



UNIVERSITÀ DEGLI STUDI DI TRIESTE
Dipartimento di Fisica

Corso di Laurea Magistrale in FISICA
Curriculum in Microfisica e Struttura della Materia

STUDY OF THE LINEARITY OF THE
ELECTROMAGNETIC CALORIMETER OF THE
COMPACT MUON SOLENOID EXPERIMENT
USING J/ψ , Y AND Z RESONANCES

Author:

Federico Vazzoler

MATRICOLA: SM2300236

Supervisor:

Prof. Giuseppe Della Ricca

Co-Supervisor:

Dott. Emanuele Di Marco

Anno Accademico 2015/2016





UNIVERSITÀ DEGLI STUDI DI TRIESTE
Dipartimento di Fisica

Corso di Laurea Magistrale in FISICA
Curriculum in Microfisica e Struttura della Materia

STUDIO DELLA LINEARITÀ DEL CALORIMETRO
ELETTROMAGNETICO DELL'ESPERIMENTO COMPACT
MUON SOLENOID
UTILIZZANDO LE RISONANZE J/ψ , Y E Z

Candidato:

Federico Vazzoler

MATRICOLA: SM2300236

Relatore:

Prof. Giuseppe Della Ricca

Co-Relatore:

Dott. Emanuele Di Marco

Anno Accademico 2015/2016

Elementary particle physics is the quadrant of nature whose laws can be written in a few lines with absolute precision and the greatest empirical adequacy.

RICCARDO BARBIERI (1944)

Italian born, theoretical physicist and professor

Introduzione

Il Modello Standard delle particelle, introdotto agli inizi degli anni '50 del secolo, ha collezionato nel corso del tempo numerosi successi. Esso permette sia di spiegare un grande numero di risultati sperimentali sia di fornire predizioni teoriche, introducendo un esiguo numero di particelle. I costituenti essenziali della materia sono rappresentati da dodici fermioni, chiamati leptoni e quarks, mentre le forze tra i medesimi sono mediate da quattro bosoni, chiamati bosoni di gauge (γ , g , Z and W).

Sebbene il Modello Standard abbia ricevuto molte conferme nel corso degli anni, esso va ancora molto migliorato essendo molti e vari i fenomeni fisici non spiegati. Molte risposte alle domande ancora aperte possono essere trovate solo ad alte energie e luminosità: per questo motivo, la ricerca in fisica delle particelle viene portata avanti grazie all'aiuto di macchine acceleratrici sempre più grandi e potenti. Il Large Hadron Collider, un collisionatore protone-protone operativo dal 2008, è una di esse. Esso ospita quattro esperimenti, ATLAS, LHCb, ALICE e CMS, l'ultimo dei quali è un rivelatore ad ampio spettro, capace di studiare vari aspetti delle collisioni protoniche, rilevando sia particelle neutre che cariche. Il calorimetro elettromagnetico di CMS è ottimizzato per una precisa identificazione e ricostruzione di muoni, elettroni e fotoni. Uno tra gli obiettivi principali di CMS era quello di verificare il modello teorico, proposto negli anni '70 del secolo scorso da Peter Higgs e altri, capace di spiegare, nel contesto del Modello Standard, come i bosoni di gauge, i quarks ed i leptoni possano acquisire la propria massa. L'esistenza del bosone predetto dalla suddetta teoria è stata confermata nel 2012 quando una particella neutra con spin-0 e una massa approssimativamente di $125 \text{ GeV}/c^2$ è stata scoperta grazie agli sforzi congiunti delle collaborazioni ATLAS e CMS.

La misura della massa del bosone di Higgs risulta essere particolarmente importante: una volta conosciuta con precisione, tutte le proprietà del bosone, come ad esempio la sezione d'urto di produzione o le larghezze di decadimento, possono essere facilmente predette. Molti sforzi sono volti ad aumentare la precisione della misura della sua massa, contestualmente alla riduzione delle incertezze su quest'ultima. Il canale di decadimento in quattro leptoni, $H \rightarrow ZZ^* \rightarrow 4l$, è quello preferito nello studio del bosone di Higgs, principalmente grazie all'eccellente risoluzione nella misura dei momenti dei leptoni di decadimento e, soprattutto, per il ridotto numero di eventi di fondo presenti. L'unico svantaggio risiede nella ridotta statistica disponibile. L'upgrade Run 2 di LHC ha messo a disposizione un'energia nel centro di massa pari a $\sqrt{s} = 13 \text{ TeV}$ contestualmente ad un'alta luminosità, rispondendo alla necessità di aumentare la statistica disponibile e contribuendo quindi a ridurre l'errore statistico associato alla misura della massa. Per ottenere una buona misura della massa di una risonanza che decade in quattro leptoni,

è altresì fondamentale una precisa calibrazione della scala energetica e della risoluzione con cui vengono misurate le energie dei leptoni. L'obiettivo di questo lavoro è quello di studiare la scala energetica con cui è calibrato il calorimetro elettromagnetico di CMS, ECAL : particolare attenzione è stata posta a possibili effetti di non-linearità nella scala energetica, effetti che rimangono anche dopo la calibrazione del sotto-rivelatore e contribuiscono alle incertezze sull'energia ricostruita nel calorimetro. A tal scopo vengono studiati gli elettroni provenienti dal decadimento di risonanze note del Modello Standard, quali il bosone Z ed i mesoni Y e J/ψ . Le correzioni energetiche sono ottenute confrontando eventi provenienti da dati sperimentali e simulazioni e sono quindi applicate alle energie di elettroni derivanti dal decadimento di bosoni di Higgs simulati, fornendo una stima dell'incertezza sistematica associata con la misura della sua massa.

Il lavoro di tesi è diviso in quattro capitoli. Nel primo capitolo vengono introdotti il Modello Standard, le interazioni deboli e la teoria associata al campo di Higgs. I principali canali di produzione e decadimento insieme ai risultati noti riguardo la massa del bosone di Higgs vengono trattati brevemente. Nel secondo capitolo l'acceleratore LHC e l'esperimento CMS vengono descritti, focalizzando l'attenzione sulle caratteristiche peculiari del calorimetro elettromagnetico. Il ruolo di ECAL nell'identificazione e ricostruzione degli elettroni provenienti dal decadimento del bosone di Higgs in quattro leptoni è fondamentale. Nel terzo capitolo vengono presentati gli algoritmi di ricostruzione energetica e le calibrazioni utilizzate per correggere l'energia ricostruita degli elettroni. Nel quarto capitolo vengono discusse le correzioni aggiuntive alla scala energetica e gli studi sulla residua non linearità del calorimetro. La funzione di correzione all'energia degli elettroni viene applicata alla misura della massa del bosone di Higgs utilizzando eventi simulati, ricavandone contestualmente l'incertezza sistematica.

Contents

Introduction	1
1 The Standard Model of Particle Physics	3
1.1 Particle content	3
1.2 Weak interactions	6
1.2.1 Local gauge invariance	6
1.2.2 Formulation of the Electroweak Standard Model	8
1.3 Higgs field and spontaneous symmetry breaking	10
1.3.1 Higgs boson searches	13
1.3.2 State of the art	17
2 The LHC accelerator and the CMS detector	21
2.1 The European Organization for Nuclear Research	21
2.2 The Large Hadron Collider accelerator	21
2.2.1 LHC: acceleration complex	22
2.2.2 LHC: structure overview	22
2.3 The Compact Muon Solenoid experiment	25
2.3.1 Structure overview	25
2.3.2 Tracker	27
2.3.3 Electromagnetic calorimeter	28
2.3.4 Hadronic calorimeter	31
2.3.5 Superconducting solenoid	32
2.3.6 Muon chambers	32
2.3.7 Data triggering system	33
3 Electron reconstruction and energy calibration	35
3.1 Electron reconstruction in CMS	35
3.1.1 Clustering of the electron energy in ECAL	36
3.1.2 Electron track reconstruction	39
3.1.3 Electron classification	43
3.2 Energy calibrations and corrections	46
3.2.1 Correction for changes in response	48
3.2.2 Single channel intercalibration	49
3.2.3 Calibration of the preshower	52
3.2.4 Energy corrections	52
3.2.5 Absolute energy calibration	54

4	Fine tuning of calibration	57
4.1	Residual data/MC discrepancies	57
4.2	Fine-tuning of the energy scale with $Z \rightarrow e^+e^-$ events	59
4.2.1	Validation of the energy calibration and corrections	59
4.3	Fine-tuning of the simulated resolution	60
4.4	Energy scale and resolution uncertainty	61
4.4.1	Electron scale and resolution measurements as a function of category and pileup	61
4.4.2	Electron scale linearity measurement from Z , J/ψ and $Y \rightarrow ee$ decays	62
4.5	Data and simulation samples	63
4.6	Correction to ECAL linearity using known SM resonances	65
4.7	Effects of the electron energy scale non-linearity on the Higgs boson mass measurements	79
5	Summary and conclusions	91
A	Fits of the resonances invariant mass distribution for the 74X release	93
B	Fits of the resonances invariant mass distribution for the 76X release	111

Introduction

During the past five decades, since it was proposed in order to reorganize the "particle zoo" originated from the first particle physics experiments, the Standard Model of particle physics (SM) has achieved numerous successes. In the context of this theory, a large number of experimental processes and predictions are justified with only a limited number of particles, representing the building blocks of the matter so far discovered. Twelve fermions, called leptons and quarks, are currently the fundamental matter constituents, while the forces between them are mediated by four bosons, called gauge bosons (γ , g , Z and W).

Despite being a successful theory, since various confirmations were provided during the years, the room for improvement is still large and numerous are the open questions. The answers to many of them can only be studied at high energies and luminosities, that is why the experimental particle physics research is currently pushed forward with powerful particle accelerators like the Large Hadron Collider (LHC). LHC is a proton-proton collider which began operations in 2008, hosting four experiments: ATLAS, LHCb, ALICE and CMS. The latter is a general-purpose detector, capable of studying many aspects of the proton collisions. It detects the widest range of neutral and charged particles, while being optimized for precise identification of muons, electrons and photons. One of the ultimate goal of the CMS detector was to verify the SM theoretical mechanism, proposed in the sixties by Peter Higgs et al., through which massless gauge bosons, together with quark and leptons, could acquire their masses. The existence of the boson predicted by the Higgs mechanism was confirmed in 2012 when a neutral spin-0 particle with a mass of approximately $125 \text{ GeV}/c^2$ was discovered by the joint efforts of the ATLAS and CMS Collaborations.

Since the Higgs boson discovery, the measurement of its mass is an important quantity: with m_H known, all of the properties of the Higgs boson, such as its production cross section and partial decay widths, can be predicted. Many efforts are made to increase the precision of the mass measurements contextually reducing the associated uncertainties. The $H \rightarrow ZZ^* \rightarrow 4l$ decay is considered the golden-plated mode for the study of the Higgs boson decay, due to its low background signals and the excellent momentum resolution. However, the drawback of this channel is the reduced available statistics. The upgrade of LHC to Run 2 at $\sqrt{s} = 13 \text{ TeV}$ and high luminosities responds to the need of more statistics in order to reduce the statistical error related to the boson mass. Moreover, to obtain a precise measurement of the mass of a resonance decaying in four leptons, it is also crucial to precisely calibrate the individual lepton momentum scales and resolutions to a level such that the systematic uncertainty on the measured mass is substantially smaller than the statistical ones. The main purpose of this work

is to study the energy scale calibration of the electromagnetic calorimeter of the CMS experiment (ECAL), focusing on the non-linear energy scale effects that still remain after the calibration of the detector and worsen the systematic errors on the leptons reconstructed energies. The electrons from the decays of few well-known SM resonances, the Z-boson and the Y- and J/ψ -mesons are studied both in data and simulations. Energy corrections are derived and applied to electrons from simulated Higgs boson decays in order to study the systematic uncertainty on its measured mass.

The work is divided into four chapters. In the first chapter the Standard Model is introduced, together with a short introduction on the Weak interactions and the Higgs theory. The Higgs boson main production and decay channels are presented, as well as the state of the art concerning its mass measurements. In the second chapter the LHC accelerator is briefly presented and the CMS experiment is described, focusing on the characteristics of its electromagnetic calorimeter ECAL, which is crucial for the identifications and measurements of the electrons in the four leptons Higgs decay channel. In the third chapter the algorithms for the electron energy reconstruction are presented. The various detector calibration methods are considered as well. In the fourth chapter the fine-tuning of the ECAL energy scale are presented together with the study of the residual non-linearity. The correction function to be applied to the electron energy from the decay of simulated Higgs boson events are derived and the systematic uncertainty on its reconstructed mass is evaluated.

1 | The Standard Model of Particle Physics

In this chapter the Standard Model of particle physics is introduced. After a brief presentation of the general theory, the weak interactions together with the Higgs mechanism are reviewed. The state-of-the-art knowledge regarding the Higgs boson particle, its main production channels and decay channels and its mass measurements are presented as well.

1.1 Particle content

The Standard Model of particle physics (SM) is a theoretical model which reflects our current knowledge of elementary constituents of matter and their interactions at scales of about 1 fm and below [1, 2]. Up to now, four kind of different interactions have been discovered: the electromagnetic, the gravitational, the strong and the electro-weak force. All of them, excluding gravity, are described by means of a *quantum field gauge theory*.

The SM has been built thanks to the joined efforts of many scientists that, from the second half of the twentieth century, have developed a theory unifying the electromagnetic, weak and strong forces [3, 4]:

1. **Electromagnetic interaction:** is described by Quantum Electrodynamics (QED), a generalization of Maxwell's classical theory of electromagnetism to a relativistic quantum field theory. The QED represents the first historical hint that the principle of gauge invariance under the group $U(1)$ for the electromagnetic force (yet present at the classical level), is also a key ingredient of the related quantum theory.
2. **Weak interaction:** is responsible for various particles decays, such as the β decay of the neutron or the charged pion decay $\pi^\pm \rightarrow \mu^\pm \nu_\mu$. In the SM Lagrangian weak and electromagnetic interactions are unified by mean of a $SU(2) \times U(1)$ gauge symmetry. The electro-weak symmetry is spontaneously broken by the Higgs field vacuum. Electromagnetic and weak interactions become very different at low momentum scales.
3. **Strong interaction:** is described by Quantum Chromodynamics (QCD), the theory of quarks and gluons (partons) interacting because of their color charge, associated with an $SU(3)$ gauge symmetry. Strongly interacting particles like

protons, neutrons and mesons emerged to be bound states of these elementary particles. The QCD provides a basis both for the quark model of hadron and for the parton model.

With the ingredients mentioned above, the SM applies not only at small scales and high energies but also in phenomena underlying larger distance scales both in atomic (as an example, QED provides a precise description of the Hydrogen energy levels) and in nuclear physics (where the strong force gives rise to the binding of protons and nucleons in atomic nuclei, allowing the observed abundance of different natural elements in our Universe, while the strong, weak and electromagnetic interactions are responsible for nuclear α , β and γ decays).

Although giving a highly successful description of the physics governing our Universe, the Standard Model still leaves open questions, and there are strong arguments calling for an extension of this description. The numerous shortcomings of the SM are one of the main motivation for the construction of collider facilities such as the LHC (cf. sec. 2.2). The most striking of the SM deficits are here briefly mentioned:

- The hierarchy problem, that is the question why the scale of electroweak symmetry breaking (or alternatively the Higgs boson mass) is $\mathcal{O}(100 \text{ GeV})$. It is a fundamental problem since, in principle, this scale should receive corrections of the order of the largest scale relevant in the SM (like the Plank scale or the grand unification scale). This discrepancy can either be explained by a fine tuning of tree-level and loop contributions to the Higgs mass or by introduction of a new symmetry (like supersymmetry).
- The large number of parameters of the Standard Model is an aesthetic and conceptual problem: in a fundamental theory one would expect explanations of the values took by certain parameters but, in the SM, the values of more than 30 parameters (masses, mixing angles, couplings) must be put "by hand". In addition, the values of some of them are rather puzzling from a phenomenological point of view (as an example the mass range of fermions starts from MeV energy scale for neutrinos and extends to more than 170 GeV for the top quark).
- The particle character of the neutrino is still open: its neutral electric charge makes it special among other fermions since it may be its own antiparticle ("Majorana" instead of "Dirac" neutrino).
- In the SM the quantization of electric charge can only be explained if magnetic monopoles exist (which, so far, have not be observed).
- The Standard model does not provide any charge or mass unification at some large unification scale.
- From a cosmological side, there is still not a renormalizable quantum theory of gravity nor an observation of a "graviton" (the boson associated with this force field) or a candidate able to explain the origin of the observed dark matter/energy content of the Universe.

The Standard Model provides the existence of two kind of elementary particles: spin- $\frac{1}{2}$ particles, called *fermions*, which obey the Fermi-Dirac statistics and the Pauli

exclusion principle and integer-spin particle, called bosons, which obey the Bose-Einstein statistics.

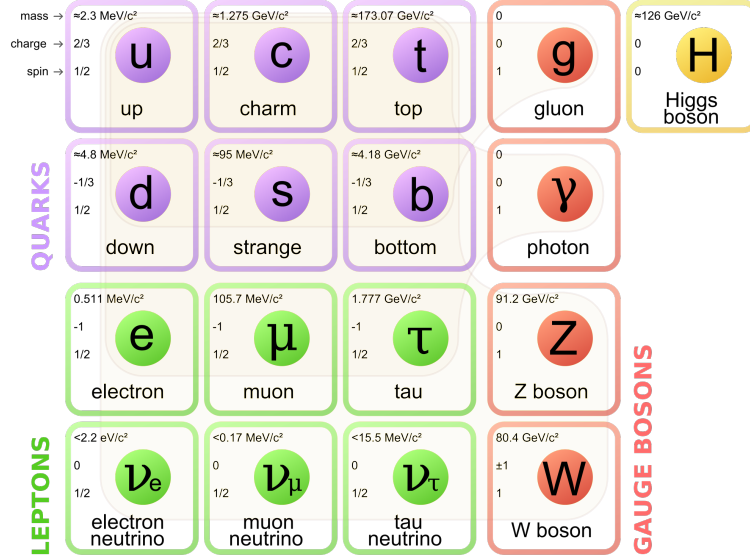


Figure 1.1: The Standard Model elementary particles content: matter particles are indicated in the first three columns, gauge bosons in the fourth and the Higgs boson in the fifth. Each fermions have an antiparticle partner, here neglected, with the same mass but different quantum numbers. In the top left corner the mass, charge and spin of the particle are reported [5].

As showing in fig. 1.1, there are 12 elementary fermions that represent all the matter constituents. The fundamental fermions observed up to now are subdivided in *leptons* and *quarks*: these two groups of particles are subdivided in three *families* (or generations) that behaves almost identically under interactions.

The three known lepton families are the electron (e), the muon (μ) and the tau (τ), each of them coming with its associated neutrino (ν_e , ν_μ , and ν_τ). The leptons do not have color charge so they are not affected by the strong force. The neutrinos, having either no electric charge, can only interact through the weak force.

The six quark flavours are labelled as up (u), down (d), charm (c), strange (s), top (t) and bottom (b). Quarks have color charge and interact via all the three fundamental forces described. Since no isolated color charges are present in nature, quarks spend most of their life bounded in color-less states: mesons particle are bound states of two quarks ($q\bar{q}$) while baryons are formed by three quarks (qqq).

All the particles listed above are accompanied by respective anti-particles with opposite quantum numbers and the same couplings of their counterparts.

Bosons are the force carriers of fundamental interactions: every interaction between fermions is described via an exchange of a boson. There are five bosons in the SM: the photon γ , the carrier of the electromagnetic force, the W and Z bosons, that mediate the weak force and eight coloured gluons (g) mediators of the strong interaction. While the gluons and photons are massless, the other two have mass and are unstable: for the Z -boson $m_Z = 91.1876(21) \text{ GeV}$ and $\Gamma_Z = 2.4952(23) \text{ GeV}$ [6] and for the W -boson

$m_W = 80.385(15)$ GeV and $\Gamma_W = 2.085(42)$ GeV [7]. In addition to these spin-1 bosons, there is the Higgs boson, a neutral massive scalar particle created as an excitation of the Higgs field. This field has a fundamental importance since it allows to understand why particles which, according to the interaction symmetries should be massless, are instead massive (cf. chp. 1.3)

Before analysing some aspects of the fundamental interactions listed so far, a brief overview of various Standard Model symmetries, which play a prominent role both in its structure and phenomena, it is given [8]:

- The local symmetry under $SU(3)_C \times SU(2)_L \times U(1)_Y$ determines the gauge interactions. The vacuum expectation value of the Higgs field breaks this symmetry giving the W and Z bosons their masses while maintaining the photon and gluon massless.
- The invariance under space-time translations, rotations and boosts (Poincare symmetry) leads to conservations of energy, momentum and angular momentum.
- The discrete symmetries parity (P) such that an inversion of coordinates takes place, time reversal (T) where $t \rightarrow -t$ and charge conjugation (C) that transform a particles into the correspondent antiparticles are involved differently. While QCD and QED conserve each symmetry individually, the electroweak gauge group $SU(2)_L \times U(1)_Y$ breaks both P and C acting differently on left and right-handed fields (cf. chp. 1.2). CP and T are broken also by the Yukawa couplings between fermions and the Higgs boson. The combination CPT is instead conserved in every known interaction, assuring that the mass and lifetimes of all particles is equal to that of the correspondent antiparticles.
- The Standard Model Lagrangian is invariant under a global common phase transformation of all lepton fields. The associated charges are baryon number ($B = \pm 1/3$ for quarks and antiquarks) and lepton number ($L = \pm 1$ for leptons and anti-leptons)

1.2 Weak interactions

1.2.1 Local gauge invariance

The general principle of non-Abelian gauge invariance, which dictates the structure of the interactions between fermions and vector bosons as well as the vector boson self-interactions, is the generalisation to non-Abelian symmetry groups of the Abelian gauge symmetry found in QED [9]. The Lagrangian for a free fermion field ψ with mass m has a global $U(1)$ symmetry such that:

$$\mathcal{L}_0 = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \quad (1.1)$$

is invariant under the phase transformation $\psi(x) \rightarrow \psi'(x) = e^{i\alpha}\psi(x)$ for any arbitrary real number α . QED is easily derived extending this global symmetry to a *local* transformations, such that α depends on the space-time coordinate $\alpha \rightarrow \alpha(x)$. In order

to maintain eq. 1.1 invariant under the now local gauge transformation, a covariant derivative together with a vector field A_μ must be introduced:

$$\partial_\mu \rightarrow D_\mu \equiv \partial_\mu + ieA_\mu \quad (1.2)$$

where e is identified as the elementary charge. The local gauge transformations form the electromagnetic gauge group $U(1)$. As a consequence of eq. 1.2 the interaction of the vector field A_μ with the electromagnetic current $j_{em}^\mu = e\bar{\psi}\gamma^\mu\psi$ can be defined in the form $j_{em}^\mu A_\mu$. In order to add the kinetic term for the boson associated with A_μ field (the photon) the following expression from classical electrodynamics is exploited:

$$\mathcal{L}_A = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (1.3)$$

where $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$. It is possible to summarize the steps followed so far as:

1. Identifying the global symmetry of the free Lagrangian.
2. Making the symmetry local by substituting the partial derivative ∂_μ with the covariant derivative D_μ , which embed a vector field.
3. Adding a kinetic term for the vector field.

This procedure can be extended to the case of non-Abelian symmetries as follows:

1. The non interacting system is now described by a generic multiplet of fermion fields of mass m , defined as $\Psi = (\psi_1, \psi_2, \dots, \psi_n)^T$. The non interacting Lagrangian is:

$$\mathcal{L}_0 = \bar{\Psi}(i\gamma^\mu\partial_\mu - m)\Psi \quad (1.4)$$

and the global transformation is defined to be a unitary matrix U parametrised by N parameters $\Psi(x) \rightarrow U(\alpha^1, \dots, \alpha^N)\Psi(x)$. The set of matrices U generally form a non-Abelian group G under multiplication $U(\alpha^1, \dots, \alpha^N) = \exp(i\alpha^a T^a)$. The T^i matrices are Hermitian, they are called "generators", and satisfy $[T^a, T^b] = if^{abc}T^c$. Physically important examples for G are the groups $SU(n)$, unitary matrices with determinant equal 1: $SU(2)$ has $N = 3$ generators $T^a = \sigma^a/2$ while $SU(3)$ has $N = 8$ generators $T^a = \lambda^a/2$ (σ^a and λ^a are the so called Pauli and Gell-Mann matrices).

2. In order to extend the global symmetry to a local one the constant α^a are converted into real functions $\alpha^a(x)$ while introducing a covariant derivative:

$$\partial_\mu \rightarrow D_\mu \equiv \partial_\mu - ig\mathbf{W}_\mu \quad (1.5)$$

and the vector field $\mathbf{W}_\mu$ together with a coupling constant g are introduced as well. Expanding $\mathbf{W}_\mu(x) = T^a W_\mu^a(x)$ a set of N fields enter the Lagrangian in eq. 1.4 giving the interaction term:

$$\mathcal{L} = \mathcal{L}_0 + \underbrace{g\bar{\Psi}\gamma^\mu\mathbf{W}_\mu\Psi}_{\mathcal{L}_{int}}$$

describing the interaction between the current $J^{a,\mu} \equiv -g\bar{\Psi}\gamma^\mu T^a\Psi$ and the gauge field W_μ^a . The Lagrangian for a multiplet of scalar fields $\Phi = (\phi_1, \dots, \phi_n)^T$ is obtained with an analogous construction, substituting:

$$(\partial_\mu\Phi)^\dagger(\partial^\mu\Phi) - m^2\Phi^\dagger\Phi \rightarrow (D_\mu\Phi)^\dagger(D^\mu\Phi) - m^2\Phi^\dagger\Phi \quad (1.6)$$

This finds its application in the Higgs sector of the SM (cf. chp. 1.3).

3. The kinetic term for the W field should be added. This can be obtained from a generalization of the electromagnetic field strength tensor $F_{\mu\nu}$:

$$\mathbf{F}_{\mu\nu} = T^a F_{\mu\nu}^a = \partial_\mu \mathbf{W}_\nu - \partial_\nu \mathbf{W}_\mu - ig[\mathbf{W}_\mu, \mathbf{W}_\nu]$$

Under the gauge transformation, the field strength transform as $\mathbf{F}_{\mu\nu} \rightarrow \mathbf{F}'_{\mu\nu} = U\mathbf{F}_{\mu\nu}U^{-1}$: as a consequence $\text{Tr}(\mathbf{F}_{\mu\nu}\mathbf{F}^{\mu\nu})$ is gauge invariant, providing the so-called Yang-Mills Lagrangian:

$$\mathcal{L}_W = -\frac{1}{2}\text{Tr}(\mathbf{F}_{\mu\nu}\mathbf{F}^{\mu\nu}) = -\frac{1}{4}(F_{\mu\nu}^a F^{a,\mu\nu}) \quad (1.7)$$

The part of $\mathcal{L}_W$ quadratic in the W fields describes the free propagation of the gauge bosons while the cubic and quartic term describe self-interactions between the gauge bosons that are completely determined by the gauge symmetry. The Lagrangians in eq. 1.3 and 1.7 describe *massless* vector bosons: mass terms for the gauge fields of the form $\frac{1}{2}m^2 A_\mu A^\mu$ and $\frac{1}{2}m^2 W_\mu^a W^{a,\mu}$ are not invariant under a gauge transformation and thus break the gauge symmetry.

1.2.2 Formulation of the Electroweak Standard Model

As already mentioned, the $SU(2) \times U(1)$ symmetry transformation acts differently on left- and right-handed fermion fields. It is useful to project out the left and right hand composition of the general Dirac field ψ defining:

$$\psi_L = \frac{1 - \gamma_5}{2}\psi, \quad \psi_R = \frac{1 + \gamma_5}{2}\psi \quad (1.8)$$

The $SU(2)$ part of the electroweak symmetry group is called weak isospin group, with associated quantum numbers I and I_3 : left-handed fields have $I = 1/2$ and form doublets which transform with a unitary matrix $U = \exp(i\alpha^a \sigma^a/2)$ (as an example, the electron doublet is $(\nu_e, e)_L$) while right-handed fields have $I = 0$ and form singlets invariant under the weak isospin transformation (such as e_R, u_R, d_R). For each doublet or singlet, the $U(1)$ part corresponds to multiplication by a phase factor $\exp(i\alpha Y/2)$ and the quantum number associated Y is defined as the weak hyper-charge. The electromagnetic gauge group appears as a subgroup of the electroweak symmetry group obtained combining a hyper-charge transformation with a particular isospin transformation. The electric charge of a particle is given by the Gell-Mann-Nishijima formula:

$$Q = I_3 + \frac{Y}{2}$$

The assignment of quantum numbers to the lepton and quark fields is given in tab. 1.1

	ν_{eL}	e_L	e_R	$\nu_{\mu L}$	μ_L	μ_R	$\nu_{\tau L}$	τ_L	τ_R	u_L	d_L	u_R	d_R
	$\nu_{\mu L}$	μ_L	μ_R	$\nu_{\tau L}$	τ_L	τ_R	t_L	b_L	t_R	b_R			
I_3	$+\frac{1}{2}$	$-\frac{1}{2}$	0	$+\frac{1}{2}$	$-\frac{1}{2}$	0	$+\frac{1}{2}$	$-\frac{1}{2}$	0	0			
Y	-1	-1	-2	$+\frac{1}{3}$	$+\frac{1}{3}$	$+\frac{4}{3}$	$+\frac{2}{3}$	$+\frac{2}{3}$	$+\frac{4}{3}$	$-\frac{2}{3}$			
Q	0	-1	-1	$+\frac{2}{3}$	$-\frac{1}{3}$	$+\frac{2}{3}$	$-\frac{1}{3}$	$+\frac{2}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$			

Table 1.1: The weak isospin and hyper-charge of the left-handed and right-handed leptons and quarks groups. The rows correspond to the different generations.

The structure just described can be embedded in a bigger one, corresponding to the gauge invariant field theory of the unified electromagnetic and weak interactions [10] by interpreting $SU(2)_L \times U(1)_Y$ as the group of gauge transformations under which the Lagrangian is invariant. The full symmetry must be broken down to the electromagnetic gauge symmetry, a consequence of the fact that the photon is massless while the gauge bosons $W^\pm$ and Z have non-zero mass. In the Standard Model the required symmetry breaking is achieved with the Higgs mechanism via a single Higgs field, which is a doublet under $SU(2)_L$.

The full electroweak Lagrangian before the symmetry breaking due to the Higgs boson consist of two parts, the gauge and the fermions terms, as can be seen in eq. 1.9:

$$\mathcal{L}_{EW} = \mathcal{L}_{WB} + \mathcal{L}_F \quad (1.9)$$

which will be discussed soon. A complete list of the Feynman rules corresponding to $\mathcal{L}_{EW}$ can be found in [11].

Gauge field Lagrangian $\mathcal{L}_{WB}$

The gauge field Lagrangian is constructed following the method developed in chp: 1.2.1 for each factor of the group $SU(2)_L \times U(1)_Y$. The results are three vector fields W_μ^a , associated with the three generators of $SU(2)_L$, and one vector field B_μ , associated with the hyper-charge group $U(1)_Y$. The corresponding field strength tensors are:

$$\begin{aligned} W_{\mu\nu}^a &= \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g_2 \epsilon^{abc} W_\mu^b W_\nu^c \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu \end{aligned} \quad (1.10)$$

where the Levi-Civita tensor ϵ^{abc} is the structure constant of $SU(2)$. From the field strength tensors (eq. 1.10) the pure gauge field Lagrangian can be constructed:

$$\mathcal{L}_{WB} = -\frac{1}{4} W_{\mu\nu}^a W^{\mu\nu,a} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \quad (1.11)$$

It is important to emphasize the fact that the $\mathbf{W}_\mu$ and B_μ fields so far described are not the fields associated with the vector bosons predicted by the SM (namely $W^\pm$, Z^0 and γ) which are obtained as a linear combination of the gauge fields components.

Fermion field Lagrangian $\mathcal{L}_F$

The kinetic part of the fermion Lagrangian, together with the interaction terms between fermions and gauge bosons, can be obtained from the appropriate covariant derivatives for isospin doublets and singlets, finding:

$$\begin{aligned} \mathcal{L}_F = & \sum_{j=1}^6 \bar{\psi}_L^j i\gamma^\mu \left(\partial_\mu - ig_2 \frac{\sigma^a}{2} W_\mu^a + ig_1 \frac{Y}{2} B_\mu \right) \psi_L^j + \\ & + \sum_{j=1}^9 \bar{\psi}_R^j i\gamma^\mu \left(\partial_\mu + ig_1 \frac{Y}{2} B_\mu \right) \psi_R^j \end{aligned} \quad (1.12)$$

In eq. 1.12 the term ψ_L^j stands for the left-handed fermions doublets while ψ_R^j for the right-handed singlets. The terms with ∂_μ and B_μ are understood to contain the two-dimensional unit matrix in isospin space. It is very important to notice that there are no fermion mass terms at this stage of the formulation: inserting these terms by hand require to add to the Lagrangian objects with the form $\bar{\psi}\psi = \bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L$ that would mix left-handed and right-handed fields and thus explicit break gauge invariance. Fermion masses will be generated dynamically by gauge invariant Yukawa interactions with the Higgs field (cf. chp. 1.3).

1.3 Higgs field and spontaneous symmetry breaking

There is a gauge invariant way to recover the fermions and bosons masses via a spontaneously broke of the local $SU(2) \times U(1)$ symmetry. In the SM the electroweak symmetry breaking without disrupting gauge invariance is achieved through the Brout-Englert-Higgs mechanism [12].

Spontaneous symmetry breaking

There is a simple example showing how the mass can arise from a broken symmetry. Starting from a scalar real field ϕ with a Lagrangian:

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \underbrace{\frac{1}{2} \mu^2 \phi^2 - \frac{1}{4} \lambda \phi^4}_{V(\phi)} \quad (1.13)$$

if the mass term μ^2 is positive, the potential $V(\phi)$ is also positive if the self-coupling λ is positive. The minimum of this potential is located in $\langle 0|\phi|0\rangle \equiv \phi_0 = 0$ as shown in fig. 1.2. In turn, if $\mu^2 < 0$ the potential has a minimum when $\frac{\partial V}{\partial \phi} = \mu^2 \phi + \lambda \phi^3 = 0$, that is to say when:

$$\langle 0|\phi^2|0\rangle \equiv \phi_0^2 = -\frac{\mu^2}{\lambda} \equiv v^2$$

The quantity $\pm v$ is called the vacuum expectation value (vev) of the scalar field ϕ . If $\mu^2 < 0$ the Lagrangian in eq. 1.13 is no more the Lagrangian of a particle of mass μ . To interpret correctly the theory, we must expand around one of the minima v by defining the field σ as $\phi = v + \sigma$. The new Lagrangian becomes:

$$\mathcal{L} = \frac{1}{2} \partial_\mu \sigma \partial^\mu \sigma - (-\mu^2) \sigma^2 - \sqrt{-\mu^2 \lambda} \sigma^3 - \frac{\lambda}{4} \sigma^4 + \text{const.}$$

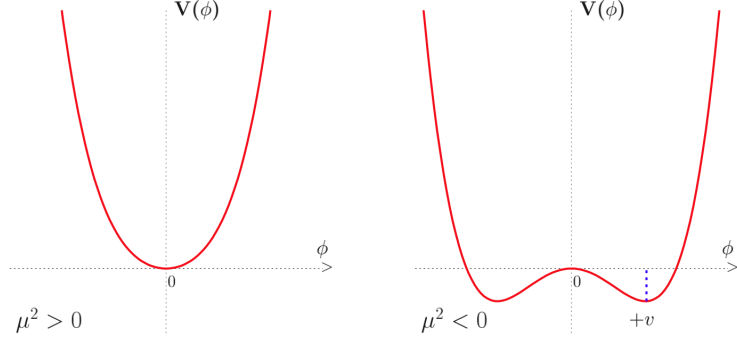


Figure 1.2: The potential of a scalar field ϕ in the case $\mu^2 > 0$ (left) or $\mu^2 < 0$ (right) [13].

