ratio* where photon transverse energy cut is 25 GeV and $|\eta| < 1.4442$. $\epsilon_{top}^{t\bar{t}\gamma Vis} \cdot A_{top}^{t\bar{t}\gamma Vis} = 0.1618 \pm 0.0009(stat)$

Acceptance and efficiency of signal sample includes term for photon selection. For cross section ratio this term is a ratio of number of events passing top and photon selection and

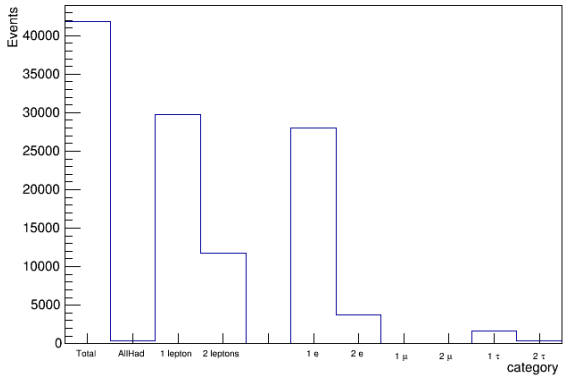

Figure 6.16: *Classification of TTGamma signal events passing top selection.*

having reconstructed photon matched to generator photon over number of events passing top selection. $\epsilon_{\gamma}^{t\bar{t}\gamma} A_{\gamma}^{t\bar{t}\gamma} = 0.1329 \pm 0.0025(stat)$

For visible cross section ratio photon selection term includes only photon reconstruction efficiency because by definition we are considering events where generator photon passes analysis level p_T and η cuts. Efficiency is calculated as ratio of events passing top and photon selection with reconstructed photon matched to generator photon over number of events passing top selection and with isolated generator photon passing p_T and η cuts for photon selection. $\epsilon_{\gamma}^{t\bar{t}\gamma Vis} = 0.281 \pm 0.004(stat)$

6.8 Systematic uncertainties

The measurement relies on simulation in several aspects.

Signal acceptance is pure simulation based quantity. Its uncertainty can be estimated by comparing different event generators and hadronizers as in [32]. In that analysis the main source of uncertainty was due to disagreement in the photon isolation templates and MC truth matched templates which is not the case here.

Background normalization is important part of the measurement. We use theoretical cross sections for di-boson and single top samples, but cross-check normalization of other

samples using different template fits to data.

These uncertainties are t as independent and their effect on the ratio measurement is summed up in one row of the table 6.2.

Estimation of systematic uncertainties related to modeling is done in the following way. All quantities that are known to have some uncertainty are varied and the measurement is repeated. The cross section ratio is then compared to nominal. In this way we can estimate the effect it has on the final result. In the table 6.2 we list the uncertainties in decreasing order of their effect on the cross section ratio.

Method is uncertainty on calculated number of signal events. It includes photon purity uncertainty, uncertainty on top purity after photon selection and statistical uncertainty on number of events in data. By changing each of these parameters up and down within their uncertainties we can get an idea of how they contribute to the total uncertainty on number if signal events (26%). Top purity uncertainty gives the biggest contribution (24%). Photon purity uncertainty translates to 13% of change of the signal normalization. Statistical uncertainty on number of events in data gives variation of 4%. It means the uncertainty of this measurement is systematics dominated, mainly because of the uncertainty on Wgamma and Zgamma processes (the main reason to perform M3 fit after photon selection).

JEC jet energy correction (scale) uncertainty originates from parton energy reconstruction. Energy of the original parton that produced jet can't be precisely reconstructed due to different detector response to various types of particles in the shower and non-linearities in the detector. Simulation-based procedure was developed by JetMET group to calibrate the response to various energies of incoming particles and regions of the detector. As a result one has a calibration constants with uncertainties. We vary all the calibration constants by increasing or decreasing them by the uncertainty.

JER jet energy resolution uncertainty.

Multijet uncertainty is estimated by varying data-driven multijet sample normalization by factor of 2 and 0.5 after the fit results. M3 fit is then done using this modified multijet

template.