This is now the theory of a scalar field of mass $m^2 = -2\mu^2$, with σ^3 and σ^4 being the self-interactions: since there are now cubic terms, the reflexion symmetry $\phi \rightarrow -\phi$ of eq. 1.13 is not anymore apparent in $\mathcal{L}$. The symmetry is spontaneously broken.

Higgs field

The Higgs sector of the SM consists of an isospin doublet of complex-valued scalar fields:

$$\Phi(x) = \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix}$$

This doublet has hyper-charge $Y = 1$ so that the appropriate covariant derivative is:

$$D_\mu = \partial_\mu - ig_2 \frac{\sigma^a}{2} W_\mu^a + ig_1 \frac{1}{2} B_\mu$$

As a result of the Gell-Mann-Nishijima relation, ϕ^+ carries electric charge $+1$ while ϕ^0 is electrically neutral. Following eq. 1.6 the Higgs Lagrangian reads:

$$\mathcal{L}_H = (D_\mu \Phi)^\dagger (D^\mu \Phi) - \underbrace{\frac{\lambda}{4} \left(\Phi^\dagger \Phi - \frac{2\mu^2}{\lambda} \right)^2 - \frac{\mu^4}{\lambda}}_{V(\Phi)} \quad (1.14)$$

The last term of eq. 1.14 is a constant that can be dropped in the following. In the vacuum (the ground state of the theory) the potential energy of the field must be minimal. For positive μ^2 and λ values the potential $V(\Phi)$ has the form of a "Mexican-hat" (cf. fig. 1.2 noticing the slightly change of notation in the potential $V(\Phi)$). The minimum of this potential is located at $\Phi^\dagger \Phi = \frac{2\mu^2}{\lambda}$ and the vacuum selects the configuration with:

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad \text{with} \quad v = \frac{2\mu}{\sqrt{\lambda}} \quad (1.15)$$

It is easy to see that eq. 1.15 is not invariant under a transformation with $U = \exp(i\alpha^a \sigma^a / 2)$ or a phase factor $\exp(i\alpha Y)$ so that the electroweak gauge symmetry

$SU(2)_L \times U(1)_Y$ is broken by the vacuum through a spontaneous symmetry breaking. It is worth to notice that, since the charged component of $\langle \Phi \rangle$ is precisely 0 and the electrically neutral component $v/\sqrt{2}$ is invariant under an electromagnetic gauge transformation, the $U_{em}(1)$ symmetry is preserved by the vacuum. Consequently no mass is attributed to the photon which in turn appears to be massless.

To make more explicit the vacuum state it is possible to rewrite the Higgs doublet as:

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1(x) + i\phi_2(x) \\ v + H(x) + i\chi(x) \end{pmatrix}$$

where the added fields are real-valued with zero expectation values. The Higgs potential reads:

$$V = \mu^2 H^2 + \text{other terms}$$

and it easy to see that the field H describes an electrically neutral scalar particle (the Higgs boson) with a mass $M_H = \sqrt{2}\mu$. No mass terms appears for the additional χ , ϕ_1 and ϕ_2 fields.

Inserting the Higgs doublet previously defined into the kinetic part of eq. 1.14 the coupling of the Higgs to the gauge fields can be obtained together with the mass term for the vector bosons in the non-diagonal form. The physical content becomes apparent only after a transformation from the fields W_μ^a and B_μ (in terms of which the original gauge symmetry is manifest) to the physical fields:

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \\ \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} &= \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} \end{aligned} \quad (1.16)$$

In terms of these fields the mass term is diagonal and has the following form:

$$M_W^2 W_\mu^+ W^{-\mu} + \frac{1}{2}(Z_\mu, A_\mu) \begin{pmatrix} M_Z^2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ A_\mu \end{pmatrix} \quad (1.17)$$

with the mass terms $M_W = \frac{g_2}{2}v$ and $M_Z = \frac{\sqrt{g_1^2 + g_2^2}}{2}v$. The angle of rotation in eq. 1.16 is called the weak mixing angle and it is given by $\cos \theta_W = \frac{g_2}{\sqrt{g_1^2 + g_2^2}} = \frac{M_W}{M_Z}$.

The coupling of the Higgs boson field to the other physical fields of the electroweak symmetry $W_\mu^\pm$, Z_μ and A_μ is obtained from eq. 1.17 via the replacements $M_W \rightarrow M_W(v+H)/v$ and $M_Z \rightarrow M_Z(v+H)/v$ (no replacement is required for the A_μ since the photon remains massless) which gives rise to vertices HZZ, HWW, HHZZ and HHWW. The Higgs mechanism gives masses to W and Z by breaking the $SU(2)_L \times U(1)_Y$ symmetry with a non-zero expectation value leaving the original Lagrangian gauge invariant. This turns out to be essential for showing that the theory can be renormalised and is thus self-consistent.

So far the Higgs mechanism has justified the boson masses of the SM but the fermion are still massless. It turns out that the spontaneous symmetry breaking of the $SU(2)_L \times U(1)_Y$ symmetry accounts also for the fermions masses via a gauge invariant

Yukawa interaction between the Higgs and the fermion fields. For a single generation of leptons and quarks, the Yukawa term in the Lagrangian has the form:

$$\mathcal{L}_{Y,gen} = -Y_l \bar{L}_L \Phi l_R - Y_d \bar{Q}_L \Phi d_R - Y_u \bar{Q}_L \Phi^c u_R + \text{h.c.}$$

The vacuum expectation value of ϕ^0 gives rise to fermion mass terms, where the masses are related to the Yukawa couplings $m_f = Y_f \frac{v}{\sqrt{2}}$ for $f = l, d, u$. In the case of three generations a more intricate structure arise because the Yukawa terms mix fermions of different generations.

To summarize, the boson masses in the SM are accessed by:

$$\begin{aligned} M_H &= \sqrt{\lambda} v \\ M_W &= \frac{ev}{2 \sin \theta_W} \\ M_Z &= \frac{M_W}{\cos \theta_W} \\ M_\gamma &= 0 \end{aligned} \tag{1.18}$$

1.3.1 Higgs boson searches

Since the λ parameter in eq. 1.18 is not given by the theory, the Higgs boson mass must be determined experimentally. Phenomenological studies of the production and decays of the Higgs boson have started as early as the 1970s and are carried on still in these days at the Large Hadron Collider (cf. chp. 2) employing protons collisions at a center of mass energy of $\mathcal{O}(\text{TeV})$.

Production mechanism

There are four main Higgs boson production channels at the energies available at LHC. The production cross section in pp collisions at a center of mass energy of $\sqrt{s} = 7 \text{ TeV}$ can be seen in fig. 1.3, while the corresponding leading order Feynman diagrams are shown in fig. 1.4.

The **gluon fusion** process ($gg \rightarrow H$) is the dominant production process over the entire mass range accessible at the Large Hadron Collider. As can be seen in fig. 1.4 (a), it is mediated by a heavy quark triangular loop: because of the Higgs couplings to fermions, the top quark loop is the most important contributor with the next-to-leading (NLO) contributions by the bottom quark being $\mathcal{O}(m_t^2/m_b^2)$ smaller. NLO corrections have been shown to increase the LO cross section of 80-100%: these corrections are non-negligible and are computed by means of the ratio between the cross sections at an higher and lower order (K-factor). The cross section of this channel decrease with m_H as the gluon content of the proton decreases with increasing the Bjorken x variable [15, 16].

The second largest production mechanism of the Higgs boson, about one order of magnitude smaller than the previous, is the **vector boson fusion** process (VBF or $qq \rightarrow qqH$) process where the H-boson is originated from the fusion of two weak bosons radiated off from the incoming quarks. Their hadronization is responsible of two forward jets of high invariant mass that help tagging the event, differentiating it from the background. Another interesting property of this channel is the reduced hadronic

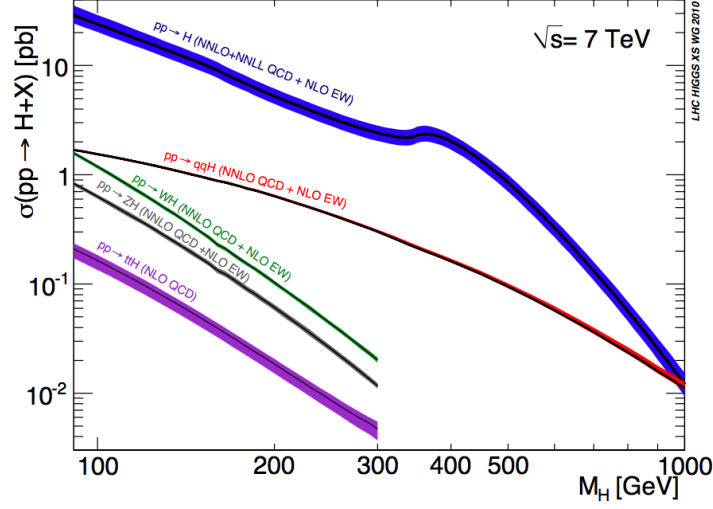


Figure 1.3: The Higgs boson production cross section at collisions at a center of mass energy of $\sqrt{s} = 7$ TeV as a function of the boson mass for different production mechanism. From top to bottom: the gluon fusion (blue), the vector boson fusion (red), the associate production with a W/Z-boson (green/grey) and finally $t\bar{t}$ production (violet). Next-to-next-leading-order (NNLO) corrections an Next-leading-order (NLO) electroweak corrections are taken into account [14].

activity in the region between the tagged jets, since they are colour connected. This channel becomes dominant in the theories of very massive Higgs [17, 18].

In the **Higgs-strahlung** ($q\bar{q} \rightarrow W\bar{H}$ or $q\bar{q} \rightarrow Z\bar{H}$) and **$t\bar{t}$ associated production** ($q\bar{q} \rightarrow t\bar{t}H$) processes the Higgs boson is produced in association with a W/Z boson or a pair of top quarks: in both cases the decay processes can be used to tag the events. Nevertheless the cross section is orders of magnitude smaller with respect to the gluon fusion. The Quantum Chromodynamic corrections are large and the K-factors range from 1.2 to 1.4, depending on the Higgs mass [19, 20].

Decay modes

Depending on the Higgs mass, different decay channels are allowed. Its total decay width and the different branching ratios depend on the Higgs couplings to the SM bosons and fermions (cf. chp. 1.3). The Higgs decays branching ratios, including also corrections from quantum chromodynamic and electroweak, can be seen in fig. 1.5(b): due to the dependence of its couplings on the other particle masses, the Higgs boson tends to decay into the heaviest particles kinematically allowed. As a consequence, new channels open up when the Higgs mass is above a dilepton/quark or vector boson threshold: light-fermion decay are dominant at low Higgs masses but when a decay in the heavy EW bosons is possible it suddenly dominate. At high Higgs masses also the decay in pairs of top quarks are relevant.

The $H \rightarrow \gamma\gamma$ decay is the main channel for the discovery of the Higgs boson at masses of about 140 GeV or below. It is the clearest channel [21], despite the low branching ratio, since the boson does not couple directly to photons at tree level, but such coupling

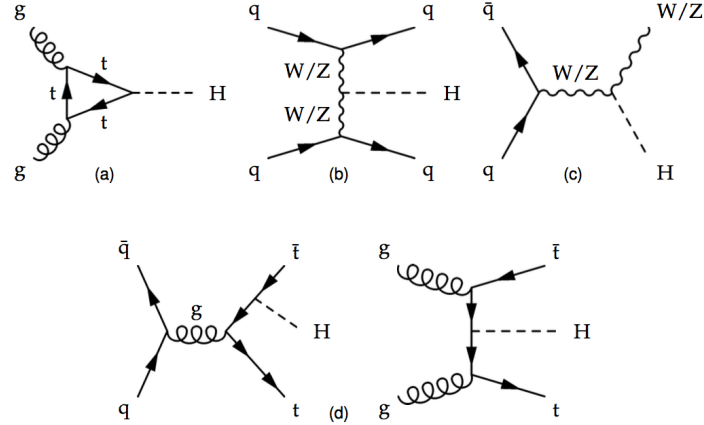


Figure 1.4: The Higgs boson production at pp collider most important low order (LO) Feynman diagrams. From top to bottom: gluon fusion (a), vector boson fusion (b), Higgs-strahlung (c) and $t\bar{t}$ associated production (d).

can arise only via fermion loops as can be seen in fig. 1.5(a). Large backgrounds come from prompt photon pairs produced by quarks and gluons in the initial state (initial state radiation), from two jets which fake the characteristic photon signature and from Drell-Yan production of electron pairs. The signal signature of this channel is represented by two photons with high transverse energy. The Higgs signal can be identified as a narrow peak over a large and smooth falling background.

The $H \rightarrow ZZ^* \rightarrow 4l$ decay is considered the golden-plated mode for the discovery of the Higgs boson [22]: the continuous backgrounds from ZZ^* , $t\bar{t}$ and $Zb\bar{b}$ can be easily suppressed in an efficient way by some requirements on the lepton isolation, transverse momentum and invariant mass, granting an excellent momentum resolution. Two pairs of leptons of same flavour are needed. The decays of the EW bosons can occur in three different configurations: $4e$, 4μ and $2e2\mu$. Various kinematic variables, with strong discriminating power, are used: among all the five emission angles, the two invariant masses of the leptonic pairs and the requirement that each lepton pair appear from the same vertex. To recover the final state radiation from leptons produced via bremsstrahlung photons are collected as well (more informations are given in chp. 3).

The $H \rightarrow WW \rightarrow 4l$ decay signature is fully dominated by the decay modes of the W-bosons [23]: two charged leptons with opposite charge and a large amount of missing transverse energy carried by the neutrinos matched in flavour with the corresponding leptons. The presence of the neutrinos complicates the analysis: the mass resolution is poor so the search strategy is then based on event counting for which an accurate knowledge of all the possible backgrounds is needed. Events are classified in categories based on the number of jets and on the flavour of the decay leptons. Events with more than two jets are rejected since the best sensitivity is expected in sub-channels with few jets where the irreducible background is given by non resonant WW pair production and the reducible by misidentified jets in $W + \text{jets}$.

The $H \rightarrow \tau^+\tau^-$ channel is directly sensitive to the Higgs coupling to the tau lepton [24]. Since the τ can decay in various different states, different sub-channels are

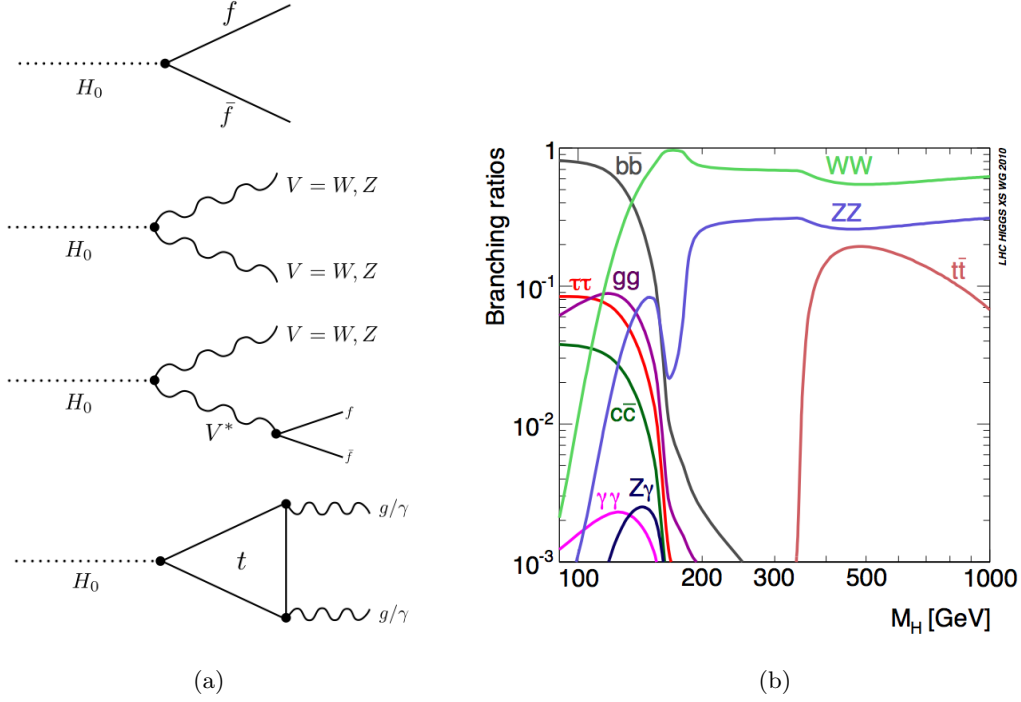


Figure 1.5: The decay branching ratios of the SM Higgs boson in different channels versus its possible mass in fig. 1.5(b) [14]. In fig. 1.5(a) the LO decay Feynman diagrams are shown.

defined: $e\mu$, $\mu\mu$, $e\tau_{had}$ and $\mu\tau_{had}$ where τ_{had} refers to the tau hadronic decays. The events are splitted on the basis of the number of jets and on the momentum of the lepton. In the totally leptonic decay channels, the main sources of background are Drell-Yan processes and $t\bar{t}$ production, while in the hadronic decay channels the events with W or Z plus misidentified τ_{had} . Other reducible background comes from $Z \rightarrow \tau\tau$, $W + \text{jets}$ and general QCD multi-jet production. Since the Higgs boson coupling is proportional to the square of the particle mass, this channel is dominant with respect to the $\mu\mu$ or ee ones.

The $H \rightarrow b\bar{b}$ is the channel with the highest branching ratio at $m_H < 135$ GeV, as can be seen in fig. 1.5(b) [25]. Despite this, it is difficult to observe the Higgs boson decay into b-quarks since the signal is dominated by QCD production of bottom quarks via $p\bar{p} \rightarrow b\bar{b} + X$. In order to suppress this background, the Higgs searching process accounts for events with a high- p_T resonance associated to a Z- or W-boson, putting an upper limit to the additional activity. This decay process is studied in different channels, including the $Z(l\bar{l})H$, the $Z(\nu\bar{\nu})H$ and the $H(L\nu)H$ processes. The b-jets must be identified using a dedicated identifying procedure and energy cuts are imposed for $W(l\nu)$ and $Z(\nu\bar{\nu})H$. Other backgrounds come from the production of vector bosons with jets, top quarks and WW , ZZ or ZW with one of the boson decaying hadronically.

1.3.2 State of the art

In 2012, the ATLAS [26] and CMS [27] Collaborations at the LHC announced the discovery of a new particle with a mass of about 125 GeV and Higgs-boson-like properties. The Higgs boson mass was measured independently by the two experiments, using the samples of pp collision data collected in 2011 and 2012, corresponding to 5 fb^{-1} at $\sqrt{s} = 7 \text{ TeV}$ and 20 fb^{-1} at $\sqrt{s} = 8 \text{ TeV}$, during the commonly referred *LHC Run 1*. Lower and upper limits that constrain the Higgs mass to be located in an interval of $114.4 < m_H < 152 \text{ GeV}$ were set at 95% confidence level by previous measurements carried on in the "pre-LHC era" at the Tevatron and the Large electron positron collider (LEP) [28] via precision electroweak measurements. The dataset of data collected at 7 TeV allowed ATLAS and CMS to determine various exclusion regions, namely 127 GeV–600 GeV, 111.4 GeV–116.6 GeV, 119.4 GeV–122.1 GeV and 129.2 GeV–541 GeV.

The discovery was based primarily on mass peaks observed in two decay channels, the $\gamma\gamma$ and the $ZZ \rightarrow 4l$ (where one or both the Z-bosons can be off-shell and the l indicates $4e$, $2e2\mu$ or 4μ). The experimental determination of the Higgs mass has allowed to precisely predict all the SM Higgs boson properties, such as its production cross section and the partial decay widths. Other measurements on the spin, parity and coupling strengths to SM particles of the discovered particle are consistent within the uncertainties with those expected from a SM Higgs boson.

Precise mass measurements of the Higgs boson are crucial for its intrinsic importance as a fundamental parameter, but also because improved knowledge of m_H yields more precise predictions for the other Higgs boson properties. As an example, in the SM the Higgs boson mass is related to the values of the masses of the W-boson and top quark through loop induced effects: taking into account other measured SM quantities, the comparison between m_H , m_W and m_t allows direct tests on the consistency of the SM model predictions and thus to search for evidence of physics beyond it. For these reasons, in order to obtain an improved mass resolution and to reduce any kind of statistical and systematic error, data from ATLAS and CMS are combined together: the combined mass measurement represents only the first step towards combinations of other quantities (such as the couplings). Since the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ \rightarrow 4l$ decay channels offer the best mass resolution, combination is performed using only these two:

- The $\gamma\gamma$ channel is characterized by a narrow resonant signal peak containing several hundred events per experiment above a large falling continuum background. Events are divided into different categories, depending on the signal purity and mass resolution in order to reduce the expected systematic uncertainties in m_H .
- The $4l$ channel yields only a few tens of signal events per experiments but the background is so low that the signal-to-background ratio is larger than 1. During the mass analysis events are separated depending on the flavour of the lepton pairs and neither attempts to distinguish between different Higgs boson production mechanism.

The measurement of m_H and of its uncertainty is based on the maximization of a profile-likelihood ratios $\Lambda(\alpha)$ in the asymptotic regime [29]. The likelihood functions are constructed using signal and background probability density functions (PDFs) that

depend on the di-photon invariant mass in the $H \rightarrow \gamma\gamma$ channel and on the four-lepton invariant mass and kinematic discriminant in the $H \rightarrow ZZ \rightarrow 4l$ channel. The signal PDFs are derived from samples of Monte Carlo (MC) simulated events while the background came from combination of simulated and experimental data in control regions.

The combined ATLAS and CMS results for m_H in the separated $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ \rightarrow 4l$ channels are:

$$\begin{aligned} m_{H \rightarrow \gamma\gamma} &= 125.07 \pm 0.29 \text{ GeV} \\ &= 125.07 \pm 0.25 \text{ (stat.)} \pm 0.14 \text{ (syst.) GeV} \end{aligned} \quad (1.19)$$

and:

$$\begin{aligned} m_{H \rightarrow ZZ} &= 125.15 \pm 0.40 \text{ GeV} \\ &= 125.15 \pm 0.37 \text{ (stat.)} \pm 0.15 \text{ (syst.) GeV} \end{aligned} \quad (1.20)$$

The likelihood ratio scans are shown in fig. 1.6. Equation 1.19 and 1.20 show that

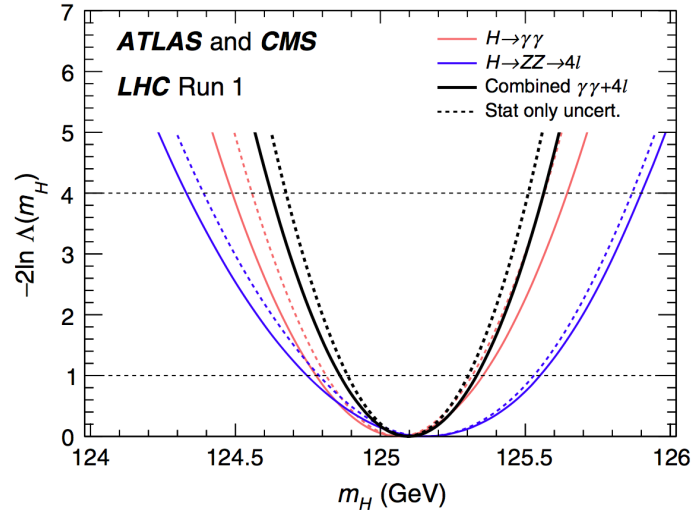


Figure 1.6: Scans of the negative log-likelihood ratio $-2 \ln \Delta(m_H)$ as function of the Higgs boson mass for the ATLAS and CMS combination of $H \rightarrow \gamma\gamma$ (red), $H \rightarrow ZZ \rightarrow 4l$ (blue) and combined (black) channels. The dashed curves show the results accounting for the statistical uncertainties only. The 1 and 2 standard deviation limits are indicated by horizontal lines at 1 and 4 respectively [30].

the uncertainties in the Higgs mass measurement are dominated by the statistical term (even if the Run 1 data sets of ATLAS and CMS are combined). In eq. 1.20, the systematic uncertainty is dominated by the absolute scale uncertainty in the momentum measurement for the muons, and in the momentum and energy measurements for the electrons. Samples with a large number of events of dilepton decays of various SM known resonances, such as the J/ψ and Y mesons together with the Z -boson are used to evaluate the absolute scales and to correct for residual misalignments in the inner tracking system (a more detailed description is given in chp. 3.2). The treatment and understanding

of the electron, photon and muons channels measurements systematic uncertainties is an important aspect of the the individual measurements and their combination. To evaluate the relative importance of the different sources of systematic uncertainty, the fit parameters are grouped into three broad classes of systematic uncertainties:

1. **Scale:** are uncertainties in the energy or momentum scale and resolution for photons, electrons and muons.
2. **Theory:** are theoretical uncertainties such as the uncertainties on the Higgs boson cross section and branching ratios. Uncertainties in the normalization of SM background processes are included as well.
3. **Other:** are the other experimental uncertainties not belonging to the previous categories.

These classes are hence subdivided into finer groups associated with specific underlying effects and the impact of each group on the overall mass uncertainty is evaluated. The contribution of each sub-group is defined as the mass shift δm_H observed when reevaluating the profile-likelihood ratio after fixing the parameter in question to its best-fit value increased or decreased by 1σ (standard deviation). The results for the CMS experiment establish that the largest systematic effects for the mass uncertainty are those related to the determination of the energy scale of the photons ($\delta m_H^{photon} = 0.1$) and those associated with the determination of the electron and muon momentum scales ($\delta m_H^{electron} = 0.12$, $\delta m_H^{muon} = 0.11$). The electron and muon scale uncertainties are dominant in the $H \rightarrow ZZ \rightarrow 4l$ channel, while the other in the same class (Scale) are negligible. Other and theory uncertainties are found to contribute only $\delta m_H^{other/theory} = 0.01$ [30]. Figure 1.7 gathers all the systematic uncertainties impacts δm_H associated with the indicated effects, both for the ATLAS and CMS experiment and the combined result.

The combination of the Higgs boson mass measurements in the $H \rightarrow \gamma\gamma$ channel (eq. 1.19) and in the $H \rightarrow ZZ \rightarrow 4l$ channel (eq. 1.20) is found to be [30]:

$$\begin{aligned}
m_H &= 125.09 \pm 0.24 \text{ GeV} \\
&= 125.09 \pm 0.21 \text{ (stat.)} \pm 0.11 \text{ (syst.) GeV} \\
&= 125.09 \pm 0.21 \text{ (stat.)} \pm 0.11 \text{ (scale.)} \pm 0.02 \text{ (other)} \pm 0.01 \text{ (theory) GeV}
\end{aligned} \tag{1.21}$$

where the total uncertainty is dominated by the statistical term, with the systematic uncertainty dominated by effects related to the photon, electron and muon energy or momentum scales and resolutions. Compatibility tests were performed as well to ascertain whether the measurements are consistent to each other, both between the two experiments and the different decay channels. The measurements are compatible in the two experiments within 2σ , as can be see in fig. 1.8. Nowadays, this is the most precise measurement of the Higgs boson mass.

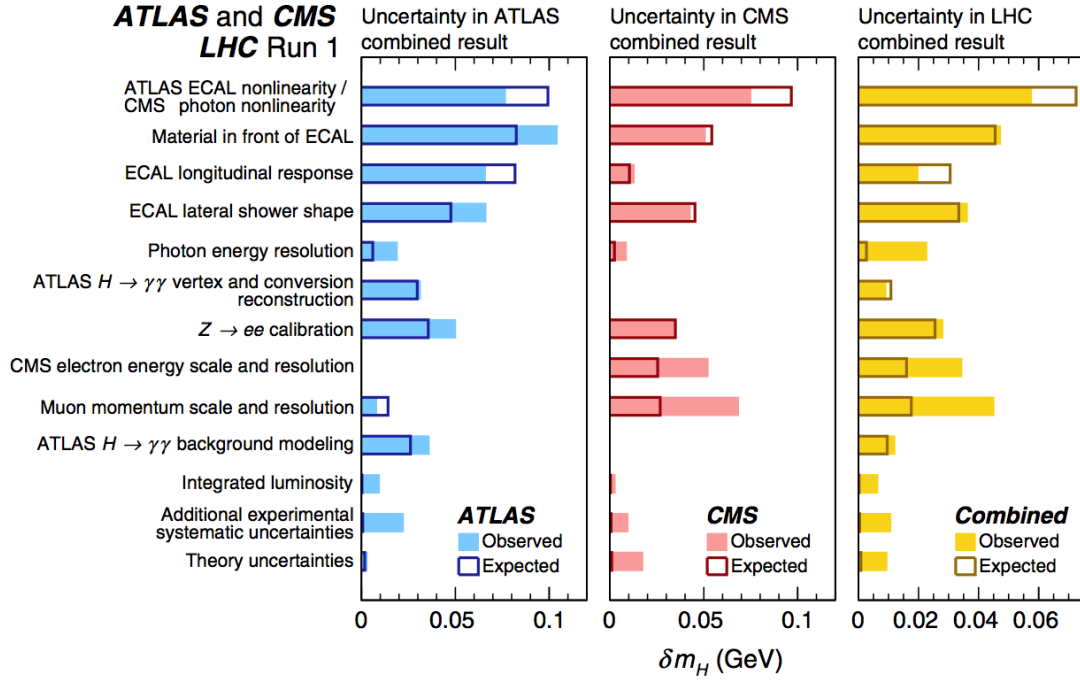


Figure 1.7: The impacts δm_H of the fit parameters groups on the ATLAS (left), CMS (center) and combined (right) mass measurement uncertainty. The observed results are shown as solid bars while the expected by empty ones [30].

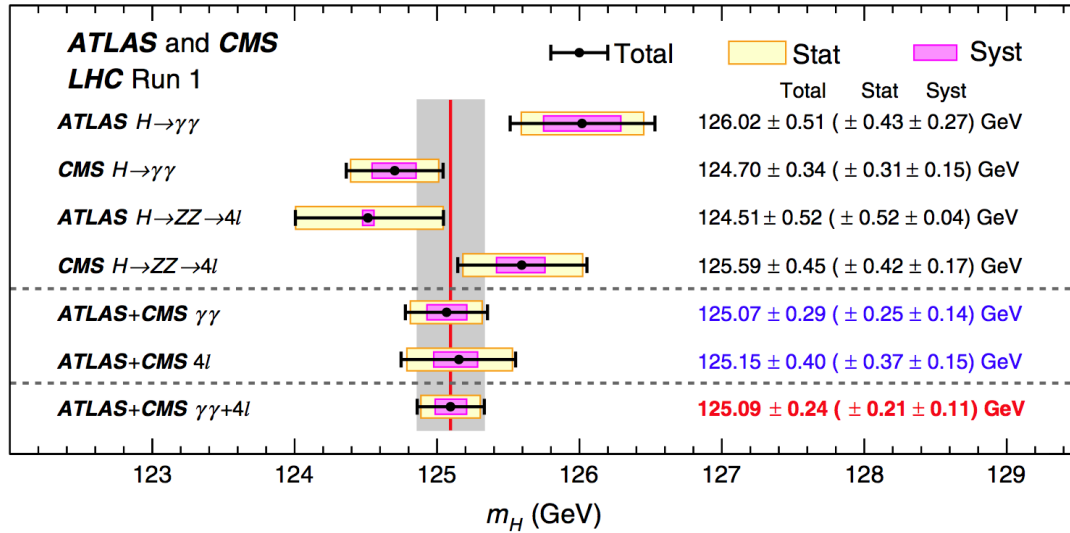


Figure 1.8: Summary of the Higgs boson mass measurements from the individual (ATLAS and CMS) and combined analysis. The systematic (magenta), statistical (yellow) and total (black) uncertainties are indicated. The red vertical line indicates the central value while the gray shaded region indicates the total uncertainty of the combined measurement.

2 | The LHC accelerator and the CMS detector

In this chapter the Large Hadron Collider accelerator complex is presented. The Compact Muon Solenoid, one of the main experiments along the accelerator ring, is described focusing on the characteristics of its electromagnetic calorimeter.

2.1 The European Organization for Nuclear Research

The European Organization for Nuclear Research (CERN), founded in 1954 by twelve member states, is one of the biggest laboratory in the world. Today it involves more than twenty countries and 10000 cooperating people in particle physics research. Located at the border between France and Switzerland, its aim is to provide the environment and the fundamental tools necessary for particle physics research activity. In order to do so, five accelerators were built during the years, each one, after having accelerated the particles to the requested energy, injects the particles into the next. As a result, from 1954 to 2009, the maximum center of mass energy achievable at CERN facilities has been raised from 50 MeV up to 14 TeV.

2.2 The Large Hadron Collider accelerator

The Large Hadron Collider (LHC) accelerator [31] is the largest and most powerful hadron collider ever built. Inaugurated on the 21st October 2009, it was installed in the same tunnel previously occupied by the Large Electron Positron Collider (LEP) [32]. The tunnel has a circular shape with a total length of 27 km and it is located under the city of Geneva, across the Franco-Swiss border.

LHC is a machine dedicated, in particular, to the discovery of new physics so, unlike its predecessor, it was built as a hadron collider. The transitions from leptonic to hadronic collider was mandatory since:

- hadrons are, commonly, heavier than leptons (electrons/positrons) so they lose less energy by synchrotron radiation while circulating in the collider. Consequently, a higher energy is achievable in the center of mass (CM) frame. LHC allow us to test the widest mass range possible so far, from B-physics (GeV/c^2) to new-physics (TeV/c^2).

- hadrons (protons) are composite so the CM energy is not fixed and a wide energy spectrum can be explored simultaneously in every collision (however, the production of many low-energy particles results in a complex experimental environment).

LHC is designed to collide two counter-circulating beams of particles in four stations, called *interaction points*: every interaction point corresponds to a detector able to collect and analyse the results of the beams collisions. CMS [33] and ATLAS [34] are general purpose detectors, so they cover the largest possible spectrum of energy allowed by LHC. They are both designed for Higgs boson detection and beyond the Standard Model particles: two "similar but different" detectors allow an useful cross-check of any possible discovery. LHCb [35] is a forward detector designed to study CP violations in the EW theory and the asymmetry of matter and anti-matter in our Universe. Finally ALICE [36] is designed to study collisions between lead ions in order to create an exotic compound, the Quark-gluon plasma (QGP) which is thought to represent the condition of our Universe few seconds after the "Big Bang".

2.2.1 LHC: acceleration complex

The Large Hadron Collider is able to deliver bunches of protons at 7 TeV each (which correspond to a total amount of 14 TeV in the center of mass frame) or heavy ions (typically lead nuclei) at 2.76 TeV/nucleon in the center of mass frame, both used at CMS, ATLAS, LHCb and ALICE. The acceleration mechanism is divided into stages that combine different types of sub-accelerators, used in other CERN's experiments in the past. Everything begins with a *Duoplasmatron*, a device that removes electrons from hydrogen gas; the protons are sent to the first linear accelerator, called *Linac 2* which is 36 m long and it is able to deliver a 30 μ s pulsed beam (corresponding to an electrical current of 180 mA) up to an energy of 50 MeV. After the linac, the bunches of particles enter the first circular accelerator, the Proton Synchrotron Booster (PSB), formed by four superimposed rings for a total circumference of 157 m, that raises the energy to 1.4 GeV injecting then the bunches into the Proton Synchrotron (PS). The PS has a circumference of 628 m and the final output results are particles bunches interspersed by 25 ns (40 MHz) with an energy of 25 GeV delivered to the Super Proton Synchrotron, a 7 km circular accelerator that finally raises the energy up to 450 GeV. The SPS is the last stage before the LHC main ring (fig. 2.1) in which two beams of particles circulating in the opposite directions are formed and then accelerated to the final energy of interest.

2.2.2 LHC: structure overview

In order to span the widest mass region possible, the Large Hadron Collider increases the proton momentum to the limit of the current technology using cyclic electrical. While the particles gain momentum, cyclic magnetic fields are applied to maintain the particles trajectories along the synchrotron ring, according to the formula:

$$p[\text{GeV}/c] = 0.3 \cdot B[\text{T}] \cdot r[\text{m}] \quad (2.1)$$

Since r is fixed by the tunnel dimension, B is the only free parameter which is in turn only constrained by superconducting magnetic present technology. The average magnetic

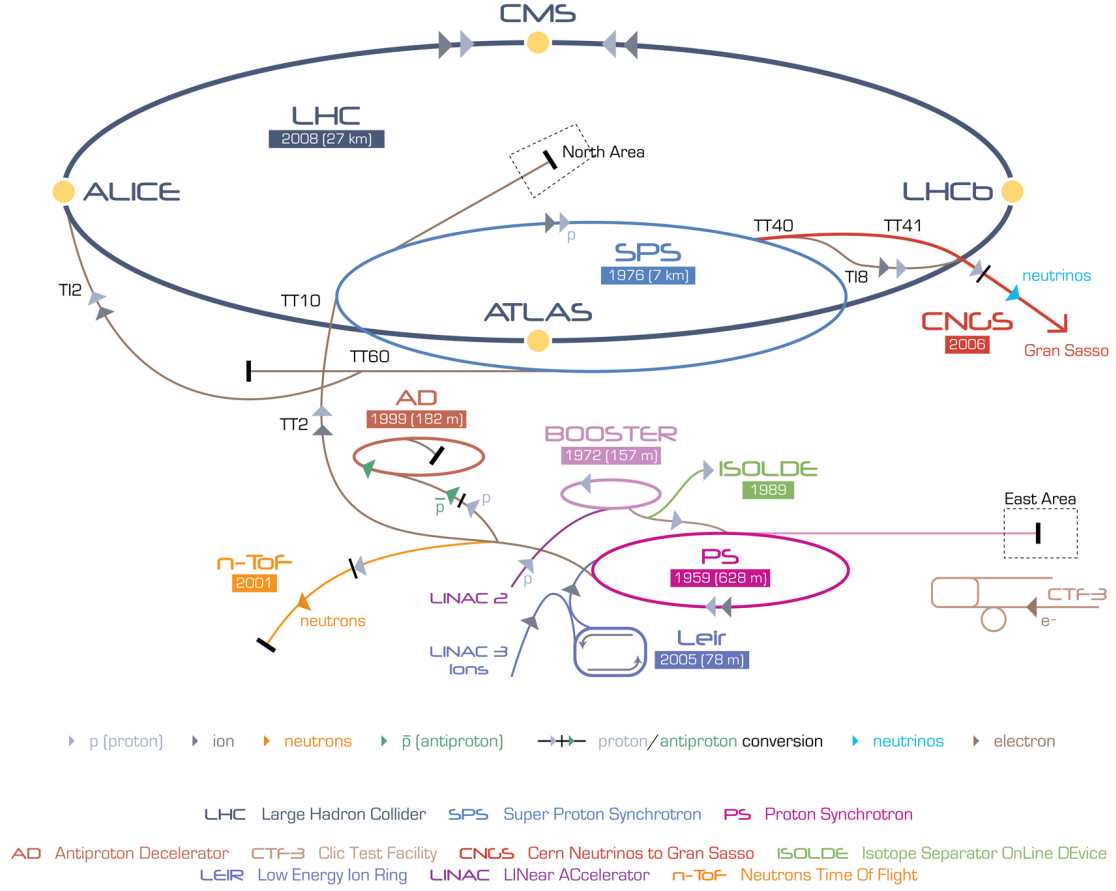


Figure 2.1: The CERN accelerators complex that delivers protons bunches to the LHC experiments [37].

field along the LHC ring is very high since $r = 27$ km, $p = 7$ TeV/c so, using eq. 2.1, $B \approx 5.4$ T.

The bunches of particles travel inside two parallel tubes, under strong vacuum: the tubes are arranged into a semi-circular shape composed by 8 arcs, called *octants*, with a radius of 2.84 km in which are located 1232 bending dipole magnets and eight long straight sections where particles are injected and Radio Frequency cavities (RF) or experiments are hosted. In the following a brief description of the main components of the LHC is discussed:

- **Magnets:** up to 9600 superconducting magnets at 1.9 K are used in the LHC. Each octant contains 14.3 m long dipole magnets to bend particles trajectories along the circumference with a magnetic field of 8.33 T plus three hundred and eighty-six focusing and defocusing quadrupoles to maintain the beam orbit stable. A small amount of higher order multipoles gives additional minor corrections to the beam direction.
- **Vacuum system:** the stability of the protons beams against minor spurious collisions with other particles is achieved circulating the particles into strong vacuum tubes, up to 10×10^{-13} atm. Such challenging low pression is achieved

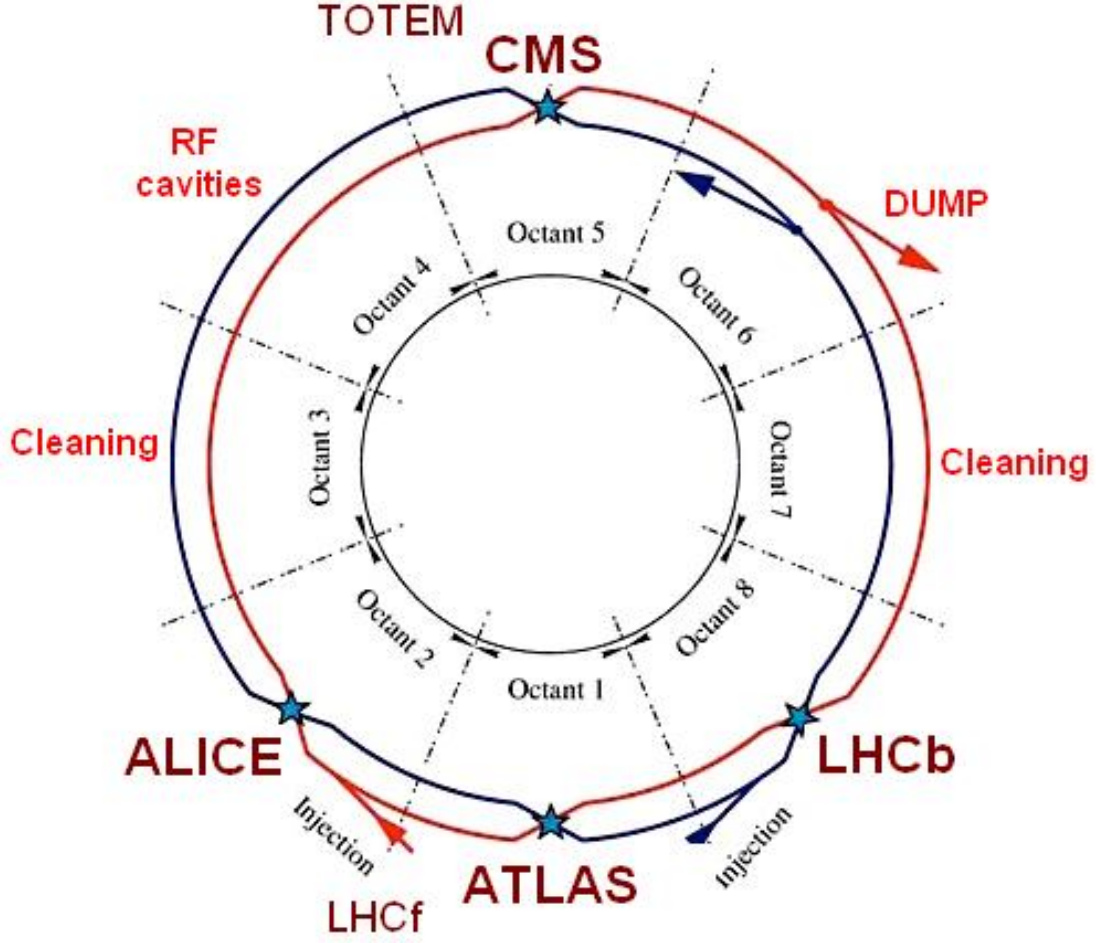


Figure 2.2: The structure of the LHC octants. Straight sections are use for inject/dump the beams. The positions of the experiments and other facilities are highlighted [38].

with the primary vacuum system of LHC. There are two other vacuum systems to insulate the superconductive magnets and the liquid helium supply system.

- **Radio-frequency acceleration system:** located in the cavern of the ALEPH experiment [39], consists in sixteen superconducting radio-frequency cavities with a carrier electromagnetic wave of 400 MHz and a field of 5.5 MV/m. Providing 0.5 MeV/revolution it takes about twenty minutes to obtain 7 TeV energy; after that the RF cavities are used only to supply the energy lost by synchrotron radiation during every revolution (about 7 keV).

LHC is able to spread 2800 bunches of 10^{11} particles along the main ring, each separated by 25 ns with a maximum energy of 7 TeV per beam. The instantaneous luminosity of the collider, defined by eq. 2.2:

$$\mathcal{L} = f \cdot \frac{n_1 \cdot n_2}{4\pi \cdot \sigma_x \sigma_y} \quad (2.2)$$

where f , n_i and σ_i are respectively the frequency, the number of particles and the

transverse dimension of the bunches, is as high as possible in order to increase the chance of spotting new/rare events with low cross sections, according to eq. 2.3:

$$N = \sigma \cdot \mathcal{L} \quad (2.3)$$

where N are the particles produced per unit time in a certain interaction and σ the cross-section.

2.3 The Compact Muon Solenoid experiment

The Compact Muon Solenoid experiment (CMS) [40] is one of the two multipurpose detectors at LHC: it shares with the ATLAS experiment the same scientific goals (searching for the Higgs Boson, Dark Matter and physics beyond the Standard Model) but they adopt different technical solutions, especially for the magnetic system design. Located at Cessy (France) it is placed in a cavern 100 m underground. CMS is built to bear pp collisions at TeV energy scale and, even if is capable of studying many aspects of proton collisions detecting the widest range of neutral and charged particles, it is optimized for precise identification of muons (μ), electrons (e) and (γ). CMS it is one of the largest international scientific collaboration in history [41].

Coordinate system

The coordinate system in CMS is a right-handed Cartesian system (x, y, z) , with the origin fixed in the interaction point inside the detector. The z -axis defines the beamline direction, the y -axis points upward to the surface while the x -axis points to the center of the LHC ring. Since the CMS detector has a cylindrical symmetry it is sometimes useful to use cylindrical coordinates (r, ϕ, θ) (fig. 2.3): the r -coordinate defines the distance from the z -axis, the ϕ -coordinate represents the azimuthal angle in the xy plane and θ is the angle between the particle three-momentum p and the positive direction of the beam axis. In experimental particle physics θ is often replaced by the *pseudo-rapidity*, defined as:

$$\eta = -\ln \left[\tan \frac{\theta}{2} \right] \quad (2.4)$$

In the limit where the particle is travelling close to the speed of light (or when the mass of the particle is negligible), the pseudo-rapidity converges to the *rapidity* usually defined in relativity. There are many reason to prefer the pseudo-rapidity over the polar angle θ since particle production is constant as function of η and differences in rapidity are Lorentz invariant under boost along the beam axis: this turns out to be extremely useful in hadrons colliders where the colliding partons carry different longitudinal momentum fractions.

2.3.1 Structure overview

The Compact Muon Solenoid detector is a 21.6 m long cylinder with a diameter of 15 m and a weight of 12 500 tons. The most impressive constructive element is represented by the huge solenoidal superconducting magnet that surrounds all the inner detectors, with the only exception of the muon chamber on the outside. This magnet provides

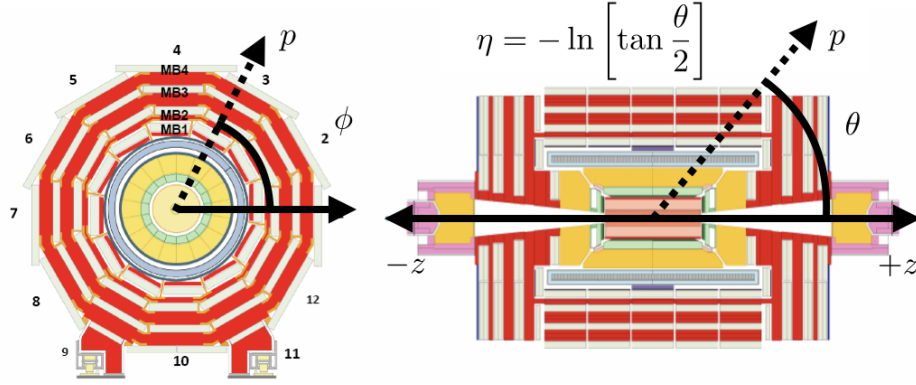


Figure 2.3: A schematic view of the polar coordinates used at CMS [42].

a 3.8 T internal constant magnetic field which is used to bend the charged particles trajectories in order to measure their momenta. The dodecagonal base prism structure of CMS is divided into two main regions: the *barrel* which represents the central part of the prism and two *endcaps* on both side to ensure the maximum tightness. In fig. 2.4 the CMS structure is depicted:

1. **Interaction point:** it is the place where interactions between bunches of particles take place. This thesis is based on data taken in 2015 with an integrated luminosity of 10^{34} fb^{-1} : the consequence of a so high collision rate is reflected in a large number of sumperimposed interactions at the same time (this effect is also called *pile-up* and hampers the reconstruction of the particles tracks). To overcome the huge pile-up expected at high luminosity CMS provides a good time resolution and high granularity of the system.
2. **Silicon tracker:** is the inner part of CMS and it extends from the interaction point to $|r| < 1.2 \text{ m}$ covering a pseudo-rapidity range $|\eta| < 2.5$. Over an active area of 215 m^2 are distributed *silicon pixel vertex* detectors surrounded by *silicon microstrip* detectors able to reconstruct particles tracks and vertices.
3. **Electromagnetic calorimeter:** it extends over $1.2 \text{ m} < r < 1.8 \text{ m}$, covering an angle $|\eta| < 3$. It is composed by scintillating crystals of lead-tungstate (PbWO_4) able to detect the trajectories and energies of photons and electrons.
4. **Hadronic calorimeter:** it extends over $1.8 \text{ m} < r < 2.9 \text{ m}$ covering a large pseudo-rapidity range up to $|\eta| = 5$. It is a sampling calorimeter composed by brass layers alternated with plastic scintillators able to detect charged hadrons.
5. **Superconducting solenoidal magnet:** it generates an internal uniform magnetic field of 3.8 T along the beam direction, able to deflect the particles trajectories to allow momentum measurements. The solenoid surrounds all the previous sections and it extends from $2.9 \text{ m} < r < 3.8 \text{ m}$ over $|\eta| < 1.5$.
6. **Muon system:** it is embedded in the magnet return yoke so it is surrounded in a residual 1.8 T magnetic field. It extends from $4 \text{ m} < r < 7.4 \text{ m}$ over $|\eta| < 2.4$ and it is composed by *drift tubes* in the barrel and *cathode strip chambers* in the endcaps.

Additional *resistive plate chambers* are used in both regions. The muon system is able to reconstruct very well tracks produced by muons that pass through it.

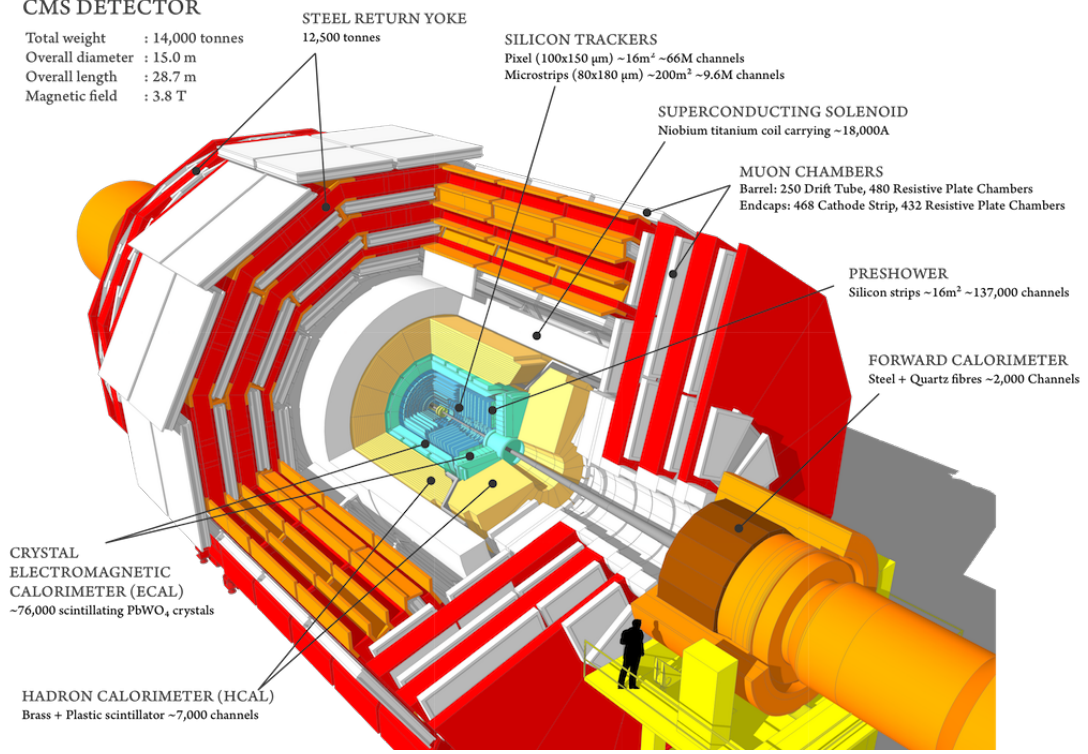


Figure 2.4: A schematic view of the structure of the CMS experiment from [43].

2.3.2 Tracker

The tracker [44] is the detector closest to the beam collision point. Its main task consists in the reconstruction, with the highest possible resolution, of the trajectories of charged particles originated in the interaction point and the identification of the positions of secondary vertices in events containing short mean-life particles (such as b-quarks hadrons). This is not an easy task since at nominal luminosity (approximately 10^{34} fb^{-1}) there are, on average, 1000 tracks every 25 ns. To make the pattern easier to reconstruct, the tracker was built with the following characteristics:

- **Low occupancy:** is achieved providing the highest possible granularity in $\Delta\eta \times \Delta\phi$ enhancing the resolution of the detector.
- **Large redundancy of measured point:** reducing the ambiguity in assigning hits to the corresponding track.
- **Limited amount of material:** to avoid compromising the measurements, since an excessive amount of material hampers the accuracy of the hits points positions in the tracker sections both from *multiple scattering* phenomena and possible *electron-positron conversion* from high energy photons entering the device.

The CMS tracker has a cylindrical shape, it is 6 m long with an outer diameter of 2.6 m. It covers both the barrel and endcaps regions (to have maximum tightness) and it is optimized for $12 \div 14$ hits per track. The tracker can be divided into two subdetectors, both made by silicon. The *silicon pixel vertex* detector is closer to the collision point, it is essential for reconstruction of primary and secondary vertices and it is composed by three concentric cylindrical sections containing 768 modules: every module contains several units of highly integrated and segmented silicon sensors that provide a resolution in the hit reconstruction of $10 \mu\text{m}$ – $15 \mu\text{m}$ in the barrel and $15 \mu\text{m}$ in the endcaps. The *silicon microstrip* detector surrounds the pixel vertex detector and consists of 15400 modules or *microstrips* composed by sensors stuck on a carbon fiber support. The sensors are distributed into four main regions: the *TIB* (tracker inner barrel) and the *TOB* (tracker outer barrel) in the barrel region and the *TID* (tracker inner disk) and the *TEC* (tracker endcap) in the endcaps. The general structure of the tracker is depicted in fig. 2.5

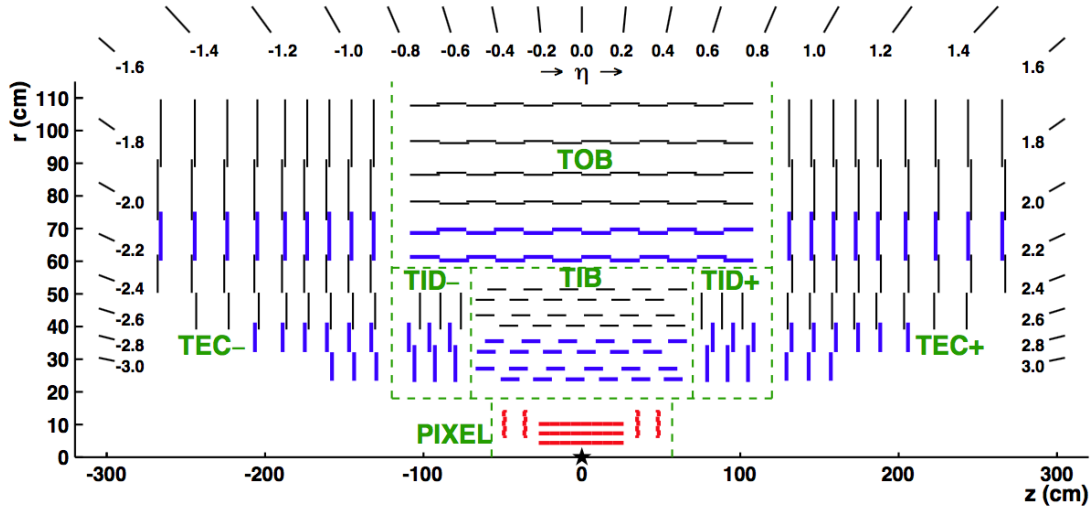


Figure 2.5: The CMS tracker of the CMS experiment [44].

Once every trajectory is properly associated to the corresponding charged particle, since the tracker is surrounded by the magnetic field inside CMS, it is possible to reconstruct the particle momentum from the track curvature using eq. 2.1.

2.3.3 Electromagnetic calorimeter

The CMS Electromagnetic Calorimeter (ECAL) is the subdetector that measures accurately the energy of electrons and photons [46]. It is a homogeneous calorimeter with cylindrical geometry (fig. 2.6): the barrel contains 61200 scintillating crystals of lead-tungstate (PbWO_4) while the endcaps consist in 15000 crystals each for a total of over 76000 crystals. The barrel has an inner radius of 129 cm, a length of 630 cm and it is composed by 36 *super-modules* that cover a region of $|\eta| < 1.479$. Every super-module includes 20×85 crystals on the ϕ - η plane, sub-divided into four *modules* along the η direction. Every module is again segmented into *sub-modules* consisting in 5×2 crystal

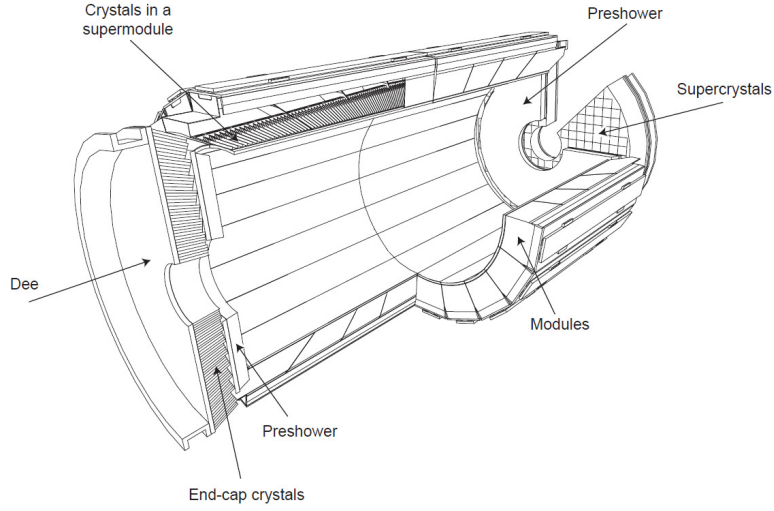


Figure 2.6: ECAL 3D view with the main components of the detector apparatus [45].

arrays mounted on a glass fiber structure. In the barrel region the crystals are cut in a truncated pyramidal shape with a total length of 23 cm which corresponds to $24.7 X_0$, a frontal and rear area of $22 \times 22 \text{ mm}^2$ and $26 \times 26 \text{ mm}^2$ and, finally, a granularity of $\Delta\eta \times \Delta\phi = 0.0175 \times 0.0175$ (1°). Crystals are grouped into 5×5 matrices called *trigger towers* and, to avoid cracks aligning with particle trajectories, crystal axes are tilted by 3° with respect to the direction of the interaction point. Following the η direction the barrel structure stops at $|\eta| = 1.479$ allowing some service cables to exit the calorimeter area as can be seen in fig. 2.7. To ensure the maximum tightness, two endcaps close the lateral areas of the detector covering a pseudorapidity angle of $1.479 < |\eta| < 3$. They consist in semicircular halves of aluminum called *dees*: each dee is composed by 3662 crystals shaped differently with respect to the barrel since they are 22 cm long with a frontal and rear area of $28.6 \times 28.6 \text{ mm}^2$ and $30 \times 30 \text{ mm}^2$. Crystals are arranged using the $x - y$ symmetry into 18 *super-crystals*, each of them containing 5×5 elementary units.

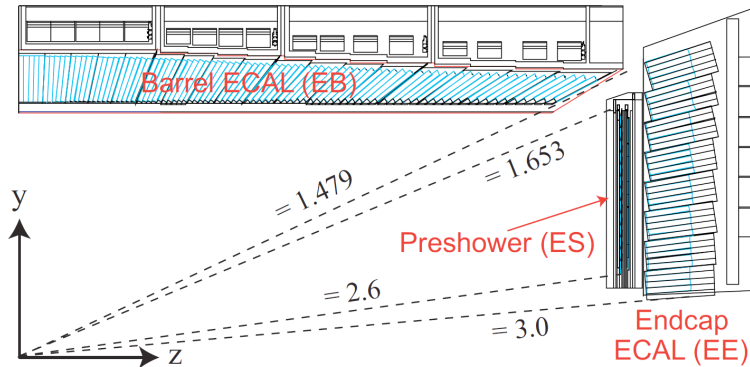


Figure 2.7: ECAL side view [45].

In order to separate the electromagnetic showers formed by primary photons from

those from primary π^0 's a *preshower* detector is placed on the inner side of every endcaps covering an angle between $1.653 < |\eta| < 2.6$ (cf. fig. 2.7). The preshower is a sampling calorimeter made by two disks of lead converter (with a length equal to, respectively, $2X_0$ and $1X_0$ which start the electromagnetic shower of the incident particle. In order to identify the shower profile and measure the released energy lead is alternate with two layers of thick silicon microstrip detectors ($320\mu\text{m}$), as in the tracker.

The ECAL calorimeter is entirely made of lead tungstate scintillating crystals and the choice of PbWO_4 as the main active material lies on various advantages:

- The PbWO_4 high density ($\rho = 8.3\text{ g/cm}^3$), short radiation length ($X_0 = 0.89\text{ cm}$) and reduced Moliere radius (which gives the scale of the transverse dimension of the fully contained electromagnetic showers initiated by an incident high energy electron or photon) of $R_m = 2.2\text{ cm}$ are good features in order to build a compact and high granulated calorimeter.
- The PbWO_4 decay scintillation time (approximately 15 ns) allows to collect about 80% of the emitted light during 25 ns , which is the time between two consecutive beams interactions at LHC.
- The good intrinsic radiation hardness lets the material to operate for years in the LHC environment with modest deterioration in performances due to radiation damages.

Among this features, the lead tungstate scintillation crystals present one, important, disadvantage: being inorganic scintillators the light yield is low, corresponding approximately to 10 photons/MeV, so an amplification of the signal is mandatory. To accomplish this task photo-multipliers collect the scintillation light of every crystal: *silicon avalanche photodiodes* (APD) [47] are used in the barrel while *single stage photomultipliers* (VPT) [48] in the endcaps.

The CMS ECAL functions as a precision calorimeter and its energy resolution is optimized to the relevant physics at the LHC. The most demanding physics channel for this detector is the two-photon decay of an intermediate-mass Higgs boson: since the background is large, the signal width must be resolved with the best resolution. For the range of energies relevant to the Higgs two-photon decay (25 GeV – 500 GeV) the energy resolution can be parametrized as [46]:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c = \frac{2.8\%\text{GeV}^{1/2}}{\sqrt{E}} \oplus \frac{12\%\text{GeV}}{E} \oplus 0.3\% \quad (2.5)$$

where E is in GeV, a is the stochastic term, b the noise term and c the constant term. The stochastic term is dominant at low energies and accounts for the statistical fluctuations in the number of photoelectrons generated and collected, the noise term, which gives a contributions that vary with the pseudo-rapidity and the luminosity, includes the electronic and pile-up noise and the constant term is dominant at high energies and involves different phenomena like the stability of the operating conditions, the presence of dead material and much more. ECAL provides one of the best energy resolution so far achieved in electromagnetic calorimetry.

2.3.4 Hadronic calorimeter

The CMS hadronic calorimeter (HCAL) [49] together with ECAL makes a complete calorimetric system for the jet energy and direction measurement. It is built as an hermetic sample calorimeter and its tightness maximizes particle containment. HCAL covers an angle of $|\eta| < 5$ and it is divided into four sub-detectors, as can be seen in fig. 2.8: the *hadronic barrel calorimeter* (HB) and the *hadronic endcap calorimeter* (HE) placed inside the solenoidal magnetic field, the *hadronic outer calorimeter* (HO) and the *hadronic forward calorimeter* (HF) in the return yoke of the solenoid. To ensure the best energy containment possible the amount of absorber material inside the magnetic coil is maximized, providing a good measurement of the features of non-interacting particles via the extraction of the missing transverse energy¹ E_T^{miss} . Since the inner part of HCAL is placed inside the solenoidal magnetic field the absorber is made of brass which is a non-magnetic material.

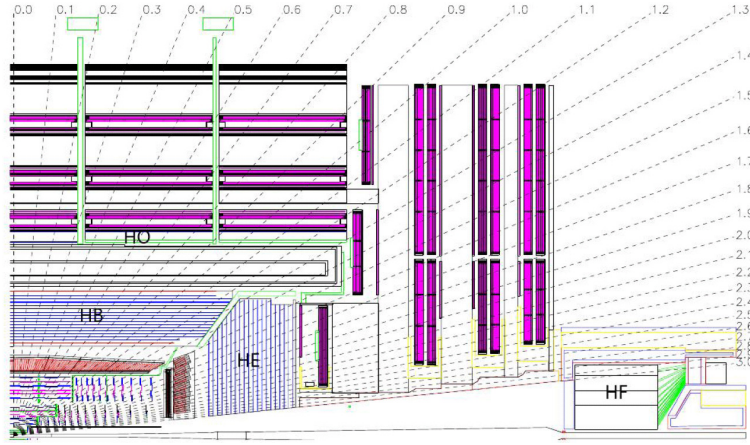


Figure 2.8: HCAL experimental apparatus side view [50].

The structure of HB and HE alternates brass plates (absorber) with plastic scintillators (active material) linked to *hybrid photodiodes photodetector* (HPD) via *wavelength-shifter* optical fibers. HB is 9m long, it covers an angle of $|\eta| < 1.4$ surrounding the ECAL barrel and it is made by 2304 *calorimetric towers*, each of them made of 15 brass layers parallel to the beams and 17 layers of plastic scintillators connected to the same photodetector to improve the output signal. HE covers an angle of $1.3 < |\eta| < 3$ overlapping partially HB and it is radially divided into 14 rings containing 2304 calorimetric towers similar to the ones used in the barrel section. HO is used to improve the measurement of higher energetic showers that can pass all through HB (since the dimensions of the last one are constrained by the solenoid structure). It is placed outside the solenoid and consists in several layers of plastic scintillators for an effective thickness of $10 \lambda_I$ and the same granularity of HB, in order to have a one-to-one correspondence between the two. Finally the HF, placed about 11.2m from the collision point guarantees the coverage of the system up to $|\eta| = 5$: its structure, consisting in two cylinders of

¹The missing transverse energy refers to energy which is not detected in a particle detector (since is carried by particles that escape the detector without being detected) but is expected due to the laws of conservation of energy and momentum.

iron absorbers and quartz fibers as active material, is optimized for very forward jets events (commonly expected in interesting physical events such as Higgs boson and SUSY particles). The energy resolution of HCAL can be seen in eq. 2.6:

$$\begin{aligned}\frac{\sigma_E}{E} &= \frac{90\% \text{GeV}^{1/2}}{\sqrt{E}} \oplus 4.5\% \text{ (barrel/endcaps)} \\ \frac{\sigma_E}{E} &= \frac{172\% \text{GeV}^{1/2}}{\sqrt{E}} \oplus 9\% \text{ (forward)}\end{aligned}\tag{2.6}$$

2.3.5 Superconducting solenoid

In order to bend particles the trajectories, the CMS experiment uses a superconductive cylindrical solenoid [51] cooled at -268.5°C . This is one of the biggest superconducting Niobium-Titanium solenoid ever built and contains the majority of the CMS subdetectors with the only exception of the muon chambers and part of the hadronic calorimeter. The solenoid stores about 2 GJ of energy, able to provide a magnetic field of 3.8 T (10^5 times the Earth magnetic field). The magnet return yoke embeds the muon chambers, surrounded by a residual magnetic field of 1.8 T.

2.3.6 Muon chambers

The muon chambers [52] detect and measure high p_T muons in combination with the tracker data. Embedded in the magnet return yoke, they exploit the 1.8 T magnetic flux and are divided into three independent sub-systems, as can be seen in fig. 2.9:

- **Drift tubes:** the drift tubes (DT) are used in the barrel region ($|\eta| < 1.2$) since the particle flux is low ($< 10 \text{ Hz/m}^2$) and the magnetic field is homogeneous. They are arranged into five iron wheels, each containing four concentric rings *stations*, made of twelve DT chambers. The basic detector element of DT is a rectangular drift tube cell with a transverse size of $13 \times 42 \text{ mm}^2$ and a length of 2 m–4 m. The volume is filled with a gas mixture (Argon and Carbon dioxide). The position resolution is about $100 \mu\text{m}$ both in $r - \phi$ and $r - z$ plane.
- **Cathode strip chambers:** the cathode strip chambers (CSC) are used in the endcaps ($0.8 < |\eta| < 2.4$), mainly due to the high particle rate and magnetic field (3.5 T) in this region. The CSC are multi-wire proportional chambers, filled with Ar and CO_2 as active gas and CF_4 as a quencher, that provides a drift path shorter than DT. They are arranged into four superimposed disks called *stations*, each formed by two rings divided into eighteen or thirty six CSC (consisting in six layers of sensitive wires each) with a resolution of $80 \mu\text{m}$ – $85 \mu\text{m}$ in the $\eta \times \phi$ plane and a precision of 0.5 cm in the r direction.
- **Resistive plate chambers:** they are placed both in the barrel and endcaps for robustness and redundancy. They are gaseous detectors with a coarse spatial resolution (since they work in avalanche mode), but good time measurements: they provide precise bunches crossing identification to the muon trigger system. The RPC are made of four bakelite planes coated with graphite as electrodes and two gas gaps between them of 2 mm. The central part of every chamber is

equipped with insulated aluminum strips in order to collect the signals generated by crossing particles.