TopPt re-weighting affects M3 shape and consequently fit results for $t\bar{t}$ and WJets normalization. Step up and down are done according to recommendation of CMS top group. Step down is not applying re-weighting at all, step up is applying it twice.

Btag B-tagging uncertainty is done according to BTV group recommendations. The scale factors and uncertainties are taken from dilepton $t\bar{t}$ events and should not be correlated with the data used in this analysis.

Photon Energy Photon Energy uncertainty mainly contributes to the acceptance. We scale energy of all reconstructed photons by 1% up and down and redo the measurement.

PU Pile Up correction is applied to all MC events and derived for average number of interactions in data. By scaling average number of pile up interactions by 5% up and down we get different event weights.

Table 6.2 lists the uncertainties and their effect on measured cross section ratio and visible cross section ratio. The latter has no signal acceptance term, so that the cross section is defined in the analysis cuts phase space, not in generator particles phase space.

Table 6.2: *Systematic uncertainties and their contribution to the cross section ratio.*

Source	Ratio change (%)	Visible Ratio change (%)
Method	26	26
JER	5.3	5.4
JEC	4.2	5.0
Multijet	3.2	3.2
Electron FakeSF	2.1	2.1
TopPt	2.1	2.1
Electron Eff	1.1	1.7
Btag	1.1	1.6
Photon Energy	1.0	1.1
PU	1.1	1.1
Electron Energy	1.1	1.1
ZJetsSF	1.1	1.1
Total	27.3	27.6

6.9 Plots

In order to check all procedures related to $t\bar{t}$ and WJets samples are scaling and multijet estimation we plot some important quantities after the events passed photon selection. Here we use theoretical cross section for $t\bar{t}+\gamma$ process. However this dataset is used only for acceptance and photon reconstruction and identification efficiency cross-check and the normalization is not used in the measurement.

6.10 Results

Cross section ratio can be calculated using the following values:

- Number of events in data after photon selection: $977 \pm 31(stat)$
- Electromagnetic object purity after photon selection: $0.564 \pm 0.063(stat + fit)$, see Section 6.4.
- Fraction of events containing top quark pair, from M3 fit after photon selection: $0.645 \pm 0.084(fit)$, see Section 6.3.
- Expected number of signal events with genuine photon after photon selection, from simulation, data-driven photon purity fit, and M3 fit: $N_{signal} = 330 \pm 85$. See Section 6.6 for details. The number is obtained by multiplying the event count form TTGamma sample with signal photon from Table 6.1 by the scale factor $t\bar{t}\gamma SF$ from Section 6.6.
- Expected number of $t\bar{t}$ events in data after top selection: $162560 \pm 1240(stat + fit)$, obtained from M3 fit after top selection, details in Section 6.3.
- Signal acceptance times photon reconstruction and identification efficiency found from signal simulation. (Section 6.7) Calculated as ratio of number of events in signal sample passing photon selection with reconstructed photon matched to generator photon over