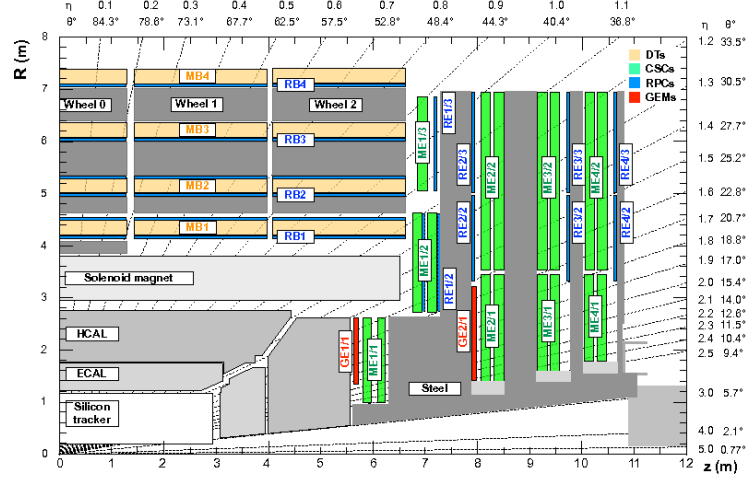


Figure 2.9: The structure of the muon detector system at CMS [52].

2.3.7 Data triggering system

The LHC provides proton-proton or heavy-ion collisions at a high rate: the protons bunches are interspersed by 25 ns corresponding to an event rate of 40 MHz, impossible to handle for the currently available technology. In order to analyze and store such stream of data a data reduction is applied via a triggering system [40]. The CMS triggering system is divided into two main stages: the *Level-1 trigger* (L1) [53] which is composed by custom-designed largely programmable electronics and the *High Level Trigger* (HLT) [54], a software system implemented in a filter farm of one thousand commercial processors. The output of the L1 stage corresponds to a data stream of 30 kHz–100 kHz while the rate reduction of the two systems combined achieve 10^6 . The L1 trigger uses coarsely segmented data from calorimeters and muon system to take decisions while holding high-resolution data in pipelined memories. It is divided in three stages, as we can see in fig 2.10:

1. **Local:** in the first stage a *Trigger Primitive Generator* (TPG) takes decisions on the events based on energy deposit in the calorimeter and hit patterns in the muon chambers.
2. **Regional:** the regional step combines local informations and use pattern logic to determine ranked and sorted trigger objects (electrons/muons) in limited spatial regions. The rank is determined as a function of the energy/momentum and quality of the event.
3. **Global:** *Global Calorimeters* and *Global Muon System* determine the highest rank calorimeter and muon objects and deliver them to the *Global Trigger* that accepts or rejects the whole event for the HLT. The global decision is based both on algorithm calculations and readiness of the sub-detectors and the DAQ.

L1 has to analyze every bunch crossing so the allowed time between a given bunch crossing and the distribution of the trigger decision to the detector front end microelectronics should be small. In the CMS triggering system it corresponds to $3.2\mu\text{s}$ but, using additional pipelines to store temporarily the data, the entire process is quasi-deadtime-free. If L1 decides to accept the event, the full read-out data enters the HLT trigger that can perform complex calculations to further analyze the result. A precise definition of how the HLT works is not easy since it evolves with time. More informations can be found in [55].

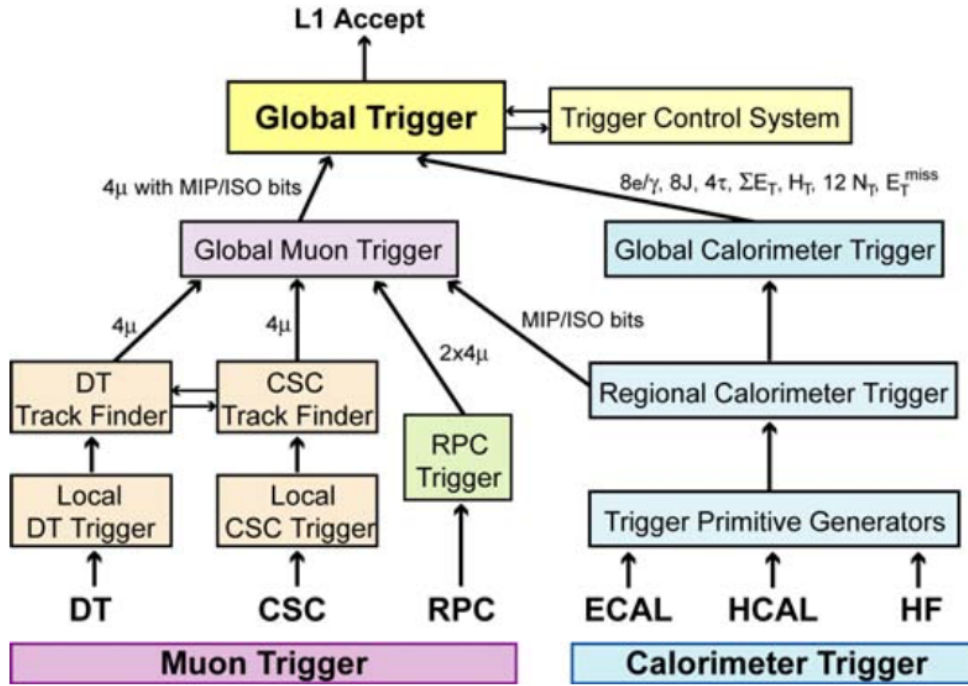


Figure 2.10: The structure of the CMS L1 triggering system [40].

3 | Electron reconstruction and energy calibration

The Standard Model Higgs boson decay $H \rightarrow ZZ^* \rightarrow 4e$ signal consists of two pairs of same flavours, opposite-sign, well-identified and isolated leptons, namely e^+e^- or $\mu^+\mu^-$, compatible with a ZZ system where one or both the Z -bosons can be off-shell, appearing as a narrow resonance on top of a smooth background in the four-electron invariant mass distribution, as shown in fig. 4.1. The main background sources to the Higgs signal in this channel are the SM ZZ production, with smaller contributions from other diboson processes, such as ZW and WW , single bosons with hadronic activity that mimic the lepton signature and top-quark-pair events. A disadvantage of the $H \rightarrow ZZ^* \rightarrow 4e$ decay is its low branching fraction of $\mathcal{O}(10^{-4})$ for $m_H = 125(200)$ GeV [56]. Due to the challenging kinematics and background conditions, to maximize the sensitivity to an Higgs boson in the mass range 110 GeV–1000 GeV, a big effort should be put on the electron reconstruction in CMS on maximizing the efficiency of lepton reconstruction and identification to enhance the sensitivity to a low-mass Higgs boson. Furthermore, the determination of the properties, as the boson mass, rely on the precision of the energy measurement of the decay leptons. This work is focused on the determination of the electron energy scale and its effect on the Higgs boson mass measurement.

In this chapter the electron energy and track reconstruction algorithms are presented. The corrections to be applied to the reconstructed energy are described.

3.1 Electron reconstruction in CMS

The reconstruction of electrons in CMS combines informations from the pixel detector, the silicon microstrips and ECAL. Different subdetectors are used together with the electromagnetic calorimeter in order to improve the quality of the reconstructed object since various effects hamper the measurement of the electron energy:

- The electrons can interact, before reaching the calorimeter, with the amount of tracker material distributed in front of it (cf. sec. 2.3.2), losing a fraction of their energy via *bremsstrahlung radiation*.
- The strong magnetic field aligned with the collider beam axis (cf. sec. 2.3.5) bends the electrons while not affecting the bremsstrahlung photons, causing a significant energy spread of the energy deposits in the crystals in the ϕ direction.

The combination of the previous effects introduces, in general, non-Gaussian contributions to the event-by-event fluctuations of the calorimeter and tracker measurements so the ECAL clustering system is specifically designed to take into account the ϕ spread, collecting all the radiated energy and associating it to the related primary electron via the electron track reconstruction. Different *classes* of electrons are defined in order to better disentangle the sources of partial energy containment in the cluster, adapt the energy scale corrections and estimate the associated errors.

3.1.1 Clustering of the electron energy in ECAL

When a high energy electron interacts with ECAL produces electromagnetic showers, depositing energy in several crystals of the calorimeter. For a single electron, most of the energy is collected in a small number of crystals due to the ECAL good lateral energy containment. Some studies and several tests were performed on ECAL super-modules before installing the calorimeter in the whole CMS detector system: as a general result, electrons with an incident energy of 120 GeV, impinging at the center of a crystal, deposit approximately 97% of their energy in a 5×5 crystal window [46]. The 5×5 crystal matrix allows for best measurement performances for electrons provided that *local containment corrections* are applied to account for the variation of the measured energy in function of the position of the particle beam with respect to the borders of the matrix in the case of a ECAL module without material in front and without magnetic field. The situation in the CMS full experiment is much more complicated for the average electron, due to the before-mentioned effects: the electron energy spray over the ϕ direction can be very high if integrated over the full particle trajectory as can be seen in fig. 3.1, and several useful informations can be extrapolated:

- About 35% of electrons radiate more than 70% of their initial energy before reaching ECAL.
- In 10% of the cases the radiation represents more than 95% of the initial energy.

Since it is essential to collect the majority of the bremsstrahlung photons, different algorithms have been developed, optimized using simulations, adjusted during data taking periods and used in the barrel and endcaps regions of ECAL. Two *clustering algorithms* were used in CMS during Run 1: the *hybrid* algorithm, in the barrel region, and the *multi- 5×5* algorithm in the endcaps. For every clustering step two directions, ϕ and η , are defined relative to the center of CMS. These two algorithms are described below for simplicity, since nowadays they are surmounted by a Particle Flow clustering of the electron tracks.

The hybrid algorithm

The hybrid algorithm [46] exploits both the geometry of the ECAL barrel and the properties of the electron shower shape. It collects the released energy in a small window in η (since the energy spread in this direction is usually negligible) and an extended window in ϕ . The starting point is the seed crystal which is the one containing most of the energy in any considered region of the barrel: if $E_{T,seed} > E_{T,seed}^{min}$ then the crystal is selected. Then arrays of 5×1 crystals in $\eta \times \phi$ are added around the seed in a range

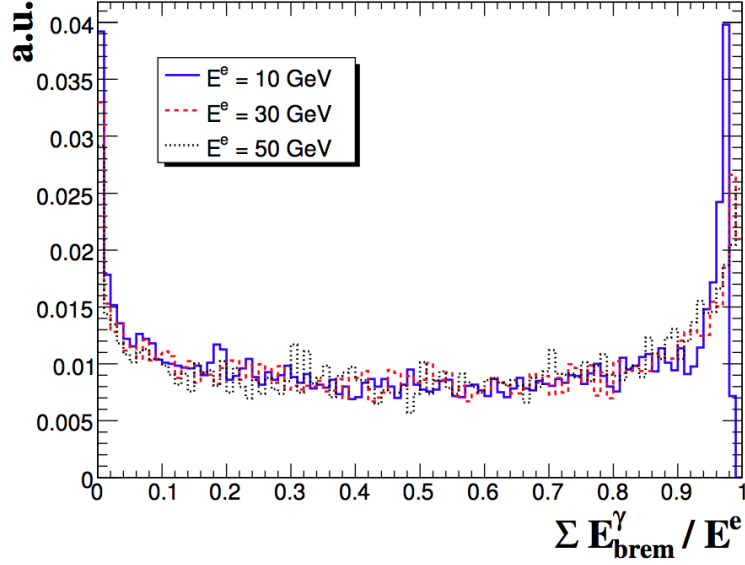


Figure 3.1: An example of the radiation effects: the simulated distribution of the fraction of the initial energy radiated by electrons before reaching ECAL for different initial particle energies 10 GeV, 30 GeV and 50 GeV. The histograms are the convolution of the bremsstrahlung spectrum, the finite path to reach ECAL and the finite initial energy [57].

of N_{steps} crystals in both ϕ directions if their energy are $E_{array} > E_{array}^{min}$. Contiguous arrays are grouped into clusters and each cluster is required to have a seed energy $E_{seed-array} > E_{seed-array}^{min}$. Finally clusters are collected in a window fixed in η and spreaded in the ϕ direction into superclusters (eg. cluster of clusters). In fig. 3.2 the cluster and supercluster construction steps are shown.

The multi- 5×5 algorithm

The hybrid algorithm is not used in the endcaps regions because here crystals are not arranged in a $\eta \times \phi$ geometry. The same idea of collecting the energy of the neighbouring crystals is then implemented in a different way, using the multi- 5×5 algorithm. The construction of a cluster starts from a seed crystal, selected as the the local maximum energy deposit relative to its four direct neighbours, chosen in a swiss-cross pattern. The seed crystal must then fullfill the condition $E_{T,seed} > E_{T,EEseed}^{min}$. When a seed crystal is found, the algorithm collects energy around it in clusters of 5×5 crystals including only the ones that do not already belong to a cluster. To allow closely overlapping showers (mainly due to bremsstrahlung) to be collected, the outer sixteen crystals of the 5×5 matrix can seed another cluster. In this case, the "shared" crystals are assigned to the cluster with the largest seed E_T . In order to produce the final supercluster, recovering all the bremsstrahlung energy, clusters are grouped into superclusters if the total energy of each cluster satisfy the condition $E_{T,cluster} > E_{T,cluster}^{min}$ in a rectangular window in the η direction of $\pm \eta^{range}$ and the ϕ direction of $\pm \phi^{range}$. This procedure is performed in descending order with respect to each supercluster energy and it is possible that only

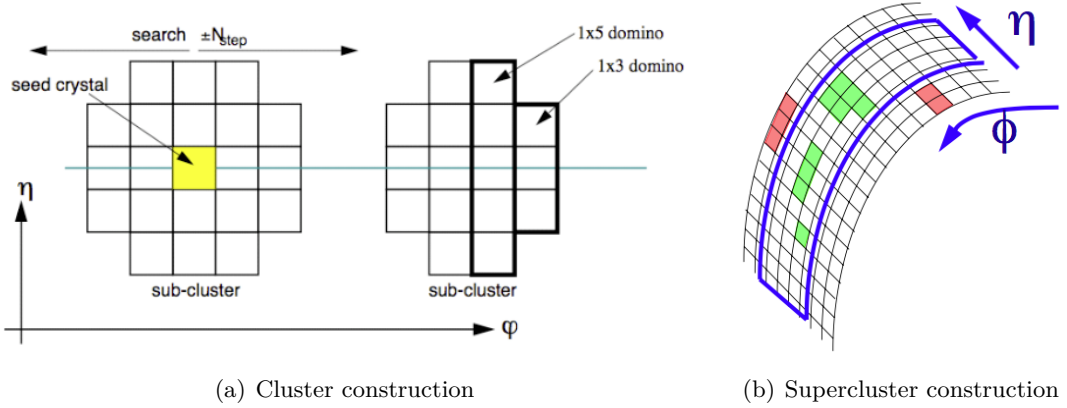


Figure 3.2: Individual steps of the hybrid algorithm [58].

a single cluster belongs to the supercluster.

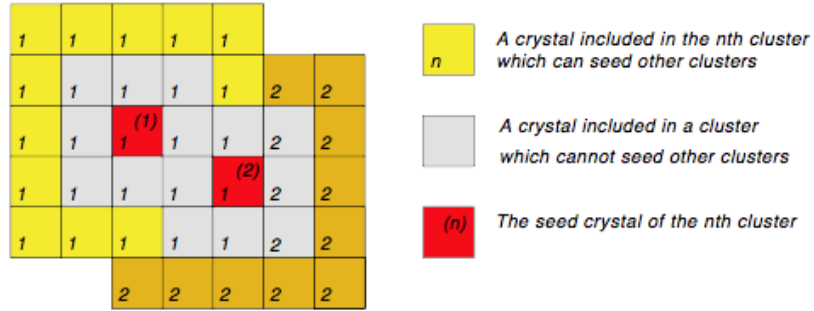


Figure 3.3: Individual steps of the multi- 5×5 algorithm. Two overlapping clusters are shown. Crystals in yellow can seed another cluster if they satisfy the seed energy condition [59].

In the endcaps regions electrons will deposit part of their energy also in the preshower detector, which covers the region $1.6 < |\eta| < 2.6$ (cf. 2.3.3). In order to properly collect this residual contribution, the energy weighted positions of all clusters belonging to a supercluster are extrapolated to the plane of the preshower with the most energetic cluster used as a reference point. The maximum distance between the clusters and their reference points are used to define the preshower clustering range along η and ϕ : the preshower energy deposits in these ranges are then added to the supercluster energy. An example of the multi- 5×5 clustering algorithm via the swiss-cross pattern is shown in fig. 3.3

In tab. 3.1 the threshold values of parameters used in the hybrid superclustering algorithm in the barrel and the multi- 5×5 algorithm in the endcaps are shown.

Barrel		Endcaps	
Parameter	Value	Parameter	Value
$E_{T,seed}^{min}$	1 GeV	$E_{T,EEseed}^{min}$	0.18 GeV
$E_{seed-array}^{min}$	0.35 GeV	$E_{T,cluster}^{min}$	1 GeV
E_{array}^{min}	0.1 GeV	η^{range}	0.07
N_{steps}	17(≈ 0.3 rad)	ϕ^{range}	0.3 rad

Table 3.1: Threshold values for the clustering algorithm [60].

3.1.2 Electron track reconstruction

The next step, after clustering the energy into ECAL, is the electron track reconstruction, a procedure able to fully determine the track of every electron associated with a calorimetric deposit. This process is divided in four sequential modular components [57]:

1. Initial track seed are looked for with a **Seed Generator** (SG).
2. The **Trajectory Builder** (TB) constructs outwards all the possible trajectories for every given seed.
3. The **Trajectory Cleaner** (TC) is able to solve ambiguities among the possible trajectories previously generated and a maximum number of tracks candidates are kept.
4. The final fit of the track is performed with the **Trajectory Smoother** (TS), that uses all the collected hits to estimate the track parameters at each layer of the pixel detector using a backward fit.

Electrons tracks can be reconstructed in the full tracker using the *Kalman Filter* (KF) algorithm [53], the standard reconstruction procedure used for all charged particles in CMS. However electrons lose energy primarily through bremsstrahlung rather than ionization so large energy losses are expected. Since the energy loss distribution is highly non-Gaussian, the standard KF (optimal when all variables have Gaussian uncertainties) is not appropriate. In addition, large radiative losses in the tracker material lead to a reduced hit collection efficiency (since hits are lost when the change in curvature is dramatic) and also hamper the estimation of track parameters. For these reasons, a dedicated tracking algorithm, consisting in the combination of two techniques that make use of informations not only from the tracker but also from the ECAL, have been developed for the seeding and bulding steps, as well as for the final smoothing. The main goals are an efficient collection of the hits while preserving an optimal estimation of the tracks parameters.

Seed generation

The seed generation is an important step in electron track reconstruction since its performances greatly affect the reconstruction efficiency. In order to build a track, a seed is created when two or three hits compatible with a given beam spot are found in the pixel detector. In order to tame the many possible hits combinations, the search

for a seed is restricted to a region compatible with a supercluster in the ECAL. Two similar approaches are possible: a simple *regional restriction*, relying on the observation of an ECAL supercluster (with the disadvantage of providing a severe fake rate) or a powerful *ECAL-based seeding*, that starts from a supercluster energy deposit position to estimate the electron trajectory in the first layers of the tracker, with the advantage of increasing the purity of the sample of candidate electron tracks. The results of another complementary algorithm are also usually combined with the ones from the ECAL-based seeding: using a *tracker-based seeding*, tracks are reconstructed via a general algorithm developed for tracking charged particles, extrapolated towards ECAL and then matched with a supercluster. This algorithm complements the seeding efficiency, especially for low- p_T or nonisolated electrons, as well as for electrons impinging in the barrel-endcap transition region [60].

The ECAL-based seeding algorithm is based on the assumption that the energy weighted average impact point of the electron and associated emitted bremsstrahlung photons, calculated using the informations collected by the ECAL clustering algorithm, coincides with the impact point measured for a no-emitting particle. This algorithm, successfully developed for robust applications at the High Level Trigger (cf. 2.3.7), works well if the radiation emitted is properly collected, the electron track is well defined and the p_T is larger than the HLT threshold. Hits in the pixel layers are predicted by propagation of the energy weighted mean position of the supercluster backward to the pixel layers: the intersection of the most probable helix trajectory with the most inner layer is looked within a loose $\Delta\phi$ and Δz interval. If no hit is found the first hit is then searched in the next to innermost layer and finally taken as the seeding hit. The backward propagation procedure is performed both for positive and negative charged particle hypothesis and the initial supercluster is selected in order to limit the possible number of misidentified seeds using the following requirement: $E_T^{SC} > 4 \text{ GeV}$ and $H/E_{SC} < 0.15$ where E_{SC} is the energy of the supercluster and H the sum of the HCAL tower energies in a cone of radius $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.15$ around the electron direction [60]. Once a seed is found the position of the primary track vertex is updated and the predicted trajectory propagated to look for other hits in the next pixel layers.

On the other hand, the tracker-based seeding procedure combines pairs or triplets of hits with the vertices obtained from pixel tracks reconstructed using a *Kalman filter algorithm* [44] when the bremsstrahlung is negligible or via a *Gaussian sum filter* (GSF) [61] in the other case: only a subset of the seeds leads eventually to a track.

The seed collections obtained by using these two methods are merged and used to initiate the electron track finding: the overall efficiency of the combined algorithms is predicted $> 95\%$ for simulated electrons from Z-boson decay [60].

Trajectory building

Starting from the seed, the electron trajectory is built and then fitted [62]. The track building algorithm is based on a combinatorial KF method which, for each seed, searches compatible hits on the next silicon layers until the last hit is found. During this process the electron energy loss is modelled using a Bethe-Heitler model. The trajectory at each layer is then defined as the weighted mean between the predicted state and the

effectively measured hit: the compatibility of the two is defined in terms of a χ^2 test and, at this stage, no specific χ^2 cuts are applied, in order to follow electrons trajectories in case of severe bremsstrahlung and to maintain a good efficiency. However, only the two tracks with smallest χ^2 are kept. Actually, if during track reconstruction several hits are found compatible with those predicted in a layer, many trajectories are grown in parallel, and at most one missing hit and a minimum of five hits are required to build a track. Finally, to avoid including hits from converted bremsstrahlung photons (photons emitted by the primary electron that interacts with the tracker material via pair production), an increased χ^2 penalty is applied to trajectory candidates with one missing hit. Simulated and experimental data from electrons originated from Z-bosons decays are used to validate the consistency of KF algorithm in hits collections: while the general behaviour is well reproduced, disagreement are observed between data and simulation mainly due to an imperfect description of the material of the tracker sensors in the simulation.

Track parameter measurement

Once hits are collected, to estimate the track parameters a GSF is performed. The energy loss in each layer is approximated by a weighted combination of Gaussian distributions. Two estimates of track properties, which gives quite different results, are usually exploited on each layer:

- The **weighted mean** of all components is an unbiased average of the momentum of the track. It provides the best sensitivity to the momentum change along the track due to radiation emission: when taking the mean, the distribution for the reconstructed transverse momentum at the vertex follows a Gaussian distribution with a slightly more pronounced tail towards high p_T values.
- The **most probable value** gives a distribution of the electron reconstructed p_T well peaked at the correct value (resulting from MC simulations) but with significantly extent at small p_T values (a characteristics of bremsstrahlung loss).

As shown in fig. 3.4, the most probable value peaks at the generated value and has a smaller standard deviation for the core distribution then it is better suited to obtain an estimation of the most probable track parameters least affected by bremsstrahlung emission. The track hits in the tracker are collected up to the ECAL so a good estimate of the track parameters at the calorimeter entrance is possible: this give a chance to improve both the estimation of the bremsstrahlung radiation and of the association between track and cluster.

Cluster-track matching

Due to bremsstrahlung emission, the matching between the track and the supercluster is done using the track parameters at vertex. Different approaches are used for ECAL-based and tracker-based seed electrons: in the former case, the ECAL cluster associated with the track is simply the one reconstructed through the hybrid or multi-5 $\times$ algorithm that led to the seed, while in the latter case the association is made with the electron Particle Flow cluster (PF). In the PF clustering algorithm, for each GSF track several

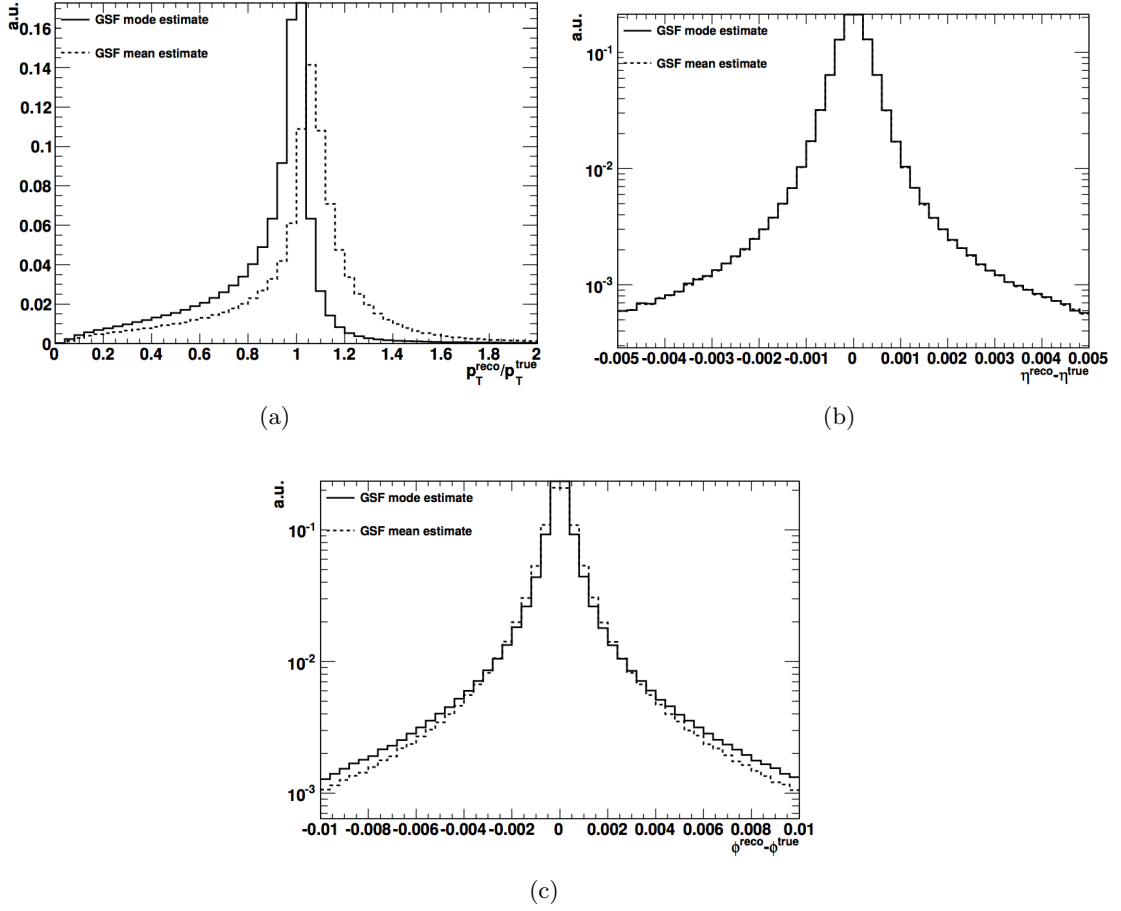


Figure 3.4: The electron track momentum parameters residual distributions for the weighted mean (dashed line) and the most probable value (solid line) estimates at the innermost track position and for electrons from $Z \rightarrow ee$ decays: in (a) the transverse momentum magnitude while in (b) and (c) the momentum η and ϕ directions [63].

PF clusters, corresponding to the electron at the ECAL surface and the bremsstrahlung photons emitted along its trajectory, are grouped together. The PF cluster corresponding to the electron at the ECAL surface is the one matched the track at the exit of the tracker.

The cluster-track matching, just like the seeding selection, is designed to preserve the highest efficiency and reduce the misidentification probability: for these reasons it is not very restrictive along the direction of the track curvature affected by bremsstrahlung. The overall efficiency for electron from Z decay is approximately 93% [60].

Bremsstrahlung estimate

The fraction of energy lost through bremsstrahlung is estimated starting from the knowledge of the point of closest approach to the beam spot, p_{in} and the momentum extrapolated to the surface of the ECAL from the track in the outermost layer of the

tracker, p_{out} . From these quantities the bremsstrahlung fraction is defined to be [57]

$$f_{brem} = \frac{p_{in} - p_{out}}{p_{in}} \quad (3.1)$$

The quantity $p_{in} - p_{out}$ is the measure of the integral amount of bremsstrahlung and the mean fraction of energy radiated along the complete trajectory is roughly proportional to the integral amount of material traversed, which varies strongly in the η direction. The variable f_{brem} is used to estimate the electron momentum entering the identification procedure, since it is linearly dependent on the radiated energy fraction. The bremsstrahlung fraction has been studied with $Z \rightarrow e^+e^-$ data and simulated events: fig. 3.5 shows that f_{brem} should be a positively defined quantity, but when a bremsstrahlung photon is emitted prior to the first three hits in the tracker or when the amount of energy radiated is very low, p_{in} is underestimate and p_{out} can be measured to be greater than p_{in} leading to a negative value for f_{brem} . In addition, regarding the dependence of f_{brem} from η , it can be seen that in the central barrel region the bremsstrahlung fraction peaks at low values since the amount of material in front of the ECAL is modest. A different behaviour is observed in the outer regions, where the material budget is higher and leads to a consistent population of electrons emitting a high fraction of their initial energy via bremsstrahlung. Another important consequence of the complex tracker material layout is the observed disagreement between data and simulation in the endcaps regions: this is imputed to an imperfect modelling of the material in simulation. The bremsstrahlung fractions turns out to be a perfect tool for study the intervening material [60]. The difference between data and simulation are relevant for updating the simulated geometry of the CMS subdetectors, through specific corrections to be applied to the electron momentum scale, resolution, identification and reconstruction efficiency.

Together with the bremsstrahlung fraction f_{brem} , another useful variable, able to assess the amount of bremsstrahlung emitted by the electrons, can be defined. The R_9 variable is defined as the ratio of the energy reconstructed in the 3×3 crystal matrix centered on the crystal with most energy and the SC energy. The category of electrons with a low level of bremsstrahlung is defined by $R_9 \geq 0.94$ and the one with a high level of bremsstrahlung by $R_9 < 0.94$.

3.1.3 Electron classification

The electron momentum is estimated using a combination of tracker and calorimeter measurements and it is particularly sensitive to the pattern of the emitted bremsstrahlung photons as well as the energy lost in the tracker material before the ECAL. The mean energy loss increases with η , as expected, that is together with the increasing amount of the tracker material: on average, the energy lost for 10 GeV electrons impinging on the edge of the barrel amounts to 7% of their initial energy [57]. This phenomenon is due to e^+e^- pairs from bremsstrahlung photon conversion so soft to be trapped in the magnetic field, losing most of their energy before reaching the ECAL or migrating away from the main supercluster: such kinds of effects are particularly expected for electrons that produce electromagnetic showers in the first stages of the tracking material.

In general, effects induced on electron energy measurements by bremsstrahlung emission, partial supercluster containment and energy lost are expected to highly

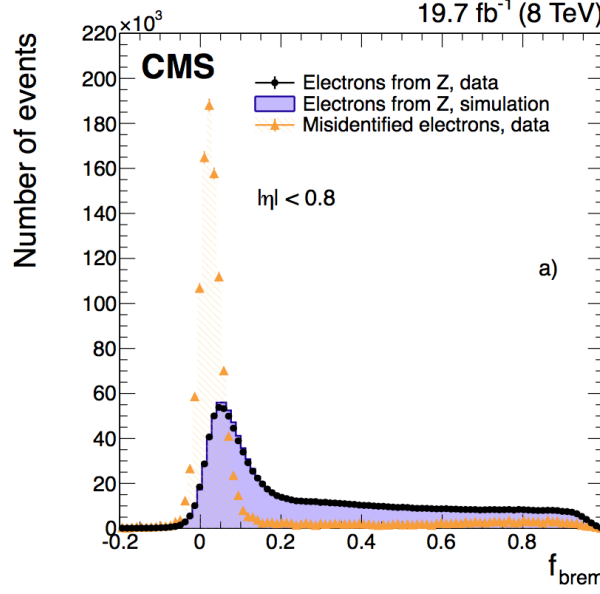


Figure 3.5: The distribution of f_{brem} for electrons from $Z \rightarrow e^+e^-$ data (dots) and simulated (solid histograms) events, and misidentified electrons from background-enriched events in data (triangles) in the central barrel region [60].

fluctuate event-by-event:

- The ϕ -spread of the energy deposit in the ECAL depends on the pattern of the bremsstrahlung emission along the electron trajectory as well as on the electron p_T . The vast majority of bremsstrahlung photons have low energy and populate a continuous band of crystals in ϕ -direction. The fluctuations in the ϕ -spread of these energy deposits can affect the energy resolution.
- Only a small number of hard photons is emitted but, if the emitter electron has low p_T the photon can deposit its energy far away from the electron forming a separate cluster. This results in additional pattern dependent fluctuation of the supercluster energy containment.
- Photons emitted in the innermost silicon layers or in η regions with a large amount of tracker material are more likely to convert before reaching the ECAL than photons emitted late (or in different η regions). This introduces additional energy measurement dependence on the bremsstrahlung emission pattern.

Since the effects on energy measurements are so variegated, various observables sensitive to the amount of bremsstrahlung radiation emitted along the trajectory and to the pattern of photons emission and conversion are used to classify electrons. These are:

- f_{brem} , the measured bremsstrahlung fraction in the tracker, introduced in eq. 3.1.
- $f_{brem}^{ECAL} = [E_{SC}^{PF} - E_{ele}^{SC}]/E_{SC}^{PF}$ where E_{SC}^{PF} and E_{ele}^{SC} are the SC energy and the electron cluster energy (measured with the PF algorithm).

- ϕ -match between reconstructed track and the SC, a quantity sensible to the bremsstrahlung collection pattern.
- The match between the SC energy and momentum measured at the tracker origin, sensible to the energy lost in the tracker
- The pattern of clusters in the superclusters.

Electrons are finally separated in different classes [57,60]:

- **Golden** electrons: are those with little bremsstrahlung, with a reconstructed track strong matching the supercluster and a well behaved supercluster pattern. They provide the most accurate estimation of the momentum and are defined by:
 - A supercluster formed by a single cluster (without observed subclusters).
 - A measured bremsstrahlung fraction $f_{brem} < 0.2$.
 - A ϕ -matching between the SC position and the track extrapolation from last point within ± 0.15 rad.
 - A value of $E_{SC}/p_{in} > 0.9$.
- **Big-Brem** electrons: this class contains electrons with a good matching between E_{SC} and p_{in} , a well behaved supercluster pattern and no evidence of energy loss from secondary photon conversion despite a very large bremsstrahlung fraction. In addition, electrons with a large amount of bremsstrahlung radiated in a single step, either very early or very late when crossing the tracker along the particle trajectory can fall into this category defined as:
 - A supercluster formed by a single "seed" cluster.
 - A measured bremsstrahlung fraction above 0.5.
 - An E_{SC}/p_{in} value between 0.9-1.1.

A complementary set of electrons, still with a good energy-momentum matching but which fails some of the previous classes criteria are defined as:

- **Narrow** electrons: are those with a significantly large bremsstrahlung fraction but not has high as for the Big-Brem, a well behaved supercluster and a not optimal track-supercluster geometrical matching. They are define as:
 - A supercluster formed by a single "seed" cluster.
 - An E_{SC}/p_{in} value between 0.9-1.1.
 - A measured bremsstrahlung fraction and/or a ϕ matching outside tha range of the previous classes.
- **Showering** electrons: are those which failed to enter all the previous classes. Several cases are included, from the electron supercluster patterns involving one or several identified bremsstrahlung sub-cluster(s) to cases where a bad energy-momentum matching is observed (usually in cases of secondary conversion of some early high energy radiated bremsstrahlung photons).

- **Crack** electrons: are those impacting in the immediate vicinity (less than about half the width in η of a crystal) of ECAL inter-module borders as well as those entering between the ECAL barrel and endcaps or at the high $|\eta|$ edge of the endcaps.

Most of the tail in the reconstructed energy spectrum of electrons between 5 and 100 GeV is localized in the Showering class electron while Big-Brem and Narrow cases appear rather well measured: this is an indication of the solidity of the energy clustering algorithm [57]. Events from Z-bosons decays give the integrated proportion of electrons in the different classes both for experimental and simulated data: 57.4% and 56.8% for Showering, 25.5% and 26.3% for Golden, 8.4% and 8.0% for Big-Brem [60].

3.2 Energy calibrations and corrections

The electron momentum can be measured in CMS by two detectors, the tracker and the electromagnetic calorimeter, with different resolution. At energies < 15 GeV, or for electrons near gaps in detectors, the track momentum is expected to be more precise than the ECAL super-cluster energy. In this energy region the track momentum is estimated from a linear combination of the tracker and super-cluster measurements, improving significantly the resolution for all electrons in the barrel up to energies of about 35 GeV. Conversely, for energies larger than 15 GeV, the resolution on the electron momentum measurement is greater in the ECAL and dominates the tracker resolution near 35 GeV, as shown in fig. 3.6. Being the ECAL momentum measurements resolution

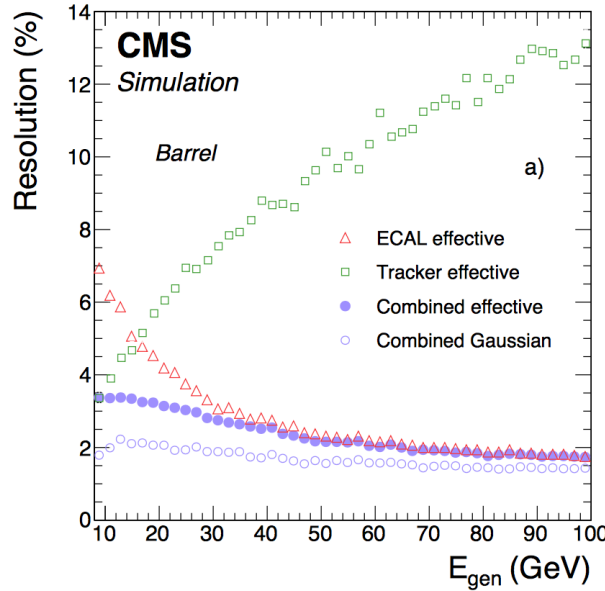


Figure 3.6: Effective resolution in electron momentum after combining E_{SC} and p estimates (solid circles), compared to those using the corrected SC energy (triangles). The ECAL measurements resolution is greater than the tracker ones for energies > 15 GeV and dominates for energies > 35 GeV [60].

important in a considerably high momentum range, the calorimeter should be carefully calibrated correcting the energy measurements.

The ECAL detector has been efficiently operating since installation with a good percentage of responding channels and a low electronic noise, corresponding to a mean energy deposit of 40 MeV – 50 MeV per channel both in the barrel and endcaps. Few channels are currently dead, however these are suppressed in the trigger and offline reconstruction.

The operating temperature of the ECAL is maintained by a dedicated cooling system, since the temperature dependence of the crystal light yield and of the APD gain demand a precise temperature stabilization in order to limit its contribution to the constant term of the energy resolution to be less than 0.2%.

The APD working point is another parameter carefully tuned in order to provide a good signal-to-noise ratio with an acceptable sensitivity to the bias voltage. The contribution of the gain variation to the constant term in the energy calibration is required to be less than 0.2%.

The ECAL response varies also under other effects, such as the formation of colour centers that reduce the transparency (and consequently the low light yield) of the lead tungstate. This "yellowing" process is a direct consequence of the high luminosity energy irradiation of the PbWO₄ crystals, it is semi-permanent in the sense that the transparency naturally partially recovers through spontaneous annealing, but must be corrected. To do so, a monitoring system, based on the injection of laser light at 440 nm in the barrel (close to the emission peak of scintillation light by the lead tungstate) is used to track and correct for response changes during LHC operations as well as to give informations on the system stability, together with ancillary lasers and LED systems in the endcaps [64]. Using quasi-online processing of the monitoring data, single-channel response corrections are available in less than two days for prompt reconstruction of the CMS data. The relative response to laser light injected in the ECAL crystals and measured by the ECAL laser monitoring system for the 2011, 2012, 2015 and 2016 data taking periods is shown in fig. 3.7

As previously seen (cf. 3.1.1), clustering algorithms are mandatory in order to collect the energy released by particles impinging the ECAL detector. After clustering together all the energy deposits in the ECAL and forming the superclusters the energy in the supercluster is evaluated using the following general expression [66]:

$$E_{e,\gamma} = F_{e,\gamma} \cdot [G \cdot \sum_i S_i(t) \cdot C_i \cdot A_i + E_{ES}] \quad (3.2)$$

In eq. 3.2 the sum is intended over the crystals i belonging to the relative supercluster. The energy deposited in each crystal is given by the *pulse amplitude* (A_i) in ADC counts multiplied by the ADC-to-GeV *conversion factors* (G) measured separately in the barrel and endcaps plus the *preshower energy* (E_{ES}), which is computed and calibrated separately. The additional terms are all corrections: C_i is the *intercalibration coefficients* between different channels, $S_i(t)$ a correction term due to *radiation-induced channel response changes* as a function of the time and finally the $F_{e,\gamma}$ term represents the *energy correction* applied to the superclusters to take into account the η and ϕ dependent geometry and material effects as well as the fact that electrons and photons showers behave slightly differently.

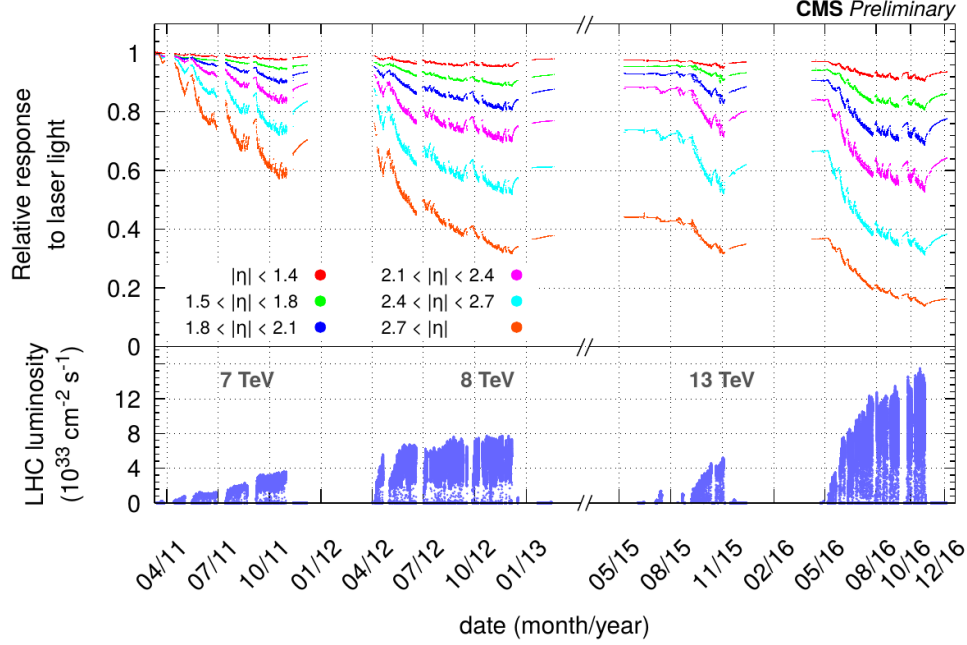


Figure 3.7: Relative response to laser light (440 nm in 2011 and 447 nm from 2012 onwards) injected in the ECAL crystals, measured by the ECAL laser monitoring system, averaged over all crystals in bins of pseudorapidity, for the 2011, 2012, 2015 and 2016 data taking periods, with magnetic field at 3.8 T. The recovery of the crystal response during the Long-Shutdown-1 period is visible. The bottom plot shows the instantaneous LHC luminosity delivered during this time period [65].

Since the ECAL calibration and performance are studied via estimation of the electrons and photons supercluster energy, the electron identification and classification, done matching the supercluster and tracker position informations previously seen, is an important step before the energy calibration. Once obtained "good" electrons, the complete set of corrections of eq. 3.2 can be investigated.

3.2.1 Correction for changes in response

Corrections to response changes in the ECAL are calculated using the light monitoring system (LM): the laser light is injected both in the front and rear area of every crystals via optical fibers, but since the spectral composition and the path for the collection for laser light at the photodetectors are different from those of scintillation light, a conversion factor is required in order to relate changes in the ECAL to responses to the laser light as can be seen in eq. 3.3 [46]:

$$\frac{S(t)}{S_0} = \left(\frac{R(t)}{R_0} \right)^\alpha \quad (3.3)$$

where $S(t)$ is the channel response to scintillation light at a certain time, S_0 the initial response, $R(t)$ and R_0 the corresponding quantities as before but for the laser light, and finally α an exponent independent on the loss for small transparency losses. The

value of α is determined from a beam test on a limited amount of crystals, showing a spread of about 10% (deriving from non-uniformity between different crystals): this spread limits the precision of the response correction for a single channel to 0.3% and 0.5% respectively in the barrel and endcaps.

Due to the presence of fluctuations, validation and tuning of the α value, using collision data, is mandatory. To do so, the energy of electrons from W-bosons decays along with the reconstructed mass from $\eta \rightarrow \gamma\gamma$ decays are used. The energy resolution is measured with $Z \rightarrow e^+e^-$ events. The η mesons are selected online by a dedicated calibration trigger then a fit is carried on the invariant mass distribution in the mass range of the meson. A large number of measurements are possible for each LHC fill session due to the high rate of the produced η events: this permits short term changes in the ECAL response to be verified before prompt data reconstruction takes place, providing a fast feedback to validate the LM corrections. While η events are used for quasi-online validation, isolated electrons from W-bosons decays are used to validate response corrections over periods of days to weeks. A reference E/p distribution is obtained from the energy measured in ECAL and the momentum measured in the tracker for the entire data set after applying LM corrections: this reference distribution is then scaled to fit E/p distributions obtained by dividing the same data in groups of 12 000 consecutive events. The scale factors provide a measure of the relative response as a function of time. This method does not depend on the absolute calibration of both the calorimeter and the tracker. Finally the validation of the response corrections is also carried out monitoring the ECAL energy resolution using events with Z-bosons decaying into pairs of electrons: the data are first selected and then fitted with a Crystal Ball (CB) parametrization building a new control variable, defined as σ_{CB}/m_Z . The optimal α values are validated minimizing the σ_{CB}/m_Z during the operation time of the CMS experiment.

Although excellent energy response and resolution stability can be achieved with these methods, residual instabilities in the barrel and endcaps remain, contributing to the constant term of the energy resolution.

3.2.2 Single channel intercalibration

The intercalibrations of the ECAL crystals are relative calibrations between channels. A number of methods are used, determining for both the barrel and endcaps different intercalibration constants, which are usually normalised by their average value, in order to obtain a set of numbers with a mean value of unity. These numbers are then combined to provide a weighted mean intercalibration constant for each channel (C_i).

The entire intercalibration process follows two main steps: first, an initial set of calibrations also known as *pre-calibrations*, are obtained from different sources like laboratory measurements, beam tests and exposure to cosmic rays before the installation of the ECAL calorimeter in the CMS structure and on a limited number of crystals both in barrel and endcaps (although all barrel channels were calibrated with cosmic-ray muons). The intercalibration constants for each method are then cross-checked for consistency and validation and then combined to provide a weighted intercalibration for each crystal: the precision of the intercalibration after the complete installation of the calorimeter into the CMS structure oscillates between 0.5% and 5%. The pre-calibration is then followed by an *intercalibration refinement* carried out via collision data following

different methods:

1. The ϕ -simmetry method is based on the expectation that the distribution of energy deposits for a large number of collected minimum bias events is the same in all the crystals in a particular η ring (with the same pseudorapidity).
2. In the π^0 and η method the invariant masses of the photon pairs produced in the decays of these mesons are used to intercalibrate the channel response.
3. The W- and Z-boson method exploits the isolated electrons from these bosons decays to compare the energy measured in the ECAL to the track momentum measured in the silicon tracker.

The first data-driven method is intended to calibrate channels at the same pseudorapidity while isolated electrons/photons from the second and third are used to extend the intercalibration to various η rings. Since these intercalibration refinements exploit experimental data, the precision of each method is studied with the aid of MC simulations while the validation comes from both the pre-calibrations and the channel-by-channel comparison of the constants derived in each method.

The intercalibration in ϕ is computed from the ratio of the total energy deposited in a crystal to the mean of the transverse energy deposited in all the crystals in the same η ring. Since this method exploits an expected simmetry of minimum bias events, special triggers with low thresholds are used to select the events. The data analysis is then performed for energy deposits with transverse energy between an upper and lower bond, applied in order to remove noise contribution and to minimize the fluctuations induced by rare fluctuations of high energy deposits in some crystals. Corrections are then applied to compensate for known azimuthal inhomogeneities of the CMS detector (intermodule gaps and tracker support system in front of ECAL). The ϕ method is used to monitor the stability of the intercalibration constants and to improve the results from the other intercalibrations methods.

The diphoton decays of the π^0 and η mesons are exploited to quasi-online intercalibrate the channels. In order to select the largest possible statistic from the copious production of such mesons while having a minimal impact on the full CMS readout bandwidth and storage space, a special data stream is used: candidate diphoton decays are directly selected online from all the events passing the single- e/γ and single jets L1 triggers but only limited data in the vicinity of photon candidates are kept. Since this method is based on the estimation of the peak of the $\gamma\gamma$ invariant mass distribution, the energy of the photon is computed summing the energy deposits in a 3×3 matrix crystal around the seed. The single-crystal energy deposits are small so corrections are applied to account for the noise effects. The transverse energy of both η and π^0 is required to be above a certain fixed threshold and, to suppress photons converted in the material in front of the ECAL, a special transverse shower shape, corresponding to an electromagnetic shower produced by an isolate photon, is required. Once the events are selected, an iterative procedure is followed: for each crystal in which a decay photon from the mesons candidate impinging the resulting invariant mass is fitted with a Gaussian function for the signal. The intercalibration constants are then updated iteratively to correct the fitted mass value in each channel. The quality of the calibration depends on the amount of selected candidates per crystal and on the signal-to-noise

ratio: the results from each meson resonance are then combined to form an average weighted by the quality.

The intercalibrations in ϕ and η directions are computed from the E/p ratio of the energy of isolated electrons, selected from W- and Z-bosons decays, measured in the ECAL to the momentum measured in the tracker. The isolated electrons data are exploited both to improve the intercalibration in ϕ and η directions:

- The intercalibration in ϕ is calculated with an iterative procedure that derives constants for all the channels. The constants are normalized to a mean value equal to unity in each ϕ ring and corrections are applied to take into account the effects of the supermodule boundaries and the variation of the tracker momentum response (variations in the amount of tracker material in ϕ direction affect the electron measurement distribution due to the bremsstrahlung losses). The invariant mass is then reconstructed in 360-bins in the ϕ direction: the square root of the invariant mass peak position in each bin is proportional to the local momentum scale for the corresponding region of the detector.
- The intercalibration along η is calculated again exploiting the E/p ratio from W- and Z-bosons decays. A reference E/p distribution is computed from MC simulation and then scaled to fit the E/p distribution from data in the same ϕ ring. Since the material budget varies with η , various reference MC simulations are carried out on different pseudorapidity regions. For each ring a specific calibration of the local momentum scale for electrons is derived from $Z \rightarrow e^+e^-$ events. Corrections to the supercluster energy are applied as well. The E/p scale factors provide a measure of the relative response to electrons along η : while in MC simulation they are consistent with unity, a residual η dependence on data indicate the need for further tuning of the energy corrections in data [66].

In order to further improve the single channel intercalibration, the previous three methods results are combined together via a weighted average on the respective precision of each constant value in fixed ϕ rings. The combined intercalibration precision turns out to be 0.4% for central barrel crystals, 0.7 – 0.8% for the rest of the barrel and 1.5 – 2% in the endcaps. A variation on the precision is observed when pseudorapidity changes: this is mainly a consequence both of the size of the used data sample and of an imperfect description of the tracker material. Fig. 3.8 shows the estimated precision of intercalibration constants in barrel and endcap as a function of pseudorapidity using the intercalibrations methods previously listed. The data set was collected in 2011, providing a precision in the intercalibration constants of 0.5 – 1% each month in the EB and of 2 – 4% each 2-3 month in the EB.

To illustrate the importance of the corrections to the bare supercluster energy so far introduced, fig. 3.9 shows the dielectron invariant mass distribution in different scenarios: without any corrections (blue), with the intercalibration correction applied (red) and finally with both intercalibration and light monitoring system calibrations applied (black). In all cases the energy correction term $F_{e,\gamma}$ (cf. 3.2.4) is included in energy computation. The corrections effects are dramatic: the resolution of the Z-boson peak is greatly enhanced both in the ECAL barrel and endcaps while in the endcaps corrections provide a remarkable shift of the peak mean position to be consistent with the nominal value of the Z-boson invariant mass $m_Z = 91.1876(21)$ GeV [6].

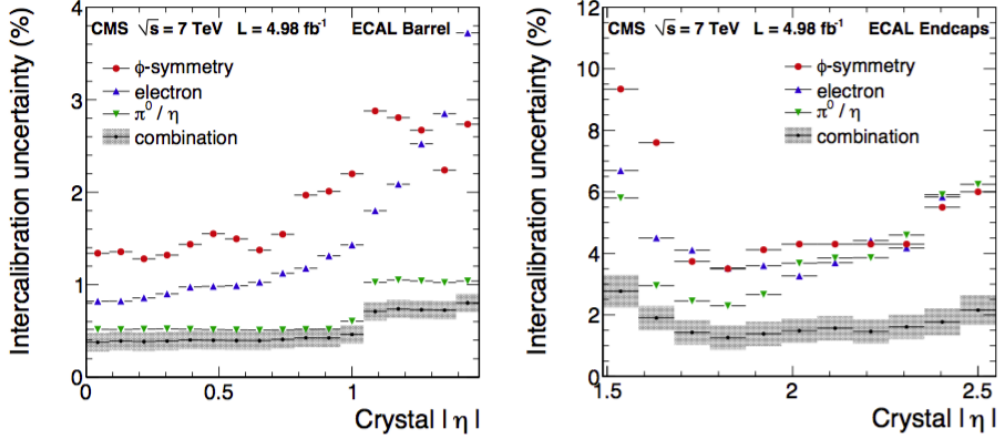


Figure 3.8: The intercalibration precision obtained using ϕ -symmetry, the E/p ratio with electrons and the π^0/η decays. The resultant combined precision is shown. The pattern along η is related to the distribution of the tracker material [66].

3.2.3 Calibration of the preshower

The precision required for the preshower energy calibration is intimately connected to the requested energy resolution in the ECAL endcaps, since a certain fraction of energy of the impinging endcaps particles is deposited in the PS subdetector. Approximately 6 – 8% of the shower energy is deposited in the preshower.

Two kinds of calibrations are performed on the preshower subdetector: prior to installation, all the sensors were calibrated with cosmic rays while, in situ they are calibrated using charged pions and muons with a momentum as large as to be minimum ionizing particles (MIPs) and a good signal to noise ratio. Since the preshower is a relatively thin detector the pulse height distribution in each channel is fitted with a Landau distribution convolved with a Gaussian function: the fitted peak is taken as the calibration.

Finally, preshower clusters are identified from the position of crystal clusters in the endcaps and the energy is evaluated according to eq. 3.4:

$$E_{ES} = G_{ES} \cdot (E_{ES}^{clust_1} + \alpha_{ES} \cdot E_{ES}^{clust_2}) \quad (3.4)$$

where $E_{ES}^{clust_i}$ are the energy deposits in the two layers of the preshower detector, expressed in ADC counts, G_{ES} the ADC-to-GeV conversion factor and α_{ES} the relative weight of the second preshower plane with respect to the first. Beam test results and W-bosons decays are used to estimate these two parameters.

3.2.4 Energy corrections

As previously seen, superclusters are used to reconstruct the energies of both photons and electrons impinging on the ECAL detector, summing together all the energy deposits spread in the ϕ direction due to interactions with both the tracker material and the solenoid magnetic field. Once grouped, a correction function $F_{e,\gamma}$ derived from MC

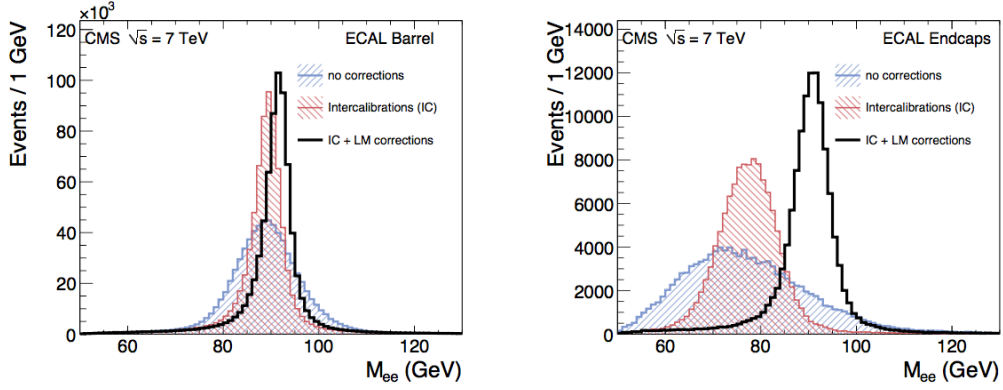


Figure 3.9: Reconstructed invariant mass distribution from $Z \rightarrow e^+e^-$ decay changes after different calibration: no corrections (blue), intercalibration correction (red) and intercalibration plus LM corrections applied (black) [66].

simulations should be applied in order to account for energy containment effects, such as the shower containment in the calorimeter and the energy containment in the supercluster for particles that showers in the material before entering the ECAL. The energy correction is tuned separately for electrons and photons, since they interact differently with matter.

The corrections for photons have been optimized using a multivariate regression technique based on a Boosted Decision Tree (BDT) implementation. The regression is trained on prompt photons from $\gamma + \text{jets}$ MC samples using the ratio of generator level photon energy to the supercluster energy, including as well the preshower energy for the endcaps. The input variables are:

- The number of primary vertices to correct for the dependence of the shower energy on spurious energy deposits due to pile-up events.
- The η and ϕ coordinates of the supercluster (global variables).
- A collection of shower shape variables to provide informations on the likelihood and locations of a photon conversion and the degree of showering in the material.
- A set of local cluster coordinates to measure the distance of the clusters from the ECAL boundaries, providing informations on the amount of energy which is likely to be lost in crystal, module gaps and cracks, and driving the level of local containment corrections predicted by the regression.

While the global and local containment corrections address different effects, these corrections are allowed to be correlated in the regression to account for the fact that a photon converted before reaching the ECAL is not incident on a single point on the calorimeter face and is therefore relatively less affected by local containment.

The primary validation tool for the regression is to compare data and MC simulation performance for electrons in Z- and W-bosons decays: a BDT with identical training settings and input variables as previously described is trained on a MC sample of electrons from Z-bosons decays. The relative corrections differs from the ones used

for the electron reconstruction in CMS (where tracker and ECAL informations are merged together) however they enabled a direct comparison of the ECAL calibration and resolution in data and MC simulation.

To illustrate the impact of the energy correction on the supercluster energy, fig. 3.10 shows the dielectron invariant mass from $Z \rightarrow e^+e^-$ reconstructed both in barrel and endcaps using a fixed-matrix 5×5 clustering algorithm (blue) with respect to the supercluster algorithm that collect most of the radiated energy (red) and the applying the energy corrections (black). For the endcaps is also shown the reconstructed supercluster energy including the preshower contributes (orange). The improvements in the Z-boson mass peak position and resolution are clearly demonstrated.

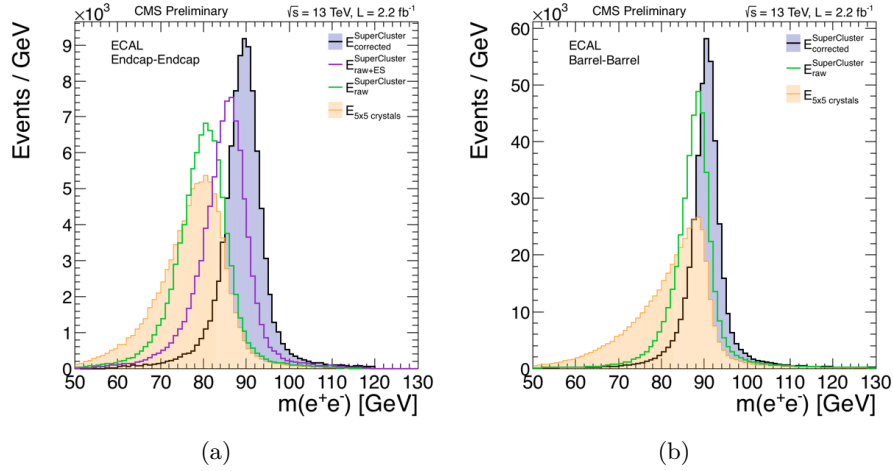


Figure 3.10: Reconstructed invariant mass from $Z \rightarrow e^+e^-$ decays applying a fixed-matrix clustering (orange), a supercluster clustering (green, purple) and the relative energy corrections (blue). The distributions are shown for the barrel (fig. 3.10(a)) and for the endcaps (fig. 3.10(b)) regions [67].

3.2.5 Absolute energy calibration

In ECAL the absolute energy calibration (G) is computed in different reference regions of the calorimeter, where the effects of upstream material and uncertainties in the energy corrections are minimal. In the barrel the calibration region is defined as the central 150 crystals in the first module of each supermodule, requiring a minimum distance of 5 crystal from the border of each module. This region presents various advantages: the material budget in front of it is small, the geometry of these crystals is very similar and the energy leakage between module or supermodules due to gaps is limited. In the endcaps, the reference region is defined as the central area of each endcap ($1.7 < |\eta| < 2.1$). Once defined the reference regions, the absolute energy calibration is worked out: in the MC simulation it is computed using 50 GeV unconverted photons (that is to say photons that have not interacted with the tracker material via pair production before entering the calorimeter) and it is defined such that the energy reconstructed in a 5×5 crystal matrix in the ECAL is equal to the true energy of the

photon in the reference region. Experimental data from the decays of Z-bosons into two electrons are then used, first to set the overall energy scale in the barrel and endcaps in data relative to MC simulation, then to validate the energy correction function F_e (cf. 3.2.4) for electrons using the Z-boson mass constrain. Finally, decays of Z-bosons into two muons where one muons radiate a photon ($Z \rightarrow \mu^+ \mu^-$) are used to cross-check the energy calibration of photons.

4 | Fine tuning of calibration

In this chapter the fine-tuning corrections to the electrons energy scale and resolution are introduced. The main analysis, concerning the study of the residual non-linearity in the electron energy scale, is presented. An electron energy correction function is evaluated and applied to simulated Higgs boson four-lepton decay. The systematic shift, due to the electron energy scale correction, is evaluated together with its statistical uncertainty.

4.1 Residual data/MC discrepancies

In the $H \rightarrow ZZ^* \rightarrow 4l$ Higgs boson decay channel the ZZ system, composed by pairs of same-flavour and opposite-sign isolated electrons, is reconstructed. The distributions of the four-lepton reconstructed mass in the full mass range for the sum of the $4e$, 4μ and $2e2\mu$ channels are shown in fig. 4.1(a) for data collected at $\sqrt{s} = 7 \text{ TeV} - 8 \text{ TeV}$ and in fig. 4.1(b) for data collected at $\sqrt{s} = 13 \text{ TeV}$.

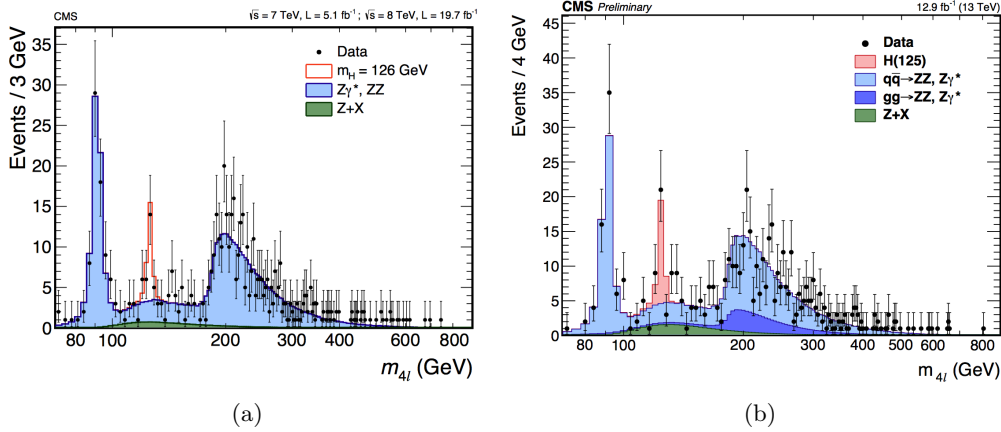


Figure 4.1: Distribution of the four-lepton reconstructed invariant mass m_{4l} in the full mass range for Run 1 [68] (fig. 4.1(a)) and Run 2 [69] (fig. 4.1(b)) data. Points with error bars represent the data and stacked histograms represent expected distributions. The 125 GeV Higgs boson signal and the ZZ backgrounds are normalized to the SM expectation, the $Z + X$ background to the estimation from data.

Despite providing a clear signal over a small background, this channel presents a

very low branching ratio (cf. sec. 1.3.2): in order to provide the best possible Higgs boson mass measurement in this channel, it is important to maintain a very high lepton selection efficiency in a wide range of momenta. Regarding the signal sensitivity, it turns out to depend on:

- The 4l invariant mass resolution, since the signal is represented by a narrow resonance on top a smooth background.
- The lepton momentum scale, since to obtain a precise mass measurement of the resonance decay, it is mandatory to calibrate the individual lepton scale such that the systematic uncertainty in the measured value of m_H is smaller than the statistical uncertainty.

As previously seen in sec. 3.2, several procedures are used to calibrate the energy response of individual crystals and to set an absolute energy calibration of the calorimeter. The energy of the supercluster is corrected for the imperfect containment of the clustering algorithm for the electron energy not deposited into ECAL, for energy leakage arising from showers near gaps between crystals or between ECAL modules and finally for the additional energy coming from pileup interactions. The supercluster energy corrections $F_{e\gamma}$, described in sec. 3.2.4, are derived from simulations based on the best knowledge of the detector conditions, encoded in the ECAL calibrations and tracker alignment. Small discrepancies between data and simulation samples remain and affect the electron momentum scale and resolution. For this reason a fine-tuning of the ECAL calibration and resolution is needed. Events in data are used to account for these discrepancies as well to correct biases. Since the energy in the individual crystals is already calibrated and the simulation of electron/photon showering in ECAL (including the measured uncertainties on the inter-calibration between crystals) is rather precise, the expected remnant calibration corrections are quite small. Indeed, large fluctuations of the variables used for the fine-tuning procedure would be an indication of a bad calibration procedure.

There are many sources of discrepancy between the energy estimated in data and in simulation, here listed in order of decreasing severity [60]:

1. An imperfect description of the tracker material in simulation can affect differently each category of electrons (cf. sec. 3.1.3).
2. The evolution of the crystals transparency and of the noise during the entire data taking time can lead to additional difference between data and MC. This effect can be mitigated by using specific run-dependent simulations.
3. The underestimation of the uncertainties correlated to the calibration constants of every crystals.
4. A residual difference in the ECAL geometry relative to the nominal one used in the simulation and can cause the SC energy corrections derived via simulated events to be inappropriate for data.

The fine-tuning corrections account both for the energy scale and for the energy resolution in data and MC simulation.

4.2 Fine-tuning of the energy scale with $Z \rightarrow e^+e^-$ events

The di-electron invariant mass in Z decays events is extracted from two quantities: the supercluster energy in ECAL, reconstructed with all the energy corrections derived from MC simulation, and the opening angle measured from tracks at the event vertex. Both energy scale and resolution are extracted from the invariant mass distribution of electron pairs with a transverse energy for each electron larger than 25 GeV in a mass range 60 GeV-120 GeV. The invariant mass distribution is then fitted with a Breit-Wigner distribution convoluted with a Crystal Ball function.

$$CB(x-\Delta m) = \begin{cases} e^{-\frac{1}{2}\left(\frac{x-\Delta m}{\sigma_{CB}}\right)^2}, & \text{if } \frac{x-\Delta m}{\sigma_{CB}} > \alpha_{CB} \\ \left(\frac{\gamma}{\alpha_{CB}}\right)^\gamma \cdot e^{-\frac{\alpha_{CB}^2}{2}} \cdot \left(\frac{\gamma}{\alpha_{CB}} - \alpha_{CB} - \frac{x-\Delta m}{\sigma_{CB}}\right)^{-\gamma}, & \text{if } \frac{x-\Delta m}{\sigma_{CB}} < \alpha_{CB} \end{cases} \quad (4.1)$$

As can be seen in eq. 4.1, among all the fitting parameters two of them estimate the displacement of the distribution peak with respect to the true Z mass (Δm) and the width of the Gaussian component of the Crystal Ball function (σ_{CB}), a measure of the energy resolution. Two other parameters, α_{CB} and γ , account for the tails of the Crystal Ball, whose shape is determined by showering electrons energy, which is not fully recovered by the clustering algorithms. A fit to the invariant mass distribution is then performed, constraining the shape of the distribution tails from MC simulation while the nominal mass and width of the Breit-Wigner distribution, relative to the Z -boson, are set to the nominal values $m_Z = 91.188$ GeV and $\Gamma_Z = 2.495$ GeV.

The fit is performed both in data and simulation: the ADC-to-GeV conversion factor G for data is adjusted such that the fitted $Z \rightarrow e^+e^-$ peak agrees with that of the MC simulation, separately for the barrel and endcaps. The systematic uncertainty associated to the absolute energy calibration is estimated to be 0.4% in the EB and 0.8% in the EE. The uncertainty is dominated by the accuracy associated to the energy correction function (F_e) [66]. The size of this uncertainty is determined deriving the energy scale from di-electron invariant mass distributions, reconstructed using the raw supercluster energy both in data and in MC events, using Monte Carlo samples generated with the tracker material budget altered within its uncertainty [70, 71]. The observed variations in the results are then taken as the systematic error. The dependence of G also on a number of additional effects, such as its stability on changes in the event selection or in the functional form used to describe the ECAL response, has been studied and it was found to be 0.2%.

4.2.1 Validation of the energy calibration and corrections

The quality of the energy corrections in data is assessed studying the variation of $E_{ECAL}/p_{tracker}$ with isolated electrons from W - and Z -bosons decays together with the fluctuation of the mass resolution in $Z \rightarrow e^+e^-$ decays, as a function of several variables that could impact on the energy reconstruction. It follows that the effects of pile-up on the shower energy can be mitigated, while a case of imperfect corrections has been identified studying E/p as a function of the impact point of the electron on the crystal. This shows that the corrections based on the MC simulation cannot fully compensate for the energy leakage in the inter-crystal gaps, yielding a residual response variation of

1% between showers impinging the center or the boundary of a crystal. These effects are estimated to contribute to the current energy resolution of the calorimeter with an RMS of about 0.3-0.5%.

4.3 Fine-tuning of the simulated resolution

This correction accounts for effects mainly related to the material in front of the calorimeter, using the smearing method. Three explanations have been suggested for the need of an additional smearing of the energy estimate in simulated events to achieve complete agreement with data [66]:

1. The amount of tracker material has been verified with different methods but the energy resolution is better for electrons that do not interact significantly with the tracker material (the ones with $R_9 > 0.94$). These observations suggest that the deteriorating effect of the material on the reconstruction algorithms should be further investigated.
2. An imperfect knowledge of the α parameter of eq. 3.3 contributes to the degradation of the resolution in data. This can be mitigated by a first-order optimization of the α parameter as shown in sec. 3.2.1.
3. The inter-calibration constants are determined with several independent methods, which are in turn combined together, as can be seen in sec. 3.2.2. An underestimation of the inter-calibration systematic uncertainty, although unlikely, may occur since in the combination it is assumed that the different methods are completely independent. However, some experimental effects, especially the ones related to the detector geometry, may lead to common systematic uncertainties, limiting the E/p inter-calibration method statistical significance.