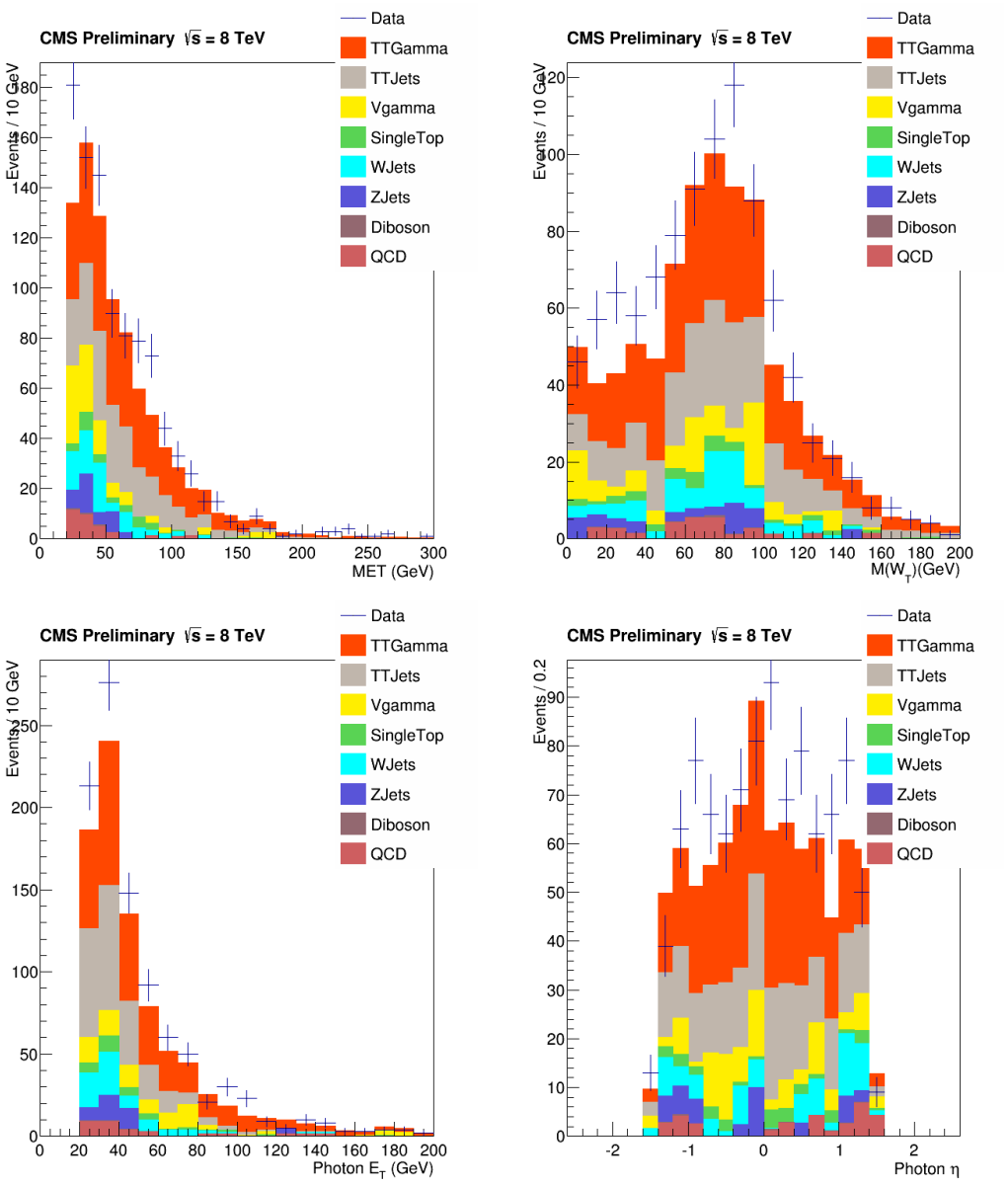

Figure 6.17: *Comparison of data and simulation after photon selection. Theoretical cross section for $t\bar{t}+\gamma$ is used. No re-weighting for TTGamma scale factor, Vgamma scale factor, jet to photon scale factor is done.*

number of events passing only top selection (1 or 2 lepton final state). This gives $\epsilon_{\gamma}^{t\bar{t}\gamma} A_{\gamma}^{t\bar{t}\gamma} = 0.1329 \pm 0.0025(stat)$

We get the cross section ratio:

$$R = \frac{\sigma_{t\bar{t}+\gamma}}{\sigma_{t\bar{t}}} = \frac{N_{signal}}{\epsilon_{top}^{t\bar{t}\gamma} A_{top}^{t\bar{t}\gamma} \epsilon_{\gamma}^{t\bar{t}\gamma} A_{\gamma}^{t\bar{t}\gamma}} \cdot \frac{\epsilon_{top}^{t\bar{t}} A_{top}^{t\bar{t}}}{N^{t\bar{t}}} =$$

$$= 0.0082 \pm 0.0022(stat + syst).$$

Note that $\sigma_{t\bar{t}+\gamma}$ here is not inclusive cross section, but within a phase space defined in Section 4.1 with cuts on particles transverse momenta and other quantities. Also, we consider only semi-leptonic and di-leptonic top pair decays, while lepton flavor can be arbitrary.