The smearing method allows to estimate more efficiently the effective resolution of the electromagnetic calorimeter, adding a Gaussian distributed random contribution to the energy reconstructed in simulated events. This is done in order to smear the invariant mass distribution. To do so, showers are classified in two R_9 bins in each of different barrel and endcaps pseudo-rapidity regions yielding to different shower categories. Combining different pairs of shower categories, various $Z \rightarrow e^+e^-$ invariant mass distributions are constructed and fitted both for data and simulated events. The shower energies in simulated events are then modified to accommodate for the mismatch in energy resolution: the quadratic difference between the electron resolution in data and MC is extracted and then added in quadrature as a constant Gaussian smearing to the electron MC events.

After the fine-tuning corrections, a good matching between data and simulation is obtained: the energy resolution in MC is corrected matching the one of the data while the remnant detector biases after absolute calibration are removed. fig. 4.2 shows the di-electron invariant mass for $Z \rightarrow e^+e^-$ events for data and simulation before and after the energy scale and resolution corrections.

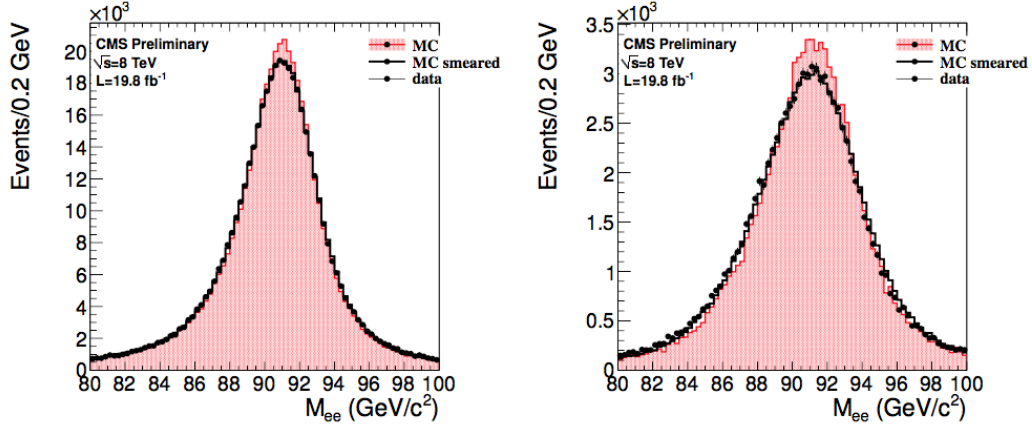


Figure 4.2: Data to simulation comparison for $Z \rightarrow ee$ events after applying the scale and resolution corrections. The red filled histograms represent the simulation before the energy smearing. The black histograms are the MC data after the smearing applied while the black dots represent data with the energy scale correction applied. On the left side electrons with $|\eta| < 1$ and $R_9 < 0.94$, on the right one electron with $|\eta| < 1$ and $R_9 < 0.94$ and the other with $1.566 < |\eta| < 2$ and $R_9 < 0.94$ [72].

4.4 Energy scale and resolution uncertainty

After these fine-tuning corrections, data to simulation irregularity may still exist mainly because the kinematic and the phase space of the calibration sample and the application sample are different. Indeed the corrections to the momentum scale and resolution, discussed above, are only obtained from correcting the SC energy in $Z \rightarrow ee$ events. This is a limitation since the check is for a range of electron energies that has only a limited overlap with the one used for the $H \rightarrow ZZ^*$ analysis, which uses electrons with p_T as low as 7 GeV [29]. As a consequence, further corrections should be applied on the electron energy scale linearity. The uncertainties related to the experimental systematic energy shifts, arising from this correction, are analyzed and described in the following.

4.4.1 Electron scale and resolution measurements as a function of category and pileup

The invariant mass is reconstructed for $Z \rightarrow ee$ events, separating the electrons in different categories in η . Additional separations for well and badly measured electrons is achieved via electron classification (cf. sec. 3.1.3) on the amount of energy radiated by Bremsstrahlung and on the reconstruction quality.

Events are then studied in low and high pileup regimes and the invariant mass distribution is fitted with a Breit-Wigner distribution with fixed parameters (to account for the Z-boson lineshape), convolved with a Crystal-ball with free parameters (to account for the detector effects).

Systematic uncertainties on electron energy scale are estimated as the maximum

deviation between data and MC of fitted invariant mass peak position in the different categories mentioned above.

Systematic uncertainties dependency on momentum scale is also studied with respect to pileup. The reconstructed $Z \rightarrow ee$ invariant mass is built for different number of vertices, in both data and MC, and fitted in the same way as above. The results are shown in fig. 4.3: the variation of the Z-boson peak with the number of vertices is not relevant. Additionally the MC simulation is well following the data.

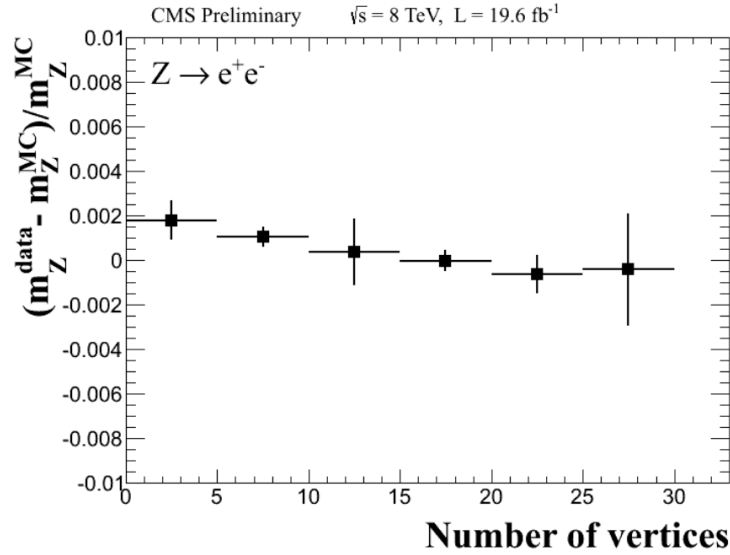


Figure 4.3: The mass peak position ratio differences between data and simulation as a function on the number of vertices [72].

4.4.2 Electron scale linearity measurement from Z , J/ψ and $Y \rightarrow ee$ decays

The studies described above were mainly checking the scale for electrons at relatively high momenta. A possible systematic shift effect on the energy scale can be due to the non-linearity in the momentum estimation differentially with respect to the Monte Carlo when propagating the electron calibration estimated at the Z-scale to the scale typical of electrons from Higgs boson decays.

This further data/MC discrepancy may again be due to the not perfect characterization of the material in front of ECAL as well as the not-perfect detector geometry description.

The electron momentum scale can be studied and validated using data from J/ψ , Y and Z decays in several bins of lepton η^1 and p_T^1 , in order to cover the full momentum range relevant for the $H \rightarrow ZZ \rightarrow 4l$ search. Fits are performed on the di-electron invariant mass distributions differentially in p_T and $|\eta|$ of the probe electron, while the tag is integrated over the full available phase space. In each fit the signal PDF is obtained from a signal template built on simulation convolved with Gaussians with floating means and standard deviations (to account for low statistics). These fits are

performed both in data and simulation. A variable sensitive to the residual energy non linearity is defined as the ratio [29, 60]:

$$R \equiv \frac{m_{data}^{peak} - m_{MC}^{peak}}{m_{PDG}} \quad (4.2)$$

where m_{data}^{peak} is the mass measured in data, m_{MC}^{peak} the mass measured in simulated events and m_{PDG} the nominal mass from the Particle Data Group [73].

The offset in the measured peak position is then assigned as a residual systematic uncertainty in the signal mass scale from $H \rightarrow ZZ \rightarrow 4l$ events. The measurement of the properties of a Higgs boson in the four-lepton final state measured at a center-of-mass energy of $\sqrt{s} = 7$ TeV and 8 TeV observed a p_T^e dependence of the momentum scale resulting in effects of 0.3% on the mass scale of the $H \rightarrow 4e$ channel [29]. This uncertainty corresponds to a systematic error, associated to the measure of the Higgs boson mass, of $\mathcal{O}(10^2 \text{ MeV})$.

4.5 Data and simulation samples

The simulated samples used in this analysis concern both the SM Higgs boson signal at generator level in the $H \rightarrow ZZ \rightarrow 4l$ channel (together with its relevant background process) and the Z-boson decays in lepton pairs. The Monte Carlo simulated samples are generated with programs based on state-of-the-art theoretical calculations.

The samples of Higgs boson signal events, produced in either gluon fusion ($gg \rightarrow H$) or vector boson fusion ($qq \rightarrow qqH$) are generated with the POWHEG generator [74] for a center of mass energy of $\sqrt{s} = 8$ TeV at next-to-leading-order (NLO) QCD accuracy. The interference between the Higgs boson signal produced by gluon fusion and the background from $\gamma\gamma \rightarrow ZZ$ are included as well, together with the theoretical uncertainty in the shape of the resonance (due to missing NLO corrections in the interference between signal and background) and the uncertainties in the electroweak corrections. The dominant backgrounds come from the SM ZZ or $Z\gamma^*$ production via $q\bar{q}$ annihilation or gluon fusion. The $ZZ/Z\gamma^*$ production via quark-antiquark(gluon-gluon) annihilation is generated at NLO with POWHEG(GG2ZZ). Other additional small background samples are generated with the MADGRAPH generator [75]. All generated samples are processed with PHYTIA [76] for jet fragmentation and showering. The simulations include pileup from overlapping pp interactions, matching the distribution of the number of interactions per LHC beam crossing observed in data.

The simulations for the $Z \rightarrow ll$ decay are obtained with the MADGRAPH generator for a center of mass energy of $\sqrt{s} = 13$ TeV, including a pileup of 25 interactions per bunch crossing and assuming a bunch spacing of 25 ns. Events are also processed through the detailed simulation of the CMS detector, based on GEANT4 [77], and are reconstructed with the same algorithms used for data.

The data samples for the Z-boson comes from proton-proton collisions at $\sqrt{s} = 13$ TeV. The samples were recorded in 2015 and collected by a *single lepton* selected at HLT: an event is selected if at least one of the lepton has a p_T larger than a certain threshold $p_T^{thr.}$. A sufficiently high $p_T^{thr.}$ is required in order to match the trigger requirements and to have a sufficiently high signal-to-noise ratio. The total statistics

analyzed correspond to an integrated luminosity approximately of 2.3 fb^{-1} of data collected by CMS in 2015 pp running.

While two data releases were obtained both for data and $Z \rightarrow \ell\ell$ simulation, the $H \rightarrow ZZ \rightarrow 4\ell$ data set is unique. Two set of reconstructions, noted by different naming of the of the datasets in the following (74X and 76X), have been used in this analysis. The 74X release includes, for the data, the following data sets:

- *Run2015C_05Oct_runs_254227_255031*
- *Run2015D_05Oct_runs_256630_258158*
- *Run2015C_v4_runs_258159_260627*

while, for the Z-boson simulation, it is used the following data set:

- *DYJetsToLL_M50*

The 76X release includes, for the data, the data sets:

- *SingleElectron_Run2015C_16Dec_runs_254227_254914*
- *SingleElectron_Run2015D_16Dec_runs_256630_260627*

For the Z-boson simulation it is used the following data set:

- *DYJetsToLL_M50*

the same used above but with different optimisations, as can be seen later on.

The releases have different characteristics, being the recent one (76X) including all the state-of-the-art corrections both in data and MC. Regarding the Z-boson Monte Carlo simulations, the photon and electron energy regression methods have been re-trained on new data samples and re-optimised in new run conditions. The performances are improved with respect of the 74X older release. The technical details of these optimisations are the following:

- A semi-parametric technique is used to obtain simultaneous prediction of the true energy and the prediction associated uncertainty.
- New particle flow based shower shapes variables are used, instead of the full- 5×5 containment.
- A new training of the supercluster energy regression have been used, where the local and global coordinates, describing the positions of the impinging electrons on the calorimeter crystals, are removed in order not to be affected by potential ECAL misalignments.
- Further studies on the electron resolution as a function of the preshower energy are carried on.

The removing of local/global coordinates shows no loss of performance in the ratio E_{corr}/E_{true} while the preshower improves the resolution by 10% when applied to simulation. The gap effects, namely large deviation from 1 of the ratio E/E_{gen} plotted

with respect to η_{SC} in particular η regions of the calorimeter, are mitigated but not completely suppressed.

After the energy scale corrections, to get more realistic performance studies, a smearing of the invariant mass distribution, tuned with $Z \rightarrow ee$ events, is applied.

Various corrections are applied also in data. Between the two releases, the energy regression is updated, similarly as before, improving the electron resolution also in data while electron energy scale corrections are applied in order to ensure a better matching in the $Z \rightarrow ee$ invariant mass in data and in simulation.

The aim of these corrections is meant to get an agreement in the scale and resolution of the Z resonance, while it is not granted that the peak in the data matches perfectly with the nominal Z -boson mass.

4.6 Correction to ECAL linearity using known SM resonances

The electron scale non-linearity was corrected by studying the di-lepton invariant mass of known SM resonances. The invariant mass is reconstructed from data and Monte Carlo, both for the 74X and the 76X release. The resonances studied are well known in particle physics, that is their mass and width is known with a good precision. The di-lepton invariant mass is reconstructed both for the Z -boson and the J/ψ and Y mesons. These resonances can be produced boosted in the laboratory frame: while the invariant mass is constrained to its value the decay electrons energies to be studied can be used to check the linearity in various energy regions.

The energy scale linearity can be studied in different ways: by comparing the invariant mass of the resonance reconstructed from data (M_{data}^{reco}) to the one reconstructed from the MC (M_{MC}^{reco}), defining the R_{MC} ratio:

$$R_{MC} = \frac{M_{data}^{reco}}{M_{MC}^{reco}} \quad (4.3)$$

or by comparing the invariant mass reconstructed from data and MC to the nominal mass of the resonance (M_{PDG}), that is the tabulated mass from Particle Data Group, using the R_{PDG} ratio:

$$R_{PDG} = \frac{M_{data}^{reco}}{M_{PDG}} \quad (4.4)$$

While the R_{MC} ratio is expected to be flat near 1, since data and Monte Carlo are tuned with all the techniques listed above (cf. sec. 4.5) to be nearly the same, the behaviour of the R_{PDG} ratio is not expected to be distributed near 1. Indeed the absolute energy calibration (G) is derived comparing data and simulation in reference regions of the calorimeter barrel and endcaps where the effects of upstream tracker material are negligible. When extending to the full barrel/endcaps region it is not granted that the di-lepton reconstructed invariant mass agrees with the nominal resonance mass ($M_{ee}^{reco}/M_Z^{PDG} = 1$).

The R_{MC} ratio could be constructed only with the Z -boson signal since this is the only simulated sample currently available for this study. On the contrary, the R_{PDG} ratio exploits the Z , Y and J/ψ resonances.

The first step of the data analysis consists in the selection of good electron pairs from which compute the invariant mass. The following procedure is pursued in every dataset, both for data and simulation:

1. An event with at least two leptons is selected.
2. The leptons are required to be electrons.
3. Since the leptons in each event are stored in descending order with respect to their energy, the two most energetic electrons (the ones likely to come from a resonance decay and not from background events) are selected.
4. Each electron is required to impinge the ECAL into the barrel region, in order to minimize the tracker material effects on its energy/momentum. This is done selecting electrons with a pseudo-rapidity $\eta \leq 1.479$.
5. Each electron is required to belong to the Golden class (cf. sec. 3.1.3), imposing the R_9 shower shape parameter to be $R_9 \geq 0.94$.

Once a good event is selected, the invariant mass of the di-electron pair is reconstructed using the informations gathered both in the calorimeter and in the tracker detector. The resonances studied are produced by two particle decay processes: the invariant mass of the initial particle can be reconstructed from the four-momenta of the decaying particles using the relativistic kinematics:

$$\begin{aligned} (M_{ee}^{reco})^2 &= (\vec{\mathbf{p}}_{1,e} + \vec{\mathbf{p}}_{2,e})^2 \\ &\approx 2(E_{1,e}E_{2,e} - p_{1,e}^{\hat{x}} \cdot p_{2,e}^{\hat{x}} - p_{1,e}^{\hat{y}} \cdot p_{2,e}^{\hat{y}} - p_{1,e}^{\hat{z}} \cdot p_{2,e}^{\hat{z}}) \end{aligned} \quad (4.5)$$

where the $p_{i,e}^{\hat{j}}$ is the projection on the $\hat{j}$ axis of the tri-momentum of the i -electron and the c constant is omitted. Moreover, in the last stage, the electron mass is neglected since its contribute to the total invariant mass of the resonances studied is negligible. The quantities needed in eq. 4.5 are extrapolated from the events selected above:

- The energy of each electron $E_{i,e}$ is extracted from the energy collected in ECAL including all the fine tuning corrections listed in sec. 4.2 and sec. 4.3.
- The components of the momentum of each electron $p_{i,e}^{\hat{j}}$, are computed starting from the emission angles ϕ and θ . Since the dataset contains the pseudo-rapidity η the theta angle is extracted using the following relation:

$$\theta = 2 \cdot \tan^{-1} [\exp(-\eta)]$$

Then the $(\hat{x}, \hat{y}, \hat{z})$ cartesian components are computed from the cylindrical coordinates of the CMS detector:

$$\begin{aligned} p_{i,e}^{\hat{x}} &= p_T^{i,e} \cdot \cos(\phi) \\ p_{i,e}^{\hat{y}} &= p_T^{i,e} \cdot \sin(\phi) \\ p_{i,e}^{\hat{z}} &= \frac{p_T^{i,e}}{\sin(\theta)} \cdot \cos(\theta) \end{aligned}$$

To obtain the distribution of the invariant mass of the resonances, histograms are filled with the number of events collected into an energy interval centered on the nominal value of each resonance (eg. the one from the PDG): for the Z-boson the interval corresponds to $60 \text{ GeV} < M_{ee} < 120 \text{ GeV}$ being $m_Z = 91.1876(21) \text{ GeV}$, for the Y-meson to $5 \text{ GeV} < M_{ee} < 15 \text{ GeV}$, being $m_Y = 9.46030(26) \text{ GeV}$ [78] and for the J/ ψ -meson, since $m_{J/\psi} = 3.096916(11) \text{ GeV}$ [79] the interval corresponds to $1 \text{ GeV} < M_{ee} < 5 \text{ GeV}$.

In order to test the calorimeter energy linearity, the invariant mass of the resonances from data and MC is reconstructed in different energy regions, defined by three kinematic variables related to the transverse momenta of the electron pairs. These kinematic variables are defined by:

- $H_T \equiv p_T^{1,e} + p_T^{2,e}$, the sum of the transverse momenta. With this kinematic variable all the informations on the single electron kinematic is not considered.
- p_{T1} the momentum of the first electron, the one with the highest energy. In this case the kinematic of the second electron is integrated over all the possible phase space.
- p_{T12} the momentum of the first or the second electron. Following this kinematic variable the histogram is filled if the transverse momentum of any of the two selected electrons belonging to the selected kinematic region. The main features of this approach lie into more available statistics and the better knowledge of the identity of the electrons, with respect to the p_{T1} method, since the information of the first(second) electron is integrated over the possible phase space of the second(first) preserving, on average, the identities of the pair.

The energy regions defined by the kinematic variables listed are chosen in order to guarantee a good statistics in each of them. Once selected, the same regions are considered both for the H_T , p_{T1} and p_{T12} sets of histograms. In tab. 4.1 the energy regions for the different resonances are shown.

Energy regions		
Resonance	Min (GeV)	Max (GeV)
Z-boson	40	60
	60	80
	80	100
	100	120
	120	140
	140	160
	160	200
Y-meson	0	20
	20	40
J/ ψ -meson	0	40
	40	100

Table 4.1: Different energy regions in which the invariant mass of the resonances are reconstructed. In these regions the invariant masses are reconstructed for both the H_T , p_{T1} and the p_{T12} kinematic variables.

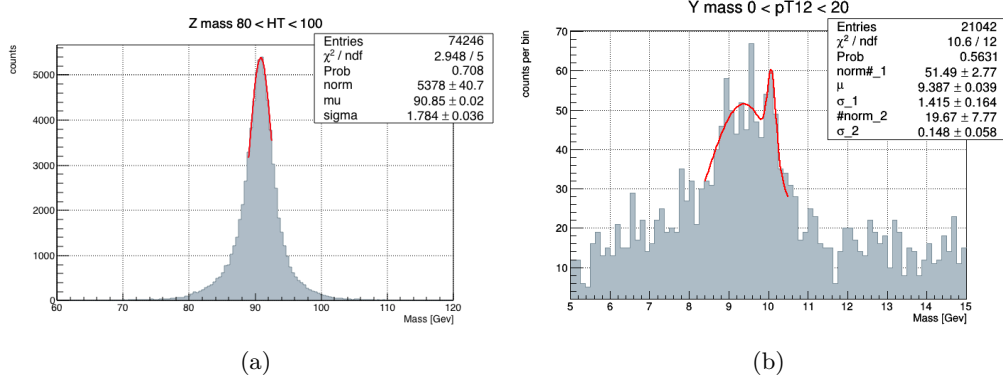


Figure 4.4: An example of fitted invariant mass histograms. In fig. 4.5(a) the Z-boson invariant mass reconstructed as a function of H_T kinematic variable in the interval $40 \text{ GeV} < H_T < 60 \text{ GeV}$, fitted with a Gaussian function. In fig. 4.4(b) the Y-meson invariant mass reconstructed as a function of p_{T12} kinematic variable in the interval $0 \text{ GeV} < p_{T12} < 20 \text{ GeV}$, fitted with the sum of two Gaussian functions.

The invariant mass in each histogram is then fitted, in order to obtain the mean value of the distribution and the related statistical error. Different approaches are used to fit the resonances signals: the Z-boson and the J/ψ -meson are fitted with a Gaussian distribution, while the Y-meson is fitted with the sum of two different Gaussian distribution: the signal appears as two peaks close to each other, the first corresponding to the $Y_{(1S)}$ at $m_{Y_{(1S)}} = 9460.30(26) \text{ MeV}$ and a mixture of the $Y_{(2S)}$ and $Y_{(3S)}$ (due to limitations to the energy resolution) at approximately $m_{Y_{(1S)}} = 10\,023.26(31) \text{ MeV}$. One of the two Gaussian functions used to fit the Y-meson signal fits the peak position of the (1S) resonance and correspond to the Y-meson mass in the following of the analysis. The mean value of the other Gaussian distribution is constrained to be at fixed distance from the first one, in order to fit the (2S) and (3S) signal.

The Gaussian distribution meets perfectly the requirements of the analysis since the only parameters that enters the R_{MC} or R_{Data} ratios are the mean invariant mass and its error. No additional informations from the tails of the invariant mass distributions (which are, of course, not Gaussian at all) are needed for the purpose of this analysis. An example of a fitted histogram containing the invariant mass distribution of the Z-boson and the Y-meson are shown in fig. 4.4.

Once the invariant mass is reconstructed and fitted, the R_{MC} and the R_{PDG} ratios are evaluated in all the energy regions for all the kinematic variables. The statistical error associated with the R_{MC} and the R_{PDG} ratios are evaluated according to eq. 4.6:

$$\sigma_{R_{MC}} = \sqrt{\frac{\sigma_{M_{data}^{reco}}^2}{(M_{MC}^{reco})^2} + \frac{\sigma_{M_{MC}^{reco}}^2}{(M_{data}^{reco})^4} \cdot \sigma_{M_{MC}^{reco}}^2} \quad (4.6)$$

$$\sigma_{R_{PDG}} = \sqrt{\frac{\sigma_{M_{data}^{reco}}^2}{M_{PDG}^2} + \frac{\sigma_{M_{PDG}^{reco}}^2}{M_{PDG}^2} \cdot \sigma_{M_{data}^{reco}}^2}$$

The R_{PDG} and R_{MC} ratios, obtained in the energy regions of tab. 4.1, are then plotted as functions of the H_T , p_{T1} and p_{T12} kinematic variables. Since data and MC are tuned to agree with each other, the ratios are supposed to be flat. The trend of the ratios in the different energy regions is then an excellent tool to investigate the residual non-linearity of the energy reconstruction in ECAL. Rather than using merely the spread of the ratios points as an estimate of the systematic uncertainty in the energy reconstruction for simulation, as previously seen in [29], seem much reasonable trying to derive a further correction, purely data dependent, in order to correct the lepton energy measured in the calorimeter and reconstructed in simulation. A correction function $f(p_T)$ is then extracted from the R_{MC} ratio and applied to data: this procedure is expected to improve the agreement of the invariant mass reconstructed from MC with the invariant mass reconstructed from data.

The correction function is applied to the simulation starting from eq. 4.7:

$$R \equiv \frac{M_{data}}{M_{MC}} = f(p_{T12}) \rightarrow M_{MC}^{corr} = M_{MC} \cdot f(p_{T12}) \quad (4.7)$$

In order to extract a correction for each single electron, the contribution in $f(p_{T12})$ of the first electron is supposed to be shared from the one of the second, according to eq. 4.8:

$$f(p_{T12}) \approx \sqrt{f(p_T^{1,e})} \cdot \sqrt{f(p_T^{2,e})} \quad (4.8)$$

The implementation above is technically legitimate only if the correction function is extracted from a ratio trend function of a kinematic variable that does not hamper at all the single electrons identities. Once splitted in two parts, it is possible to apply the correction function to the single electron energy, following the right-hand side of eq. 4.7:

$$\sqrt{E_{1,e}^{corr} E_{2,e}^{corr} - \vec{p}_{1,e} \cdot \vec{p}_{2,e}} = \sqrt{E_{1,e}^{bare} E_{2,e}^{bare} - \vec{p}_{1,e} \cdot \vec{p}_{2,e}} \cdot \sqrt{f(p_T^{1,e})} \sqrt{f(p_T^{2,e})}$$

where E^{bare} is the electron energy before the correction and E^{corr} the quantity of interest. Getting rid of the square roots and collecting the terms:

$$E_{1,e}^{corr} E_{2,e}^{corr} = E_{1,e}^{bare} f(p_T^{1,e}) E_{2,e}^{bare} f(p_T^{2,e}) + \vec{p}_{1,e} \cdot \vec{p}_{2,e} \underbrace{(1 - f(p_T^{1,e}) f(p_T^{2,e}))}_{\approx 0}$$

where the last term, containing the scalar product of the electrons tri-momenta measured in the tracker, is neglected supposing the correction function nearly flat over the energy regions. The correction to the ECAL reconstructed energy is finally shown in eq. 4.9

$$E_{i,e}^{corr} = E_{i,e}^{bare} \cdot f(p_T^{i,e}) \quad (4.9)$$

In order to extract the correction function, different approaches can be followed:

- The simpler and easier correction scheme is represented by a modification of the electron reconstructed energy considering the deviation of the ratio R bin-by-bin. The $f(p_T)$ is then tabulated in the different energy regions defined and applied to each electron when its transverse momenta belong to the corresponding energy range. The main disadvantage is a poor sensitivity on small electron p_T variations, being limited by the ranges defined in tab. 4.1.

- The extraction of $f(p_T)$ from a fit of the ratio distribution is a more solid approach. Once defined a fitting function, the best-fit parameters are chosen to represent the correction function to be applied on the simulated electrons energies reconstructed in the ECAL. Following this approach the correction depends on every single value of p_T .

The ratio distribution is assumed to be well represented by a linear function mainly because the limited number of fitted points (less than twelve) does not give an amount of statistics able to justify different functions with more parameters (polynomial functions). The linear function used is defined as:

$$f(p_T) = m \cdot p_T + q \quad (4.10)$$

The m and q are the parameters of interest and their best-fit estimation entirely define the electron correction function. Since we expect a good agreement between the invariant masses reconstructed from data and simulation, the foresee values of these parameters are approximately 0 for m and approximately 1 for q .

Moreover, while the R_{MC} ratios can be fitted with a linear function, the R_{PDG} trend does not permit any easy assumption on the data distribution: in this second instance a linear fit would not be consistent and, in order to define $f(p_T)$, a simple bin-by-bin approach is chosen.

Since the R_{MC} ratios are obtained in distinct energy regions selected from different kinematic variables (H_T , p_{T1} and p_{T12}), different correction functions can be extracted which could differ greatly from one to another and are somewhat dependent on the peculiar characteristics of the kinematic variables used to construct the R_{MC} ratios. Indeed, the correction function determined from the H_T or p_{T1} selected regions ratios do not contain all the informations about the identities of the single electrons, the kinematic variable being the sum of the electrons p_T in the former case, or the integration of the second electron transverse momentum over all the phase space in the latter. The splitting shown in eq. 4.8 cannot be easily justified if the correction function is derived with the kinematic variables described above. On the contrary, this splitting is more reasonably justified if the correction function is extracted from the p_{T12} selected regions ratio, since this kinematic variable contains, on average, the informations about the identities of every single electron. For these reasons the function used to correct the reconstructed energy in the calorimeter, in the following analysis, is chosen to be the one derived with the p_{T12} kinematic variable.

Once defined a suitable correction function, the invariant mass of the simulated electron pairs is reconstructed with the same techniques listed above but substituting the electron energies in eq. 4.5 with the corrected ones of eq. 4.9, as a function of the electron p_T . The ratio R_{MC} is then reconstructed with the corrected quantities in order to assess the quality of the corrections applied and to check the data to MC agreement. After the correction, the statistical errors on the ratios are evaluated using eq. 4.6 while the systematic errors related to the energy correction procedure, are evaluated in a different way. From the linear fit a 1σ band is constructed modifying the estimated fit parameters $m \pm \sigma_m$ and $q \pm \sigma_q$ within their errors (cf. eq. 4.10). Since they are verified to be anti-correlated at more than 95%, two additional linear functions are defined by:

$$\begin{aligned} f_a(p_T) &= (m - \sigma_m) \cdot p_T + (q + \sigma_q) \\ f_b(p_T) &= (m + \sigma_m) \cdot p_T + (q - \sigma_q) \end{aligned} \quad (4.11)$$

The simulated bare electron energy is then separately corrected using $f_a(p_T)$ and $f_b(p_T)$ from eq. 4.11 and two additional classes of corrected R_{MC} ratios are constructed. The systematic error is then defined as the absolute value of the difference of the R_{MC} ratio evaluated after the $f_a(p_T)$ and the $f_b(p_T)$ corrections applied separately, according to eq. 4.12:

$$\text{Syst. Err.} \equiv \frac{|R_{MC}^{f_a(p_T) \text{ corr}} - R_{MC}^{f_b(p_T) \text{ corr}}|}{2} \quad (4.12)$$

The total error on the corrected ratios is defined as the squared sum of the statistical and systematic errors.

The invariant mass of the resonances as well as the ratios are reconstructed and determined both for the 74X and 76X releases. Different corrections functions are extracted separately from these two datasets.

Residual electron energy non-linearity in the 74X release

Some examples of the invariant masses reconstructed from data and simulation for the 74X release are shown in fig. 4.5. The whole set of the results is shown in

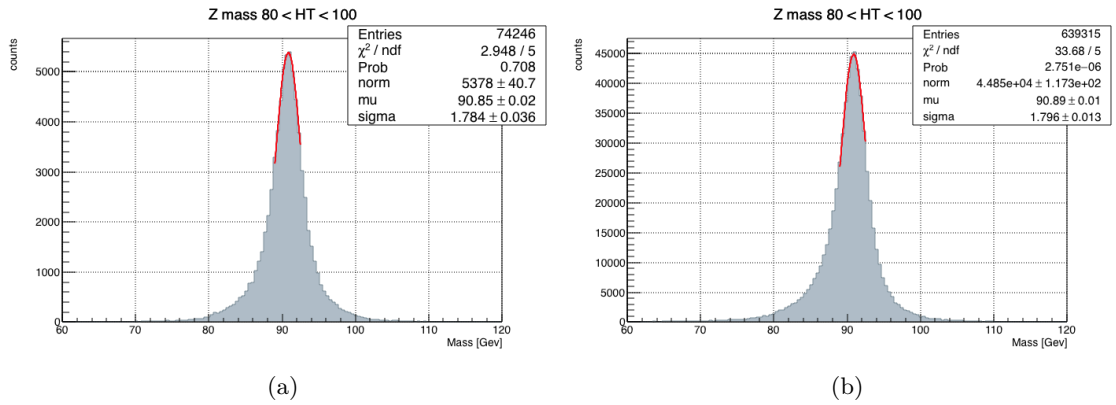


Figure 4.5: Gaussian fits of the Z-boson invariant mass distribution from data (fig. 4.5(a)) and simulation (fig. 4.5(b)) plotted as a function of the H_T of the two electrons for the 74X release.

appendix A: the figs. A.1, A.2 and A.3 gather the fits of the Z-boson invariant mass as a function of the different kinematic variables for the data while figs. A.4, A.5 and A.6 show the results for simulation. Being the position of the peak of the distribution the quantity of interest, the fit range is restricted to the Gaussian core of the invariant mass distribution. This is assumed to correspond approximately to the events included in the full-width-half-maximum of the invariant mass distribution: a suitable fit range (x_{min}, x_{max}) is then reasonably chosen in this region and the Gaussian fit performed. The possible dependency of the best-fit mean position (μ) on the fit interval has been studied modifying the fit range by changing independently the lower and upper limits of $\pm 10\%$ of the best-fit standard deviation (σ). The dependence resulted to be negligible in comparison with the statistical error of the fit.

Once fitted, the average invariant masses of each histogram and the related statistical errors are extracted from the mean parameters of the Gaussian functions (μ_Z). The

invariant masses of the Y-meson reconstructed from data are shown in fig. A.16: the fit function is represented by the sum of two Gaussian functions in order to fit the Y(1S) and the Y(2S) peaks. The average invariant mass of each histogram and the related error are extracted from the mean parameter of the Gaussian distribution fitting the Y(1S) resonance peak ($\mu_{Y(1S)}$). The statistics associated to the J/ψ -meson is too low to be fitted and displayed. Generally speaking, the statistics associated to the Y and J/ψ is very low, both for the 74X and the 76X release: this is a consequence of the single electron selection applied at the HLT stage from which the data used in this analysis are constructed. In order to store an event, the single electron selection trigger requires at least one electron with a transverse measured energy in the calorimeter higher than a certain value. The drawback in this depletion of all the low energy events is the worsening of the available statistics for all the low energy resonances, exactly the Y and the J/ψ mesons. Events selected with the double electron trigger, where at least two electrons are required to be measured in each event allowing a lower energy threshold, were not available at the time of this analysis. In any case, even the thresholds of the double electron HLT trigger path in 2015 and 2016 were too high to select a significant amount of J/ψ -meson data. For the future, a dedicated $J/\psi \rightarrow ee$ trigger is under study.

The R_{PDG} ratios are constructed only for the p_{T12} kinematic variable, being the one providing the better statistics. The results are shown in Fig. 4.6. The statistical errors are evaluated following eq. 4.12. The ratios suggest that the invariant mass of the

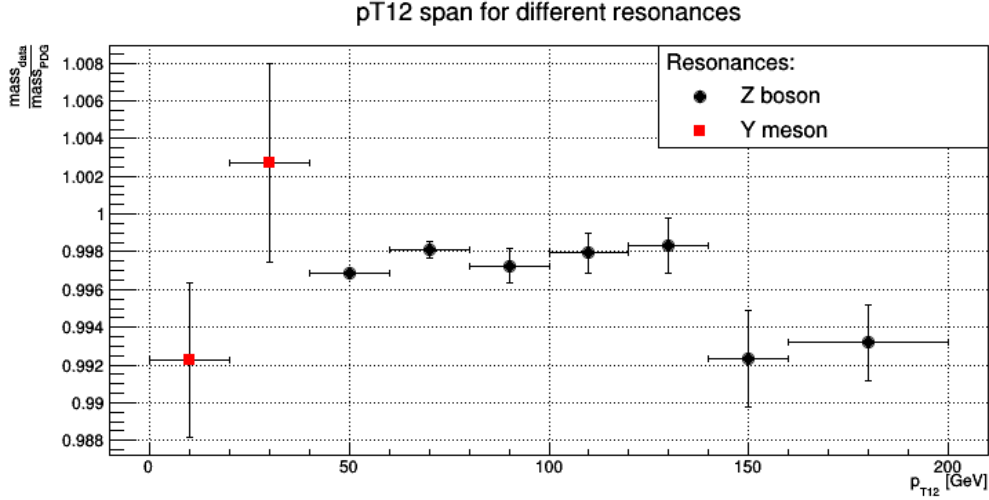


Figure 4.6: The R_{PDG} of the Y and Z resonances for the 74X release as a function of the p_{T12} kinematical variable. The statistical errors on the y coordinate are estimated using eq. 4.6 while the ones on the x coordinate are assumed to be the half of the kinematic variable energy range in which every invariant mass histogram is filled.

Z-boson reconstructed from the combination of tracker and calorimeter data (cf. eq. 4.5) is underestimated with respect to the nominal resonance mass value. The central region, included from 50 to 150 GeV appears to be nearly flat, while the last two points of the Z-boson ratio are considerably shifted from the first five. It can be noticed also that the associated statistical error tends to increase with energy: this is justified since the invariant mass statistics associated to increasing values of the p_{T12} kinematic variable

tends to decrease, hampering the peak position determination resolution. This effect is even more pronounced for the two Y-meson ratios. The R_{PDG} ratios are not expected to be compatible with 1 for the reasons explained before.

The R_{MC} ratios are constructed for each kinematic variable and the results are shown in fig. 4.7. The points are fitted, in each case, with the linear function defined in eq. 4.10 and the best fit results are shown as red lines. Once fitted, the statistical error band of the fit is constructed drawing the auxiliary linear functions (blue and green) defined in eq. 4.11. The process ends up when the parameters of the three correction functions are extracted. Each ratio shows an increasing trend of the ratio of the reconstructed invariant mass with respect to the energy. Moreover the ratio points are not equally distributed around the reference value 1 so a correction is needed. In fig. 4.7(a) the linear increment of the ratio is clearly evident and only the first point, calculated in the energy region 40 GeV to 60 GeV tends to deviate. This is a consequence of the fit resolution of the invariant mass reconstructed in data, as can be seen in fig. A.1(a), however its error bars are large enough to make it still compatible with a linear increment.

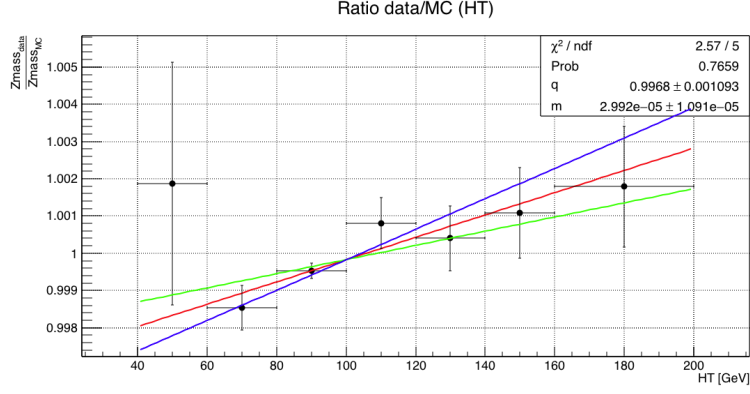
In the p_{T1} and p_{T12} ratios the linear increment is clear only for the first points. In fig. 4.7(b) the 130 GeV and 150 GeV points clearly deviate from the linear fit but are equally distributed above and below the fit function respectively, being compatible with it at least within 3σ . The trend of the last two points belonging to the p_{T12} ratio represent a clear deviation from the linear fit of the previous ones, as shown in fig. 4.7(c): this is likely to be a consequence of the reduced statistics of the related invariant mass histograms resulting in bad peak resolution, as can be seen in fig. A.6(f) and A.6(g). A similar behaviour for the ratios evaluated in high energy regions can be seen in the R_{PDG} ratios of fig. 4.6 (black points) but no further informations could be deduced.

The best fit parameters obtained in the fits of fig. 4.7(c) (the ones belonging to the p_{T12} ratios) are gathered in tab. 4.2. This corrections functions and used to correct the electron simulated energy ($f(p_T)$), according to eq. 4.9, and to asses the statistical uncertainty of this correction ($f_a(p_T)$ and $f_b(p_T)$).

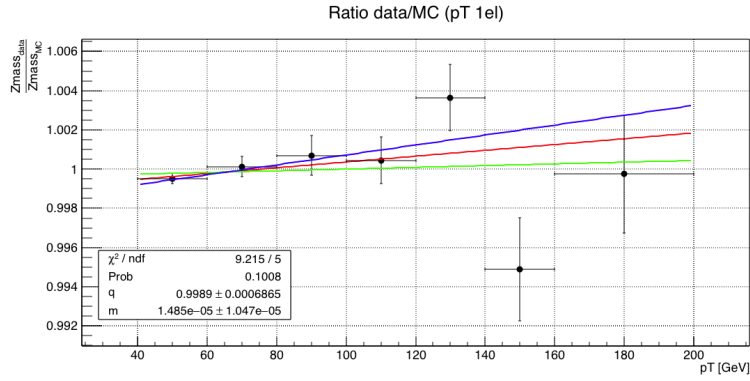
Correction	m	q
$f_a(p_T)$	8×10^{-6}	0.999
$f(p_T)$	2×10^{-5}	0.999
$f_b(p_T)$	3×10^{-5}	0.998

Table 4.2: The best fit parameters obtained from the p_{T12} ratios, used in the correction function of the electron simulated energy for the 74X release.

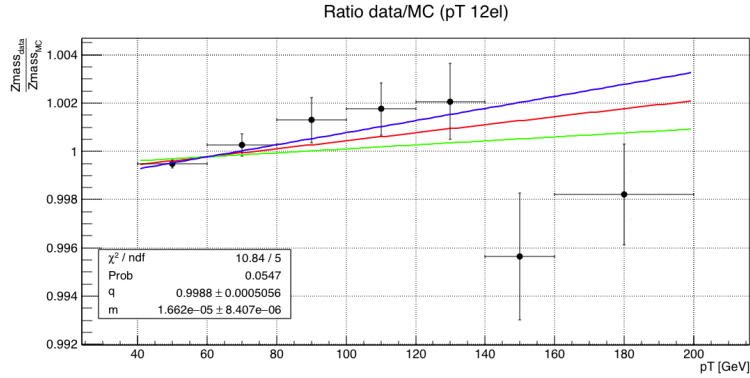
The R_{MC} ratios accessed from the invariant mass distributions obtained correcting the MC electron energy with the $f(p_T)$ function are shown in figs. 4.8(a), 4.8(b) and 4.8(c) for the H_T , p_{T1} and p_{T12} kinematic variables separately. The statistical errors are calculated according to eq. 4.6. In order to define the systematic errors, invariant mass distribution are obtained correcting the MC electron energy with the $f_a(p_T)$ and $f_b(p_T)$ functions respectively (cf, tab. 4.2) and fitted as can be seen in figs. A.10, A.11, A.12, A.13, A.14 and A.15. The systematic errors are the evaluated according to eq. 4.12. In figs. 4.8(a), 4.8(b) and 4.8(c) the quadrature sum of the statical and systematic



(a)



(b)



(c)

Figure 4.7: The R_{MC} before any corrections for the H_T (fig. 4.7(a)), the p_{T1} (fig. 4.7(b)) and the p_{T12} (fig. 4.7(c)) kinematical variable for the 74X release, fitted with a linear function (red). The additional linear functions $f_a(p_T)$ (green) and $f_b(p_T)$ (blue), derived from the best fit parameters of $f(p_T)$ in order to assess the systematic effects on the energy correction (cf. eq. 4.11) are shown, together with the statistical errors associated to each point. The statistical errors on the y coordinate are estimated using eq. 4.6 while the ones on the x coordinate are assumed to be the half of the kinematic variable energy range in which every invariant mass histogram is filled.

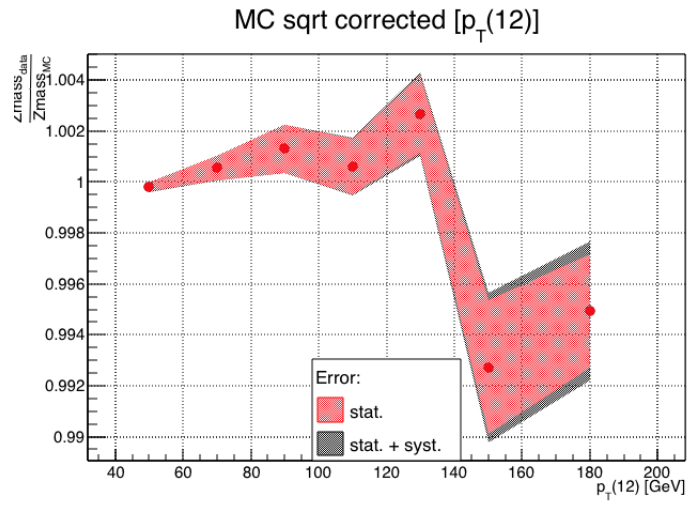
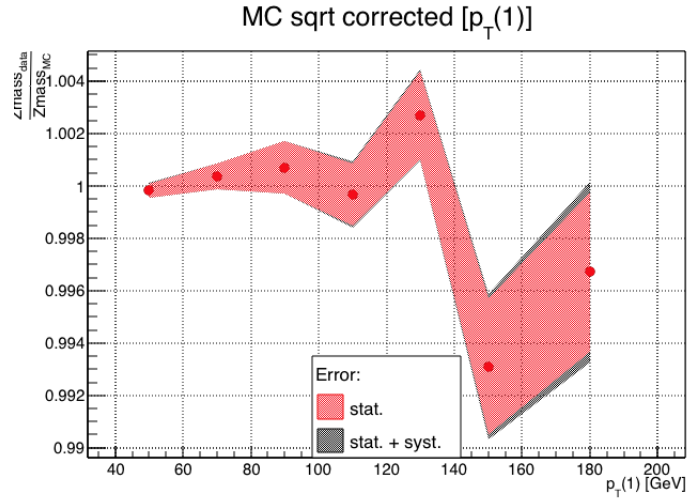
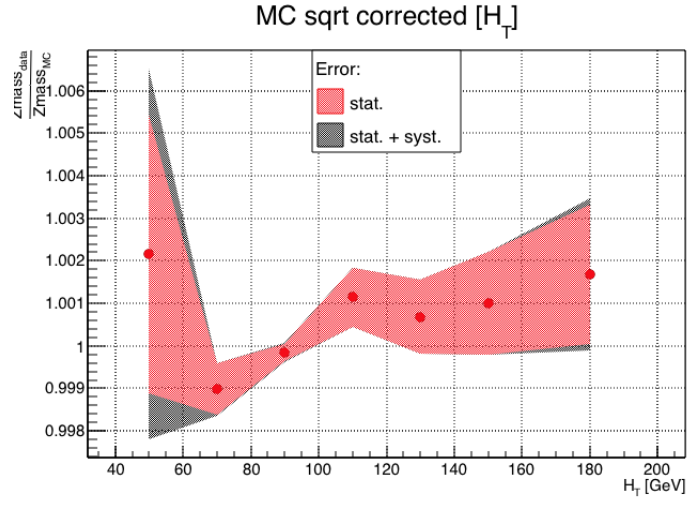


Figure 4.8: The H_T , p_{T1} and p_{T12} R_{MC} ratios after the $f(p_T)$ corrections to the simulated electron energy. The bands correspond to the statistical uncertainties (red) and to the total (stat. + syst.) uncertainties (black).

uncertainties is shown: the statistical uncertainties are higher for the points where the invariant mass peak resolution is lower.

After the energy correction in simulation, a general improvement of the compatibility of the Z-boson invariant mass extracted from data and MC is clear for each point of the H_T and p_{T1} ratios. The p_{T12} ratios are improved as well, apart from the points within the 140 GeV – 200 GeV that are worsened after the correction. This is mainly a consequence of the fact that the linear trend of the non corrected ratios R_{MC} (cf. fig. 4.7(c)) is not followed by the point reconstructed in this region, being the gap with the previous point evident. The linear correction tends to overestimate the position of these ratios so the worsening, after applying the correction, is expected. It can be argued that, for these points, a simple bin-by-bin correction could be applied instead of the linear function correction: this different approach seems not reasonable since the main goal of this analysis is trying to derive a general analytic correction function suitable to correct the energy non-linearity effects, not only for the reconstructed invariant mass of the Z-boson, but as a general approach for the electrons energy deposits in ECAL. In addition it should be noticed that the 140 GeV – 200 GeV energy region is not significantly relevant for the $H \rightarrow 4e$ boson studies being the phase space of the leptons from $m_H = 125$ GeV Higgs decay limited.

The invariant masses reconstructed from the 74X release do not show a perfect agreement between data and simulation both after and before the non-linearity corrections (cf. fig. 4.11). This is expected for many reasons, that is the statistics associated to the invariant masses distributions in every kinematic variable (H_T , p_{T1} and p_{T12}) is not enough, restraining consistently the peak resolution: even small deviations could result in large fluctuations of the R_{MC} ratios. Moreover, data and simulation in this release are not finely tuned to match one to each other as well as possible: the small non-linearity effects could be, at this stage, surmounted by lower order residual mismatching (as an example from not perfect energy containment) all of which contribute is difficult to estimate and separate from the one studied with the linear fits of Fig. 4.7.

In order to study purely the non-linearity effects of the calorimeter in the reconstructed energy, the 76X release is considered, both for its higher available statistics, and for the further corrections applied in data and MC. The results with this latter dataset are expected to be better than before, showing the best so far available agreement between M_{data}^{reco} and M_{MC}^{reco} after the energy correction with the smallest accessible statistical and systematic uncertainties.

Residual electron energy non-linearity in the 76X release

The invariant masses reconstructed from data and simulation for the 76X release are shown in appendix B. The figs. B.1, B.2 and B.3 gather the fits of the Z-boson invariant mass as a function of the different kinematic variables for the data while figs. B.4, B.5 and B.6 show the results for simulation. The fit intervals are again chosen in order to cover only the peaks of the distributions, without considering the non-Gaussian tails. Once fitted, the average invariant mass of each histogram and the related error are extracted from the mean parameter of the Gaussian function (μ_Z). The invariant masses of the Y-meson reconstructed from data are shown in fig. B.16: the fit function is represented by the sum of two Gaussian functions in order to fit the Y(1S) and the Y(2S) peaks. The average invariant mass of each histogram and the related error

are extracted from the mean parameter of the Gaussian distribution fitting the $Y(1S)$ resonance ($\mu_{Y(1S)}$). The statistics associated to the J/ψ -meson is too low to be fitted and displayed even for the 76X release.

The R_{PDG} ratios are constructed only for the p_{T12} kinematic variable, being the one providing the better statistics. The results are shown in fig. 4.9. The ratios suggest that

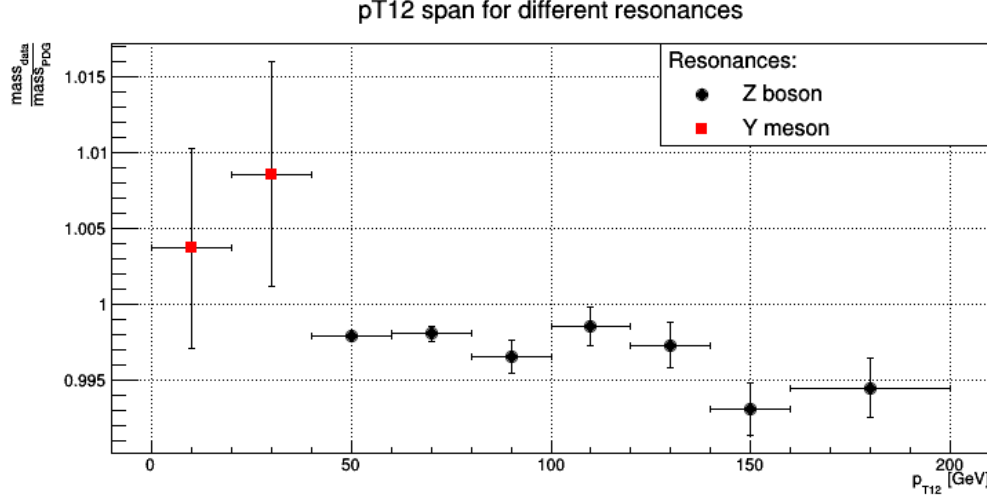


Figure 4.9: The R_{PDG} for the Y and Z resonances for the 76X release as a function of the p_{T12} . The statistical errors on the y coordinate are estimated using eq. 4.6 while the ones on the x coordinate are assumed to be the half of the kinematic variable energy range in which every invariant mass histogram is filled.

the invariant mass of the Z-boson reconstructed from the combination of tracker and calorimeter data (cf. eq. 4.5) is underestimated with respect to the nominal resonance mass value also for the 76X release even if to a lesser extent than the 74X release. The central region from 50 to 150 GeV, appears to be flat, and the last two points of the Z-boson ratio are more compatible with the first five ones. It can be noticed that the associated statistical error is reduced since the invariant mass statistics associated to increasing values of the p_{T12} kinematic variable is higher, giving a better peak position resolution. The two Y-meson ratios are now fully compatible but the associated statistics is still too low, compared to the Z-boson one.

The R_{MC} ratios are constructed for each kinematic variable and the results are shown in Fig. 4.10. The points are fitted, in each case, with the linear function of eq. 4.10 and the best fit results are shown as red lines. The ratios in the energy regions for any of the three kinematic variable are considerably improved in the 76X release, being the increasing trend visible for the 74X release suppressed. The points appear evenly distributed around the reference value 1 and compatible with a straight line. Regarding the p_{T12} ratios the linear increment is now clear for all the energy regions and the gap visible in the 74X release (cf. fig. 4.7(c)) is attenuated. The better agreement between data and MC for the 76X is likely to be a consequence of:

- the optimization of the 76X release with respect to the 74X: both data and simulation are further corrected to improve the agreement between the two.

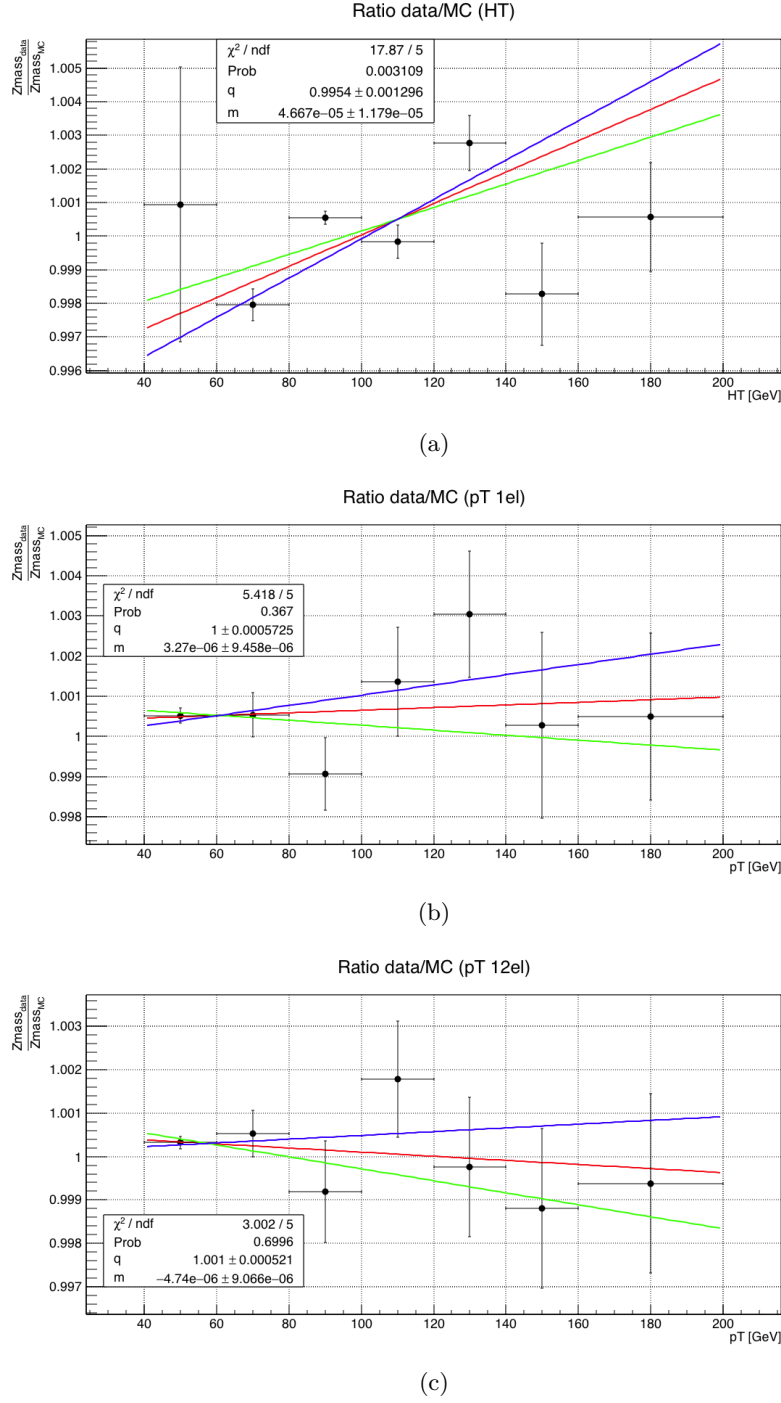


Figure 4.10: The R_{MC} before any corrections, fitted with a linear function (red). The additional linear functions $f_a(p_T)$ (green) and $f_b(p_T)$ (blue), derived from the best fit parameters in order to assess the systematic effects on the energy correction (cf. eq. 4.11) are shown together with the statistical errors associated to each point. The statistical errors on the y coordinate are estimated using eq. 4.6 while the ones on the x coordinate are assumed to be the half of the kinematic variable energy range in which every invariant mass histogram is filled.

- the 76X release offers a larger dataset of data, enhancing the peak position resolution in all the energy regions, as can be seen in fig. B.4, B.5 and B.6.

The best fit parameters obtained in the fits of fig. 4.10(c) (the ones belonging to the p_{T12} ratios) are gathered in tab. 4.3 and used to correct the electron simulated energy, according to eq. 4.9. Comparing the results of tab. 4.3 with the ones obtained in tab. 4.2

Correction	m	q
$f_a(p_T)$	-1×10^{-5}	1.0011
$f(p_T)$	-5×10^{-6}	1.0006
$f_b(p_T)$	4×10^{-6}	1.0000

Table 4.3: The best fit parameters obtained from the p_{T12} ratios, used in the correction function of the electron simulated energy for the 76X release.

the improvement of the 76X release with respect to the 74X is clear both for the m and q parameters of $f(p_T)$.

The invariant mass distributions obtained correcting the MC electron energy with the $f(p_T)$ function for the 76X release are shown in fig. 4.11(a), 4.11(b) and 4.11(c) for the H_T , p_{T1} and p_{T12} kinematic variables separately.

The statistical errors are calculated according to eq. 4.6. The systematic errors are derived in the same way as before. In figs. B.7, B.8 and B.9 the quadrature sum of the statistical and systematic uncertainties is shown.

After the energy correction in simulation, the improvement of the compatibility of the Z-boson invariant mass extracted from data and MC is clear for each point of the H_T , p_{T1} and p_{T12} ratios: this is expected since the residual differences between data and simulation in the 76X release are ascribable only to the residual non-linearity of the ECAL calorimeter. The results before the correction justified the approximations to derive the electron energy correction applied to the MC (eq. 4.9) and the assumption that, in order to estimate the correction function, a linear fit of the data can be used (cf. eq. 4.10). The latter assumption is also supported by the fact that, among the other kinematic variable, the p_{T12} ratios after the energy corrections, show the best agreement between data and simulation.

4.7 Effects of the electron energy scale non-linearity on the Higgs boson mass measurements

The position of signal peak trivially corresponding to the Higgs resonance as measured in the CMS experiment cannot be considered as the nominal invariant mass of the boson, both for the systematic uncertainties related to the events selection and backgrounds estimations, and for the systematic uncertainties of the detector in hampering the energy resolution and position of the signal itself. The Higgs boson mass in the $H \rightarrow ZZ$ decays is then estimated by fitting the peak observed in data with a line-shape obtained by MC simulations made for a given mass hypothesis which includes all the known detector biases to the best of the current knowledge. Various simulations with different H-boson nominal masses are used for this purpose, searching for the signal template

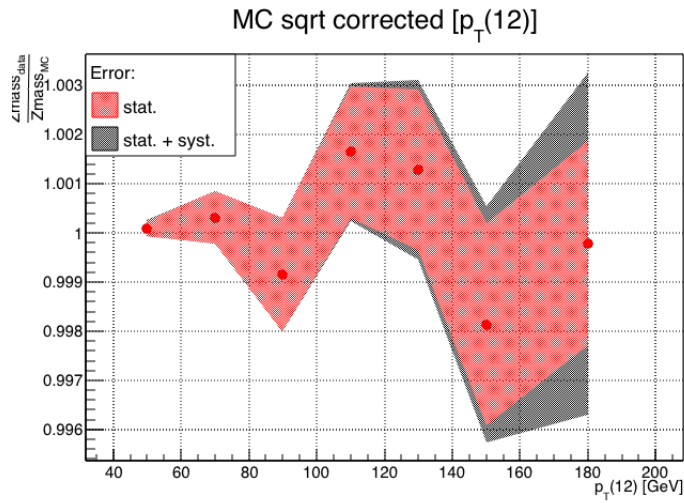
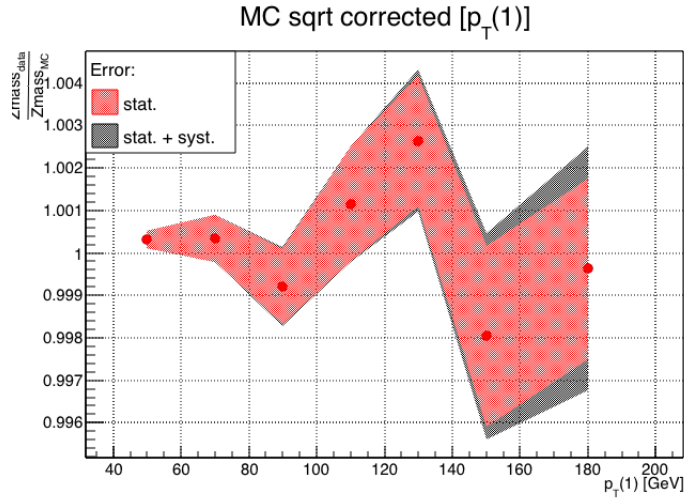
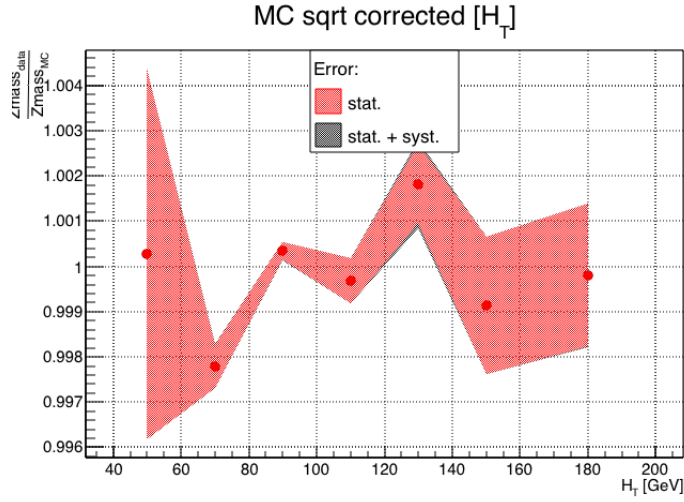


Figure 4.11: The H_T , p_{T1} and p_{T12} R_{MC} ratios after the $f(p_T)$ corrections to the simulated electron energy. The bands correspond to the statistical uncertainties (red) and to the total (stat. + syst.) uncertainties (black).

that minimize the χ^2 when compared to the observed data distribution. The nominal mass, corresponding to the signal template that better mimics the data behaviour, is selected constraining the fitted peak position to be an analytic function of the true Higgs boson invariant mass [29]. Clearly, the nominal mass value evaluated with this method should be validated in order to assess the reliability of the MC, but also to determine the systematic error associated to the MC reconstructed invariant mass. The systematic energy shift related to the MC description of the detector effects, after all the corrections, should be propagated to the Higgs boson invariant mass measurement. This can be done studying the effects of the MC simulation on the reconstructed invariant mass of well known resonances of the SM: the Z-boson and Y-meson studies of the residual non-linearity in the energy reconstruction in ECAL for simulation represents a good starting point to estimate the systematic shift of the reconstructed invariant mass for the simulations of the H-boson lepton decays.

Both the 74X and the 76X release results obtained in the previous section 4.6 are used: the correction function derived from the p_{T12} Z-boson R_{MC} ratios and the bin-by-bin correction derived from the Y-meson and Z-boson R_{PDG} are separately used to correct the Higgs boson mass in the Higgs boson simulated samples (more informations about the simulated H-boson sample used can be found on sec. 4.5). Unlike the Z-boson simulated samples, the H-boson simulated samples are at generator level since data reconstructed at the detector level were not available. This does not affects the systematic uncertainties from the Higgs boson mass obtained in the analysis.

Since they are derived from the R_{PDG} and the R_{MC} ratios obtained in different energy regions, the energy correction to the simulated electron energy can be applied to the whole p_T emission spectrum of the Higgs boson decaying electrons.

The invariant mass distribution of the Higgs boson is reconstructed from the square sum of the electron four-momenta: an histogram is filled with the invariant mass of each event belonging to a mass interval of 100 GeV-150 GeV, compatible with the so far know Higgs boson mass (cf. sec. 1.3.2). In order to evaluate the average value of the invariant mass distribution before and after the energy corrections, a Gaussian fit is performed around the peak position, without considering the distribution tails.

The energy corrections are applied to the electrons p_T according to eq. 4.13.

$$p_T^{corr} = p_T \cdot f(p_T) \quad (4.13)$$

The correction function $f(p_T)$ is linear (cf. eq. 4.10) and the parameters are defined in two different ways:

- For the bin-by-bin correction derived from the R_{PDG} ratios, when the electron transverse momentum belongs to a certain bin of the p_{T12} ratios the corresponding ratio value is used as the correction function in eq. 4.13.
- For the correction function derived from the R_{MC} ratios, the best fit parameters shown in tab. 4.2 and 4.3 are used to define the $f(p_T)$.

In order to assess the statistical uncertainty associated to the energy correction obtained from the Z-boson R_{MC} ratios, the auxiliary correction functions defined in eq. 4.11 are used to separately correct the electrons energies. The statistical error is

then defined according to eq. 4.14

$$\text{Stat. Err.} \equiv \frac{|M_{\text{H}}^{f_a(p_T) \text{ corr}} - M_{\text{H}}^{f_b(p_T) \text{ corr}}|}{2} \quad (4.14)$$

Another quantity is defined in order to determine the systematic effect related to the energy correction of the electrons from the Higgs decay:

$$\delta_m \equiv |m_{\text{H}}^{\text{no. corr.}} - m_{\text{H}}^{\text{corr.}}| \quad (4.15)$$

where $m_{\text{H}}^{\text{no. corr.}}$ is the reconstructed Higgs boson invariant mass without any corrections and $m_{\text{H}}^{\text{corr.}}$ the Higgs boson invariant mass with the $f(p_T)$ correction applied to the electrons energies. The masses are obtained from the mean value of a Gaussian fit of the peak of the invariant mass distributions.

Regarding the energy correction obtained from the Z-boson R_{MC} ratios, the statistical uncertainty related to the systematic shift $\delta_m^{R_{MC} \text{ corr.}}$ is assumed to be the one found in eq. 4.14. The statistical uncertainty of the energy correction derived from the R_{PDG} ratios cannot be assessed via eq. 4.14 since, with such method, only one correction function is defined (the bin-by-bin ratios). Using this correction method the error on the systematic shift $\delta_m^{R_{PDG} \text{ corr.}}$ is reasonably defined as half of the $\delta_m^{R_{PDG} \text{ corr.}}$ value.

The analysis is performed selecting different Higgs boson decay channels, the $\text{H} \rightarrow \text{ZZ}^* \rightarrow 4\text{e}$ where the leptons energies are corrected and the $\text{H} \rightarrow \text{ZZ}^* \rightarrow 2\text{e}2\mu$ where only the electrons energies are corrected. In this context, the systematic energy shifts related to the energy corrections are expected to be different in the two decay channel. In particular, the ones related to the $\text{H} \rightarrow 4\text{e}$ channel is expected to be larger with respect to the $\text{H} \rightarrow 2\text{e}2\mu$ ones, in the latter being corrected the energies of two (the electrons) out of four leptons.

The invariant mass distributions of the Higgs boson reconstructed from the $\text{H} \rightarrow \text{ZZ}^* \rightarrow 4\text{e}$ and the $\text{H} \rightarrow \text{ZZ}^* \rightarrow 2\text{e}2\mu$ decays without any correction, fitted with a Gaussian function, are shown in fig. 4.12. The mean value of the Gaussian distributions of fig. 4.12, together with their statistical errors, are gathered in eq. 4.16:

$$\begin{aligned} m_{\text{H} \rightarrow 4\text{e}}^{\text{no. corr.}} &= 124.44 \pm 0.05 \text{ GeV} \\ m_{\text{H} \rightarrow 2\text{e}2\mu}^{\text{no. corr.}} &= 124.68 \pm 0.02 \text{ GeV} \end{aligned} \quad (4.16)$$

Electron energy correction function from the 74X release

The bin-by-bin correction from the R_{PDG} ratios is derived from the ratio-points shown in fig. 4.6 while the $f(p_T)$ correction from the R_{MC} ratios is assessed via the parameters values gathered in tab. 4.2. The reconstructed invariant masses after the separated application of these corrections are shown in figs. 4.13 and 4.14. The mean value of the Gaussian distributions in both the two considered decay channels after the corrections, with their statistical errors are shown in tab. 4.4.

The R_{PDG} and the R_{MC} corrected values are not expected to be compatible since, as explained, data in each release are corrected in order to improve their agreement with simulation, without considering the nominal PDG value of the Z and Y resonances.

The systematic shifts of the R_{MC} energy corrections are then computed according to eq. 4.15 while their related statistical uncertainties are evaluated according to eq. 4.14.

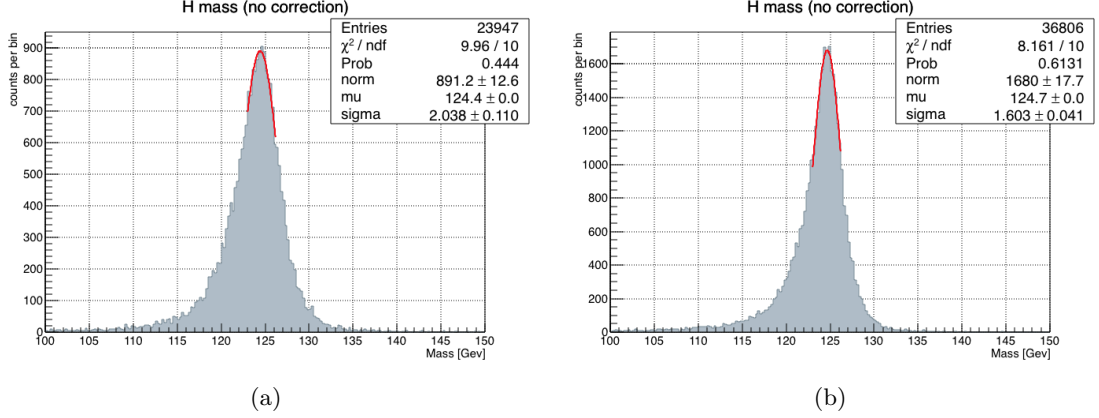


Figure 4.12: The invariant mass distribution of the Higgs boson reconstructed from the $H \rightarrow ZZ^* \rightarrow 4e$ (fig. 4.12(a)) and the $H \rightarrow ZZ^* \rightarrow 2e2\mu$ (fig. 4.12(b)) decays from simulation without any correction applied. In order to assess the peak position a Gaussian fit (red) is performed avoiding the distribution tails.