Another result is *visible cross section ratio* where we do not extrapolate the measured result to the phase space used for signal simulation. In this way we do not rely on kinematic properties of simulated signal dataset. We will need photon reconstruction and identification efficiency, but not signal acceptance for photon selection. Efficiency is calculated as ratio of number of generated signal events that passed photon selection to number of generated signal events with a generator level signal photon in the region of $p_T\eta$ space used for photon selection. (Same Pt cut and eta cut on generator photon as used for reconstructed photon)

$$R_{Vis} = \frac{N_{signal}}{\epsilon_{top}^{t\bar{t}\gamma Vis} A_{top}^{t\bar{t}\gamma Vis} \epsilon_{\gamma}^{t\bar{t}\gamma Vis}} \cdot \frac{\epsilon_{top}^{t\bar{t}} A_{top}^{t\bar{t}}}{N^{t\bar{t}}} = 0.00152 \pm 0.00041(stat + syst).$$

By using the CMS measurement of inclusive $t\bar{t}$ cross section[17] we can convert ratio to $t\bar{t} + \gamma$ cross section:

$$\sigma_{t\bar{t}+\gamma} = R \cdot \sigma_{t\bar{t}} =$$

$$(0.0082 \pm 0.0022(stat + syst)) \cdot (239 \pm 2(stat.) \pm 11(syst.) \pm 6(lum.)pb) =$$

$$1.96 \pm 0.53pb$$

All uncertainties were added in quadrature.

Visible cross section calculated in the same manner:

$$\sigma_{t\bar{t}+\gamma Vis} = R_{Vis} \cdot \sigma_{t\bar{t}} =$$

$$(0.00152 \pm 0.00041(stat + syst)) \cdot (239 \pm 2(stat.) \pm 11(syst.) \pm 6(lum.)pb) =$$

$$0.36 \pm 0.10pb$$

This result can be compared with theoretical NLO calculation[30]. For di-lepton channel the cross section is $33fb$, for lepton+jets channel it is $148fb$. This is for one lepton species. We calculated acceptance for all lepton flavors and charge combinations. For di-lepton channel there are 3 lepton options for top quark and 3 for top anti-quark, which makes 9 combinations. For semi-leptonic channel one of the top quarks can decay leptonically (3 options), while other decays hadronically (2 options, because W can decay into first and second generation quarks). The situation does not change if the top quark and antiquark are swapped, this adds factor of 2, making total 12 different combinations. Total cross section for 1 and 2 lepton final state is:

$$\sigma_{t\bar{t}\gamma}^{NLO} = 9 \cdot 0.033pb + 12 \cdot 0.148pb = 2.07 \pm 0.41(scale)pb$$

We see that experimental and theoretical results agree very well within their precision.

Chapter 7

Summary and Outlook

The analysis presented in this thesis complements a similar measurement by the CMS experiment in the muon+jets channel[32] and by the ATLAS experiment in both the electron+jets and muon+jets channels[33]. No significant deviations from Standard Model predictions were observed.

The development of analysis methodology is the main result of this work.

Kinematic properties of events are explored through the M3 variable – invariant mass of three jets that have highest transverse momentum when added together. The profile of M3 distribution in data is fitted with simulation-based shapes to find the fraction of events containing a top quark pair, called top purity for compactness. The method is cross-checked in electron+jets event selection where top pair events are dominant and their amount is known. Then the method is applied in electron+photon+jets event selection.

Photon identification efficiency is estimated by studying photon isolation profile, inspired by[48], but the method was modified for the different environment. The modified method allows extraction of signal and background isolation templates directly from data. The procedure is tested on simulated events and then data-driven shapes are used in the measurement.

Decay of Z boson to e^+e^- process was used to correct the description of electrons misiden-

tified as photons in simulation.

A simultaneous fit for 3 parameters including number of top pair and photon events is done by combining information obtained from the M3 fit and photon isolation fit.

The main sources of uncertainty of this measurement are photon purity estimation and top purity estimation. Precision of the method is currently limited by the number of events passing the selection. With more data available at LHC Run 2 it will be possible to improve the precision.

At this point the uncertainty of the measurement is slightly bigger than the theoretical uncertainty. By improving the result with more data and better analysis procedure one will be able to put theoretical calculations at more stringent test.

Other analyses like the search for production of top pair and Higgs boson, which decays in two photons, has a similar experimental signature, with two photons instead of one. Understanding top pair and photon process will be of a high importance because it will be a background for this search.

Bibliography

[1] MissMJ. Standard model of elementary particles. Own work by uploader, PBS NOVA, Fermilab, Office of Science, United States Department of Energy, Particle Data Group http://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg.

[2] D0 Collaboration. Useful diagrams of top signals and backgrounds. Public web page. URL http://www-d0.fnal.gov/Run2Physics/top/top_public_web_pages/top_feynman_diagrams.html.

[3] Christiane Lefevre. The CERN accelerator complex. Complexe des accelerateurs du CERN. https://cds.cern.ch/record/1260465/export/hx?ln=en, Dec 2008.

[4] CMS Collaboration. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical Design Report CMS. CERN, Geneva, 2006. There is an error on cover due to a technical problem for some items.

[5] CMS Collaboration. Alignment of the CMS silicon tracker during commissioning with cosmic rays. *Journal of Instrumentation*, 5(03):T03009, 2010. URL http://stacks.iop.org/1748-0221/5/i=03/a=T03009.

[6] Jun John Sakurai and Jim Napolitano. *Modern quantum mechanics*, volume 2nd ed. Addison-Wesley, 2011.

[7] Julian Schwinger. *Quantum kinematics and dynamics*. Westview Press, 2000.

[8] James D Bjorken and Sidney David Drell. *Relativistic quantum mechanics*. McGraw-Hill, 1964.

[9] Michael E Peskin and Daniel V Schroeder. *An introduction to quantum field theory*. Westview Press, 1995.

[10] Anthony Zee. *Quantum field theory in a nutshell*, volume 2nd ed. Princeton university press, 2010.

[11] Abe, F. and others. Observation of Top Quark Production in $\overline{p}p$ Collisions with the Collider Detector at Fermilab. *Phys. Rev. Lett.*, 74:2626–2631, Apr 1995. doi: {10.1103/PhysRevLett.74.2626}. URL http://link.aps.org/doi/10.1103/PhysRevLett.74.2626.

[12] Abachi, S. and others. Search for High Mass Top Quark Production in $p\overline{p}$ Collisions at $\sqrt{s} =$ 1.8 TeV. *Phys. Rev. Lett.*, 74:2422–2426, Mar 1995. doi: {10.1103/PhysRevLett.74.2422}. URL http://link.aps.org/doi/10.1103/PhysRevLett.74.2422.

[13] ATLAS Collaboration. Measurement of the Higgs boson mass from the H$\rightarrow\gamma\gamma$ and H$\rightarrow$ ZZ^* $\rightarrow 4\ell$ channels in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector. *Phys. Rev. D*, 90:052004, Sep 2014. doi: {10.1103/PhysRevD.90.052004}. URL http://link.aps.org/doi/10.1103/PhysRevD.90.052004.

[14] CMS Collaboration. Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV. Technical Report arXiv:1412.8662. CERN-PH-EP-2014-288. CMS-HIG-14-009, CERN, Geneva, Dec 2014. Comments: Submitted to Eur. Phys. J. C.

[15] Wendell, R. and others. Atmospheric neutrino oscillation analysis with subleading effects in Super-Kamiokande I, II, and III. *Phys. Rev. D*, 81:092004, May 2010. doi: {10.1103/PhysRevD.81.092004}. URL http://link.aps.org/doi/10.1103/PhysRevD.81.092004.

[16] ATLAS Collaboration, CDF Collaboration, CMS Collaboration, D0 Collaboration. First combination of Tevatron and LHC measurements of the top-quark mass. 2014.

[17] CMS Collaboration. Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 8$ TeV. *Journal of High Energy Physics*, 2014(2):24, 2014. doi: {10.1007/JHEP02(2014)024}. URL http://dx.doi.org/10.1007/JHEP02%282014%29024.

[18] CMS Collaboration. Measurement of the production cross section in pp collisions at $\sqrt{s} = 7$ tev with lepton + jets final states. *Physics Letters B*, 720(13):83 – 104, 2013. ISSN 0370-2693. doi: http://dx.doi.org/10.1016/j.physletb.2013.02.021. URL http://www.sciencedirect.com/science/article/pii/S0370269313001500.

[19] Robert Oerter. *The theory of almost everything: The standard model, the unsung triumph of modern physics*. Penguin, 2006.

[20] Englert, F. and Brout, R. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964. doi: {10.1103/PhysRevLett.13.321}. URL http://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[21] Higgs, Peter W. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964. doi: {10.1103/PhysRevLett.13.508}. URL http://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[22] Planck Collaboration. Planck 2013 results. I. Overview of products and scientific results. *A&A*, 571:A1, 2014. doi: 10.1051/0004-6361/201321529. URL http://dx.doi.org/10.1051/0004-6361/201321529.

[23] Rudolf Haag, Jan T. Lopuszanski, and Martin Sohnius. All possible generators of supersymmetries of the s-matrix. *Nuclear Physics B*, 88(2):257 – 274, 1975. ISSN 0550-3213. doi: http://dx.doi.org/10.1016/0550-3213(75)90279-5. URL http://www.sciencedirect.com/science/article/pii/0550321375902795.

[24] Nima ArkaniHamed, Savas Dimopoulos, and Gia Dvali. The hierarchy problem and new dimensions at a millimeter. *Physics Letters B*, 429(34):263 – 272, 1998. ISSN 0370-2693. doi: http://dx.doi.org/10.1016/S0370-2693(98)00466-3. URL http://www.sciencedirect.com/science/article/pii/S0370269398004663.

[25] Kobayashi, Makoto and Maskawa, Toshihide. CP-Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics*, 49(2):652–657, 1973. doi: {10.1143/PTP.49.652}. URL http://ptp.oxfordjournals.org/content/49/2/652.abstract.

[26] S. W. Herb, D. C. Hom, L. M. Lederman, J. C. Sens, H. D. Snyder, J. K. Yoh, J. A. Appel, B. C. Brown, C. N. Brown, W. R. Innes, K. Ueno, T. Yamanouchi, A. S. Ito, H. Jöstlein, D. M. Kaplan, and R. D. Kephart. Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions. *Phys. Rev. Lett.*, 39:252–255, Aug 1977. doi: {10.1103/PhysRevLett.39.252}. URL http://link.aps.org/doi/10.1103/PhysRevLett.39.252.

[27] Martin Goerner and Peter Schleper. *Differential Cross Sections for Top-Quark-Pair Production in the e/μ+Jets Final State at* $\sqrt{s} = 8$ *TeV in CMS*. PhD thesis, Hamburg U., 2014. URL https://cds.cern.ch/record/1754332.

[28] Beringer, J. and others. Review of Particle Physics. *Phys. Rev. D*, 86:010001, Jul 2012. doi: {10.1103/PhysRevD.86.010001}. URL http://link.aps.org/doi/10.1103/PhysRevD.86.010001.

[29] Jorgen D'Hondt, Steven Lowette, O L Buchmller, Susanna Cucciarelli, Frank-Peter Schilling, Maria Spiropulu, S Paktinat-Mehdiabadi, Daniele Benedetti, and Luc Pape. Fitting of Event Topologies with External Kinematic Constraints in CMS. Technical Report CMS-NOTE-2006-023, CERN, Geneva, Jan 2006. URL https://cds.cern.ch/record/926540.

[30] Melnikov, Kirill and Schulze, Markus and Scharf, Andreas. QCD corrections to top quark pair production in association with a photon at hadron colliders. *Phys. Rev. D*, 83:074013, Apr 2011. doi: {10.1103/PhysRevD.83.074013}. URL http://link.aps.org/doi/10.1103/PhysRevD.83.074013.

[31] U. Baur, A. Juste, L.H. Orr, and D. Rainwater. Probing electroweak top quark couplings at hadron colliders. *Phys.Rev.*, D71:054013, 2005. doi: 10.1103/PhysRevD.71.054013.

[32] CMS Collaboration. Measurement of the inclusive top-quark pair + photon production cross section in the muon + jets channel in pp collisions at 8 TeV. Technical Report CMS-PAS-TOP-13-011, CERN, Geneva, 2014.

[33] Georges Aad et al. Observation of top-quark pair production in association with a photon and measurement of the $t\bar{t}\gamma$ production cross section in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector. 2015.

[34] Jahred Adelman, Matthew Baumgart, Aran Garcia-Bellido, and Andrey Loginov. Determining Top Quark Couplings at the LHC: Snowmass White Paper. 2013.

[35] ALICE Collaboration. The ALICE experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08002, 2008. URL http://stacks.iop.org/1748-0221/3/i=08/a=S08002.

[36] LHCb Collaboration. The LHCb Detector at the LHC. *Journal of Instrumentation*, 3 (08):S08005, 2008. URL http://stacks.iop.org/1748-0221/3/i=08/a=S08005.

[37] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3(08):S08003, 2008. URL http://stacks.iop.org/1748-0221/3/i=08/a=S08003.

[38] CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004, 2008. URL http://stacks.iop.org/1748-0221/3/i=08/a=S08004.

[39] A. Airapetian et al. ATLAS: Detector and physics performance technical design report. Volume 1. 1999.

[40] CMS Collaboration. Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. Technical Report CMS-PAS-PFT-09-001, CERN, 2009. Geneva, Apr 2009.

[41] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl.Instrum.Meth.*, A506:250–303, 2003. doi: 10.1016/S0168-9002(03)01368-8.

[42] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 1407:079, 2014. doi: 10.1007/JHEP07(2014)079.

[43] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 0605:026, 2006. doi: 10.1088/1126-6708/2006/05/026.

[44] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 0711:070, 2007. doi: 10.1088/1126-6708/2007/11/070.

[45] Z. Was. Precision simulations with TAUOLA and PHOTOS. *Nucl.Phys.Proc.Suppl.*, 169:16–21, 2007. doi: 10.1016/j.nuclphysbps.2007.02.113.

[46] CMS Collaboration. Measurement of differential cross sections for the production of a pair of isolated photons in pp collisions at $\sqrt{s} = 7$ TeV. *Eur. Phys. J. C*, 74

(arXiv:1405.7225. CMS-SMP-13-001. CERN-PH-EP-2014-067):3129. 41 p, May 2014. Comments: Replaced with published version. Added journal reference and DOI.

[47] CMS Collaboration. Measurement of the $W\gamma$ and $Z\gamma$ inclusive cross sections in pp collisions at $\sqrt{s} = 7$ TeV and limits on anomalous triple gauge boson couplings. *Phys. Rev. D*, 89(arXiv:1308.6832. CMS-EWK-11-009. CERN-PH-EP-2013-108):092005. 45 p, Sep 2013. Comments: Replaced with published version. Added journal reference and DOI.

[48] CMS Collaboration. Measurement of differential cross sections for the production of a pair of isolated photons in pp collisions at $\sqrt{s} = 7$ TeV. *Eur. Phys. J. C*, 74 (arXiv:1405.7225. CMS-SMP-13-001. CERN-PH-EP-2014-067):3129. 41 p, May 2014. Comments: Replaced with published version. Added journal reference and DOI.