74X Correction	Higgs mass	
	$H \rightarrow ZZ^* \rightarrow 4e$	$H \rightarrow ZZ^* \rightarrow 2e2\mu$
bin-by-bin (R_{PDG})	124.15(5) GeV	124.52(2) GeV
$f(p_T)$ (R_{MC})	124.38(5) GeV	124.64(2) GeV

Table 4.4: The mean value of the Higgs boson invariant mass distribution after the corrections derived from the 74X release in both the two Higgs boson decay channels considered. The values correspond to the mean of the Gaussian distributions of figs. 4.13(b) and 4.14(b).

The values obtained for the 74X release are:

$$\begin{aligned}\delta_m^{H \rightarrow 4e} &= 60 \pm 9 \text{ MeV} \\ \delta_m^{H \rightarrow 2e2\mu} &= 40 \pm 10 \text{ MeV}\end{aligned}\tag{4.17}$$

Concerning the R_{PDG} bin-by-bin correction, the systematic energy shifts are obtained via eq. 4.15 while their related statistical uncertainties are assumed to be half of the δ_m values. The results are shown in eq. 4.18:

$$\begin{aligned}\delta_m^{H \rightarrow 4e} &= 290 \pm 145 \text{ MeV} \\ \delta_m^{H \rightarrow 2e2\mu} &= 150 \pm 75 \text{ MeV}\end{aligned}\tag{4.18}$$

The values found in eqs. 4.17 and 4.18 are considerably different, however a comparison between the two cannot be justified since they are derived under different assumptions.

Electron energy correction function from the 76X release

In the 76X release the bin-by-bin corrections from the R_{PDG} ratios are derived from the ratio-points shown in fig.4.9 while the $f(p_T)$ correction from the R_{MC} ratios is

assessed via the parameters values gathered in tab. 4.3. The reconstructed invariant masses after the separated application of these corrections are shown in figs. 4.15 and 4.16. The mean value of the Gaussian distributions in both the two considered decay channels after the corrections, together with their statistical errors, are shown in tab. 4.5.

76X Correction	Higgs mass	
	$H \rightarrow ZZ^* \rightarrow 4e$	$H \rightarrow ZZ^* \rightarrow 2e2\mu$
bin-by-bin (R_{PDG})	124.80(4) GeV	124.82(2) GeV
$f(p_T)$ (R_{MC})	124.45(5) GeV	124.71(2) GeV

Table 4.5: The mean value of the Higgs boson invariant mass distribution after the corrections derived from the 76X release in both the two Higgs boson decay channels considered.

The R_{PDG} and the R_{MC} corrected values are not expected to be compatible, as explained.

The systematic energy shifts from the R_{MC} energy corrections are then computed according to eq. 4.15 while their statistical uncertainties are evaluated according to eq. 4.14. The values obtained for the 74X release are:

$$\begin{aligned}\delta_m^{H \rightarrow 4e} &= 50 \pm 20 \text{ MeV} \\ \delta_m^{H \rightarrow 2e2\mu} &= 30 \pm 10 \text{ MeV}\end{aligned}\tag{4.19}$$

Concerning the R_{PDG} bin-by-bin correction, the systematic energy shifts are obtained via eq. 4.15 while their related statistical uncertainties are assumed to be half of the δ_m values. The results are shown in eq. 4.20:

$$\begin{aligned}\delta_m^{H \rightarrow 4e} &= 360 \pm 180 \text{ MeV} \\ \delta_m^{H \rightarrow 2e2\mu} &= 140 \pm 70 \text{ MeV}\end{aligned}\tag{4.20}$$

The values found in eqs. 4.19 and 4.20 are considerably different however, as for the values obtained with the 74X release, a comparison between the two cannot be justified since they are derived with different methods.

Comparing eqs. 4.17 and 4.19, the corrections applied on the 76X release are considerably improved, as expected, the R_{MC} ratios showing a better agreement between M_{data}^{reco} and M_{MC}^{reco} . This is justified being the derived correction function obtained for the 76X release nearly flat and the systematic effects related to the electrons energies corrections on the Higgs boson mass smaller.

During the LHC Run 1 measurement of the properties of a Higgs boson in the four-lepton final state in pp collisions at a center-of-mass energy of $\sqrt{s} = 7 \text{ TeV}$ and 8 TeV , the residual systematic uncertainty in the Higgs boson signal mass scale was assessed conservatively from the residual difference between data and simulation of the dilepton mass scale for some SM resonances (cf. sec. 4.4.2). The uncertainties correspond to approximately [29]:

$$\begin{aligned}\delta_m^{H \rightarrow 4e} &= 300 \text{ MeV} \\ \delta_m^{H \rightarrow 2e2\mu} &= 100 \text{ MeV}\end{aligned}\tag{4.21}$$

During the LHC Run 2 the uncertainties on the energy scale was nearly the same [69].

The method developed in this analysis improves the evaluation of the systematic shifts on the four-lepton Higgs boson mass, resulting from the p_T -dependence of the electron momentum scale, with respect to the results obtained in previous works. It also assigns a statistical uncertainty to the systematic shift. Even though being still room for improvements, its use is desirable in future precision measurements of the Higgs boson mass.

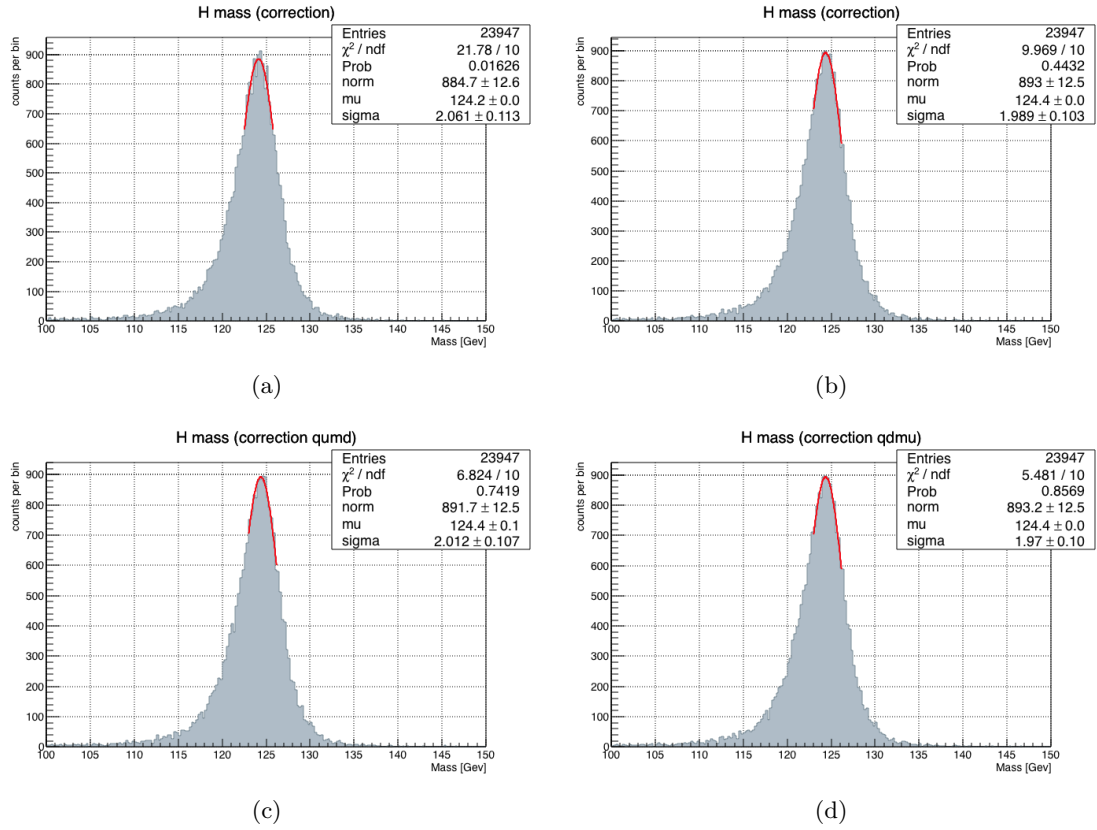


Figure 4.13: The invariant mass distribution of the Higgs boson reconstructed in the $H \rightarrow ZZ^* \rightarrow 4e$ channel, from simulation after the electrons energy correction derived from the 74X release. Fig. 4.13(a) shows the invariant mass distribution after the bin-by-bin correction derived from the R_{PDG} ratios. Fig. 4.13(c), 4.13(b) and 4.13(d) the invariant mass distributions after the energy corrections $f_a(p_T)$, $f(p_T)$ and $f_b(p_T)$ respectively, derived from tab. 4.2.

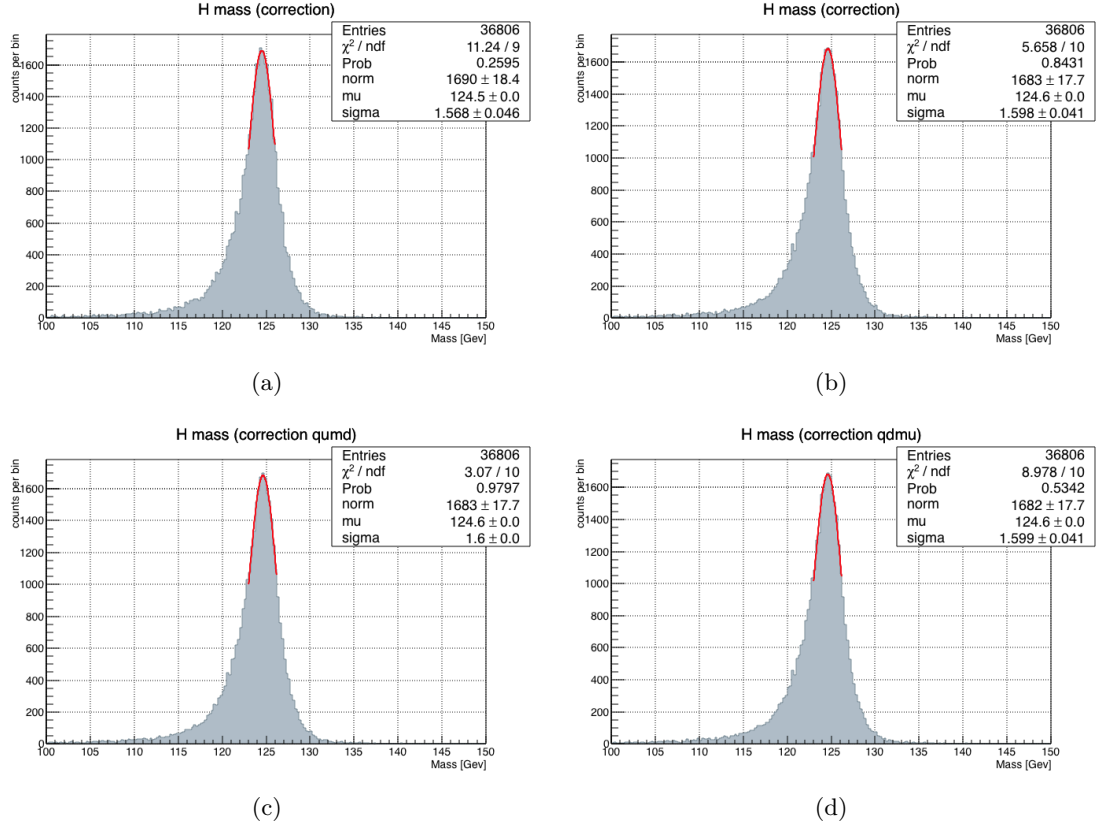


Figure 4.14: The invariant mass distribution of the Higgs boson reconstructed in the $H \rightarrow ZZ^* \rightarrow 2e2\mu$ channel from simulation after the electrons energy correction derived from the 74X release. Fig. 4.14(a) shows the invariant mass distribution after the bin-by-bin correction derived from the R_{PDG} ratios. Fig. 4.14(c), 4.14(b) and 4.14(d) the invariant mass distributions after the energy corrections $f_a(p_T)$, $f(p_T)$ and $f_b(p_T)$ respectively, derived from tab. 4.2.

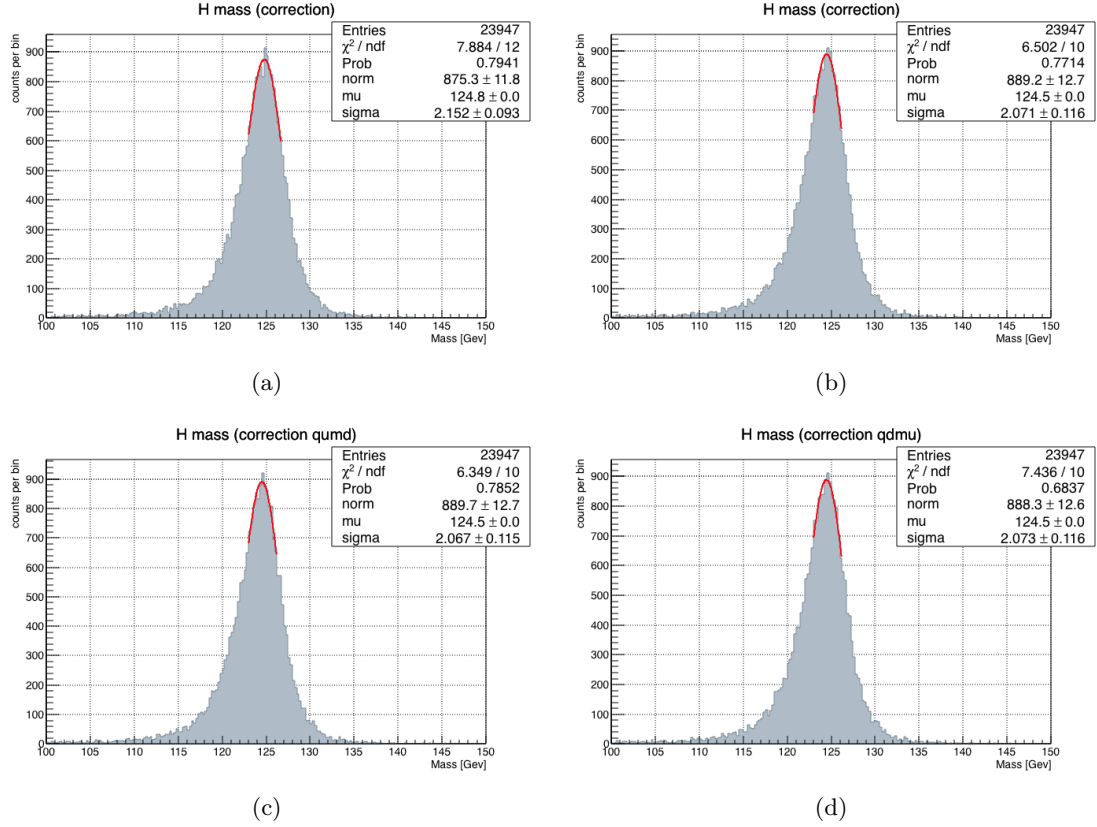


Figure 4.15: The invariant mass distribution of the Higgs boson reconstructed in the $H \rightarrow ZZ^* \rightarrow 4e$ channel, from simulation after the electrons energy correction derived from the 76X release. Fig. 4.13(a) shows the invariant mass distribution after the bin-by-bin correction derived from the R_{PDG} ratios. Fig. 4.15(c), 4.15(b) and 4.15(d) the invariant mass distributions after the energy corrections $f_a(p_T)$, $f(p_T)$ and $f_b(p_T)$ respectively, derived from tab. 4.3.

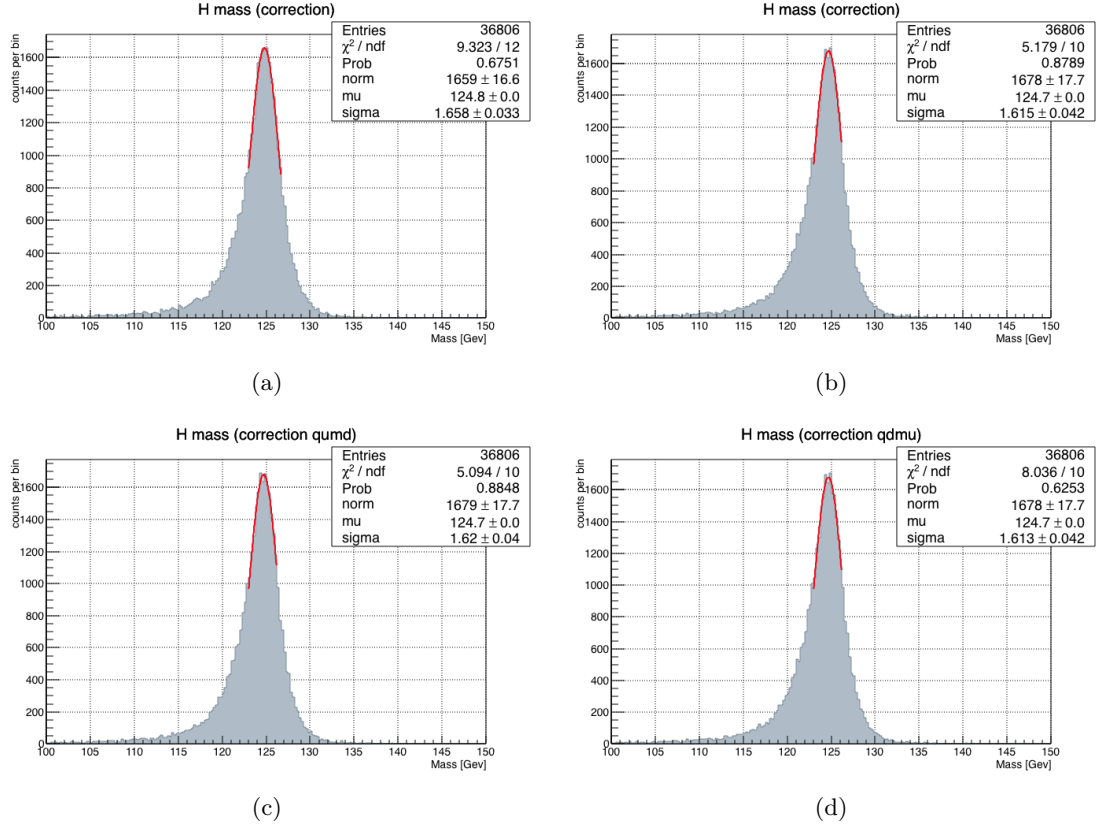


Figure 4.16: The invariant mass distribution of the Higgs boson reconstructed in the $H \rightarrow ZZ^* \rightarrow 2e2\mu$ channel from simulation after the electrons energy correction derived from the 76X release. Fig. 4.16(a) shows the invariant mass distribution after the bin-by-bin correction derived from the R_{PDG} ratios. Fig. 4.16(c), 4.16(b) and 4.16(d) the invariant mass distributions after the energy corrections $f_a(p_T)$, $f(p_T)$ and $f_b(p_T)$ respectively, derived from tab. 4.3.

5 | Summary and conclusions

This work is focused on the study of the linearity of the energy scale of the CMS electromagnetic calorimeter (ECAL), using known Standard Model resonances. While several methods are used to calibrate the calorimeter and improve the detector knowledge encoded in the Monte Carlo simulator, small residual discrepancies between data and simulations are observed. The residual non-linearity effect plays a fundamental role in estimating the systematic uncertainty on the energy scale for reconstructed electrons.

The fine-tuning calibration and its effects on the detector energy scale are studied in the environment of the LHC Run 2 upgrade, using data and simulations of resonances decays at a center of mass energy of $\sqrt{s} = 13$ TeV for a luminosity of approximately 2 fb^{-1} . The electrons energies are reconstructed using two different software releases, 74X and 76X, both including only the simulation of Z-boson decays, the latter including all the state of the art improvements for data and MC agreement. The electrons energy scale can be studied in different energy regions defined by three kinematical variables. Two quantities are defined, the ratio of the invariant mass reconstructed from data to the ones reconstructed from simulation (R_{MC}) and the ratio of the invariant mass reconstructed from data to the nominal mass of the resonance, as tabulated from the Particle Data Group (R_{PDG}). In each ratio, a residual mismatching between data and simulation is observed. The R_{MC} shows an increasing trend of the reconstructed MC invariant mass for all the kinematical variables, for both software releases. A linear correction function is extracted and applied to the reconstructed electron energies from simulated events: the R_{MC} ratios are then studied after the corrections, together with the systematic errors related to this method. The best agreement are observed for the 76X release, the systematic effects being relevant only at high transverse energies.

An energy correction function is then applied to the electrons energies from simulated Higgs boson decays in the gold-plated channel $H \rightarrow ZZ^* \rightarrow 4l$. This is the more valuable channel to obtain precise measurements of the Higgs boson mass: these measurements are important since from m_H all the Higgs boson properties can be determined. To study the effects of the energy correction, the invariant mass is reconstructed before and after the corrections: the systematic effects on the energy scale is estimated to be few tens of MeV. The analysis provides a good estimate of the energy scale systematic effects.

This work represents only the first step towards the fine-tuning calibration of the energy scale of the ECAL and several improvements are suggested: it could be interesting to study the R_{MC} ratios also for other SM resonances, together with the Z-boson, in order to improve the reliability on the energy correction function extracted. Moreover, datasets with an increased statistics of the Y- and J/ ψ -meson signals are needed, being

the latter not analysed in this work due to the lack of data. Additionally, new simulations of the Higgs boson four-leptons decays at 13 TeV and at a detector level are needed to validate the preliminary systematic uncertainty on the energy scale estimated in this work.

A | Fits of the resonances invariant mass distribution for the 74X release

The Z-boson invariant mass distribution is fitted using a Gaussian function. The H_T set of histograms of the invariant mass for data and MC are shown in fig. A.1 and fig. A.4, the p_{T1} set of histograms in fig. A.2 and fig. A.5 and the p_{T12} set of histograms in fig. A.3 and fig. A.6 respectively.

The histograms in fig. A.7, A.8, and A.9 contain the invariant mass of the Z-boson for the H_T , p_{T1} and p_{T12} scans, with the energy correction applied to the decay electrons.

The histograms in fig. A.10, A.11 and A.12 contain the invariant mass of the Z-boson for the H_T , p_{T1} and p_{T12} scans, with the energy correction applied to the decay electrons corresponding to the lower error band of fig. 4.7 (green). In figs. A.13, A.14 and A.15 are shown the histograms with the energy correction applied to the decay electrons corresponding to the upper error band of fig. 4.7 (blue).

The Y-meson invariant mass distribution is fitted in data using the sum of two Gaussian function. The H_T set of histogram of the invariant mass is shown in fig. A.16(a) only for the $20 \text{ GeV} < H_T < 40 \text{ GeV}$ due to the lack of statistics in the $0 \text{ GeV} < H_T < 20 \text{ GeV}$ energy interval. The p_{T1} set of histograms are shown in fig. A.16(b) and fig. A.16(c) and the p_{T12} set of histograms in fig. A.16(d) and fig. A.16(e) respectively.

The J/ ψ -meson invariant mass distribution cannot be fitted due to the lack of statistics.

APPENDIX A. FITS OF THE RESONANCES INVARIANT MASS DISTRIBUTION FOR THE 74X RELEASE

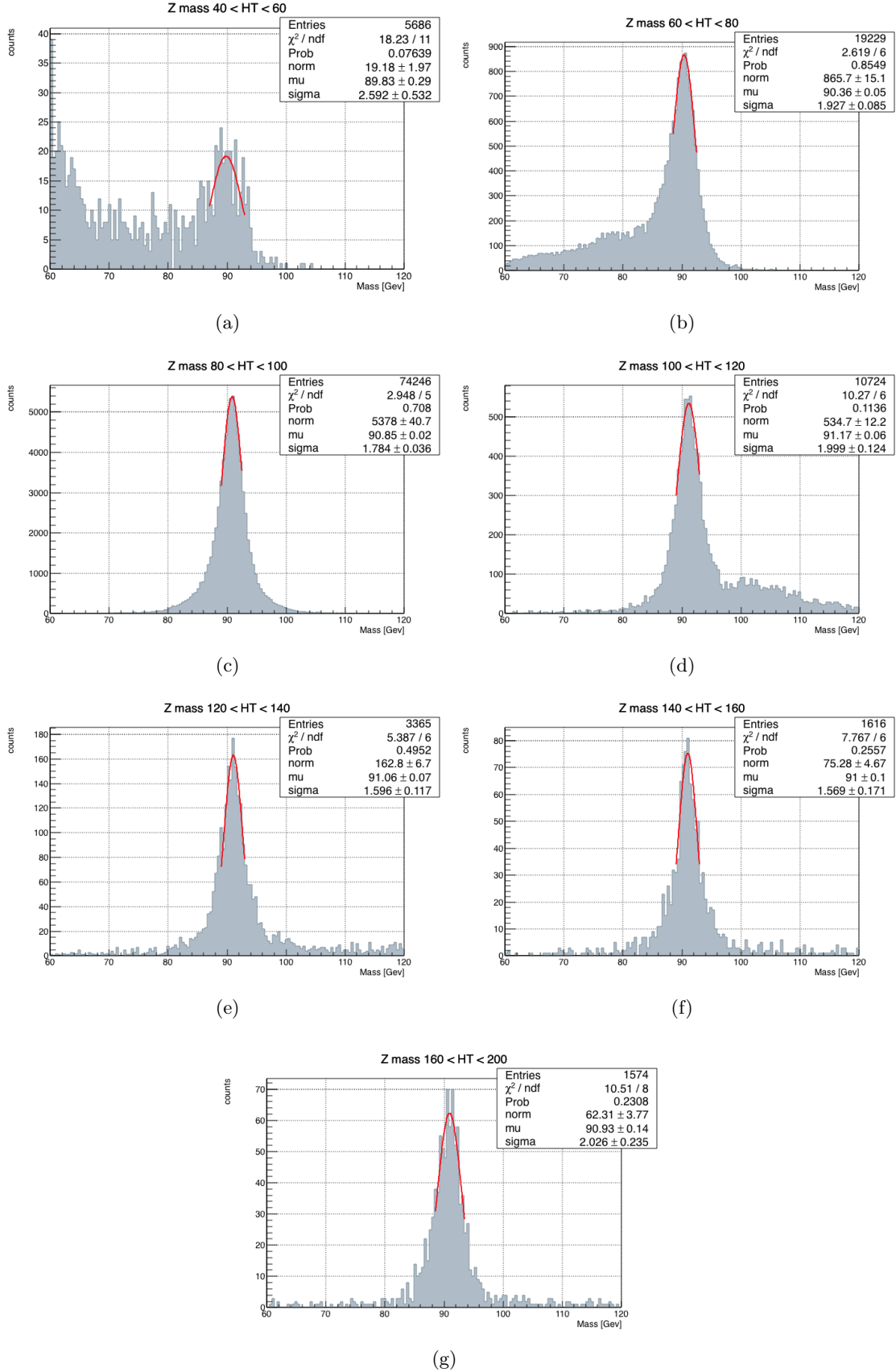


Figure A.1: Gaussian fits of the Z-boson invariant mass distribution plotted as a function of the H_T of the two electrons for the 74X release in data.

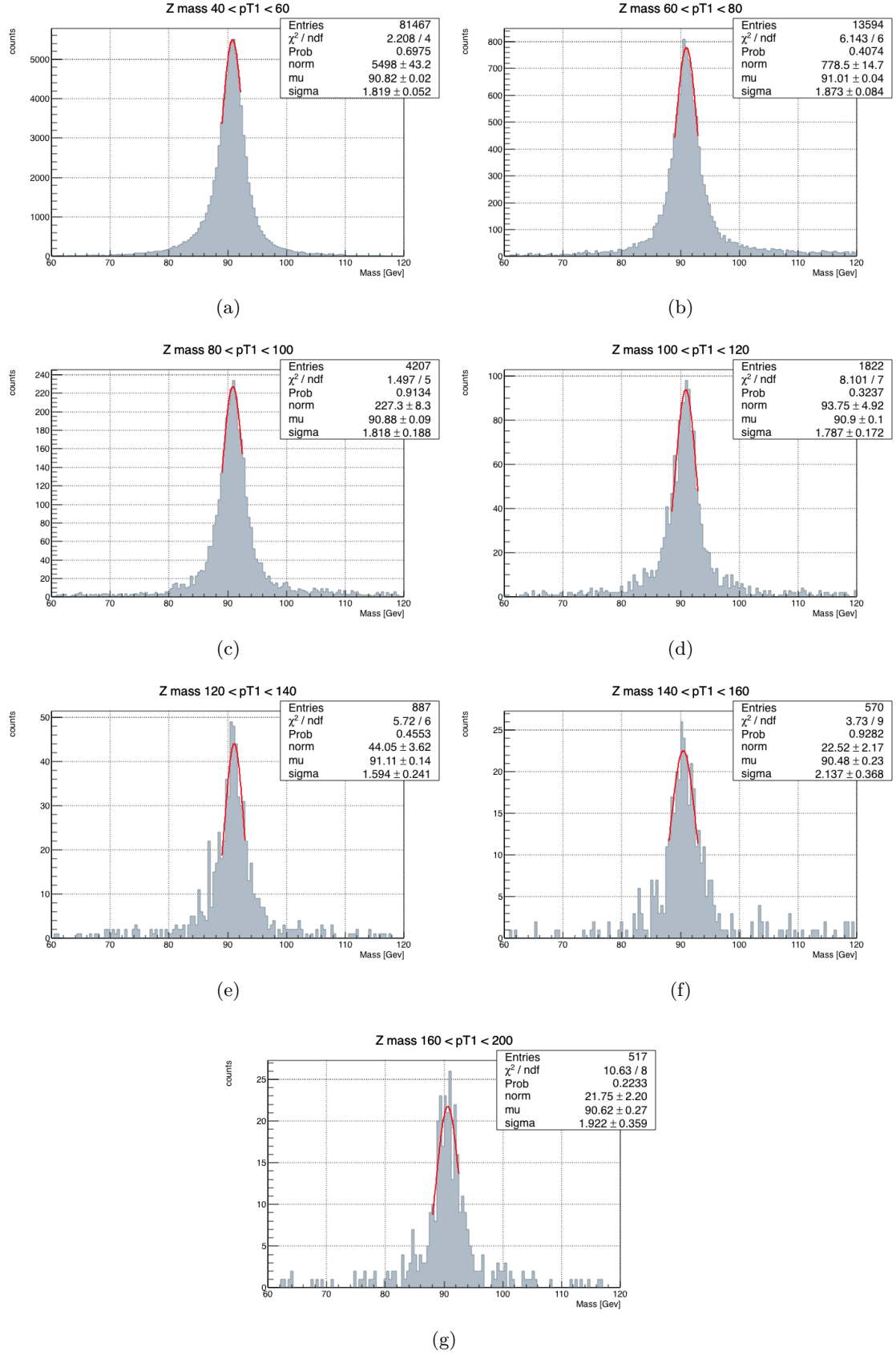


Figure A.2: Gaussian fits of the Z-boson invariant mass distribution plotted as a function of the p_T of the first electron for the 74X release in data.

APPENDIX A. FITS OF THE RESONANCES INVARIANT MASS DISTRIBUTION FOR THE 74X RELEASE

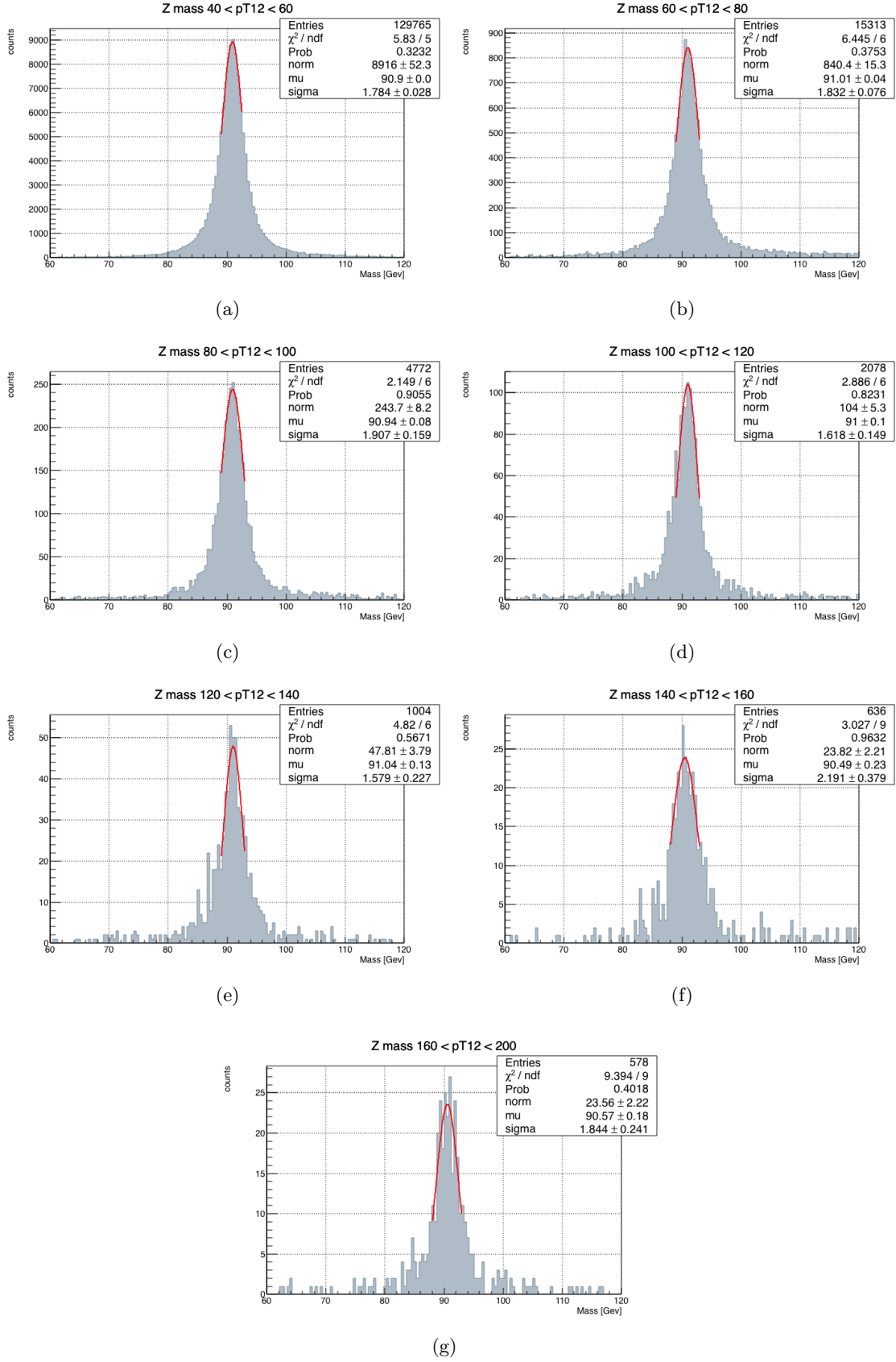


Figure A.3: Gaussian fits of the Z-boson invariant mass distribution plotted as a function of the p_T of one of the two electrons for the 74X release in data.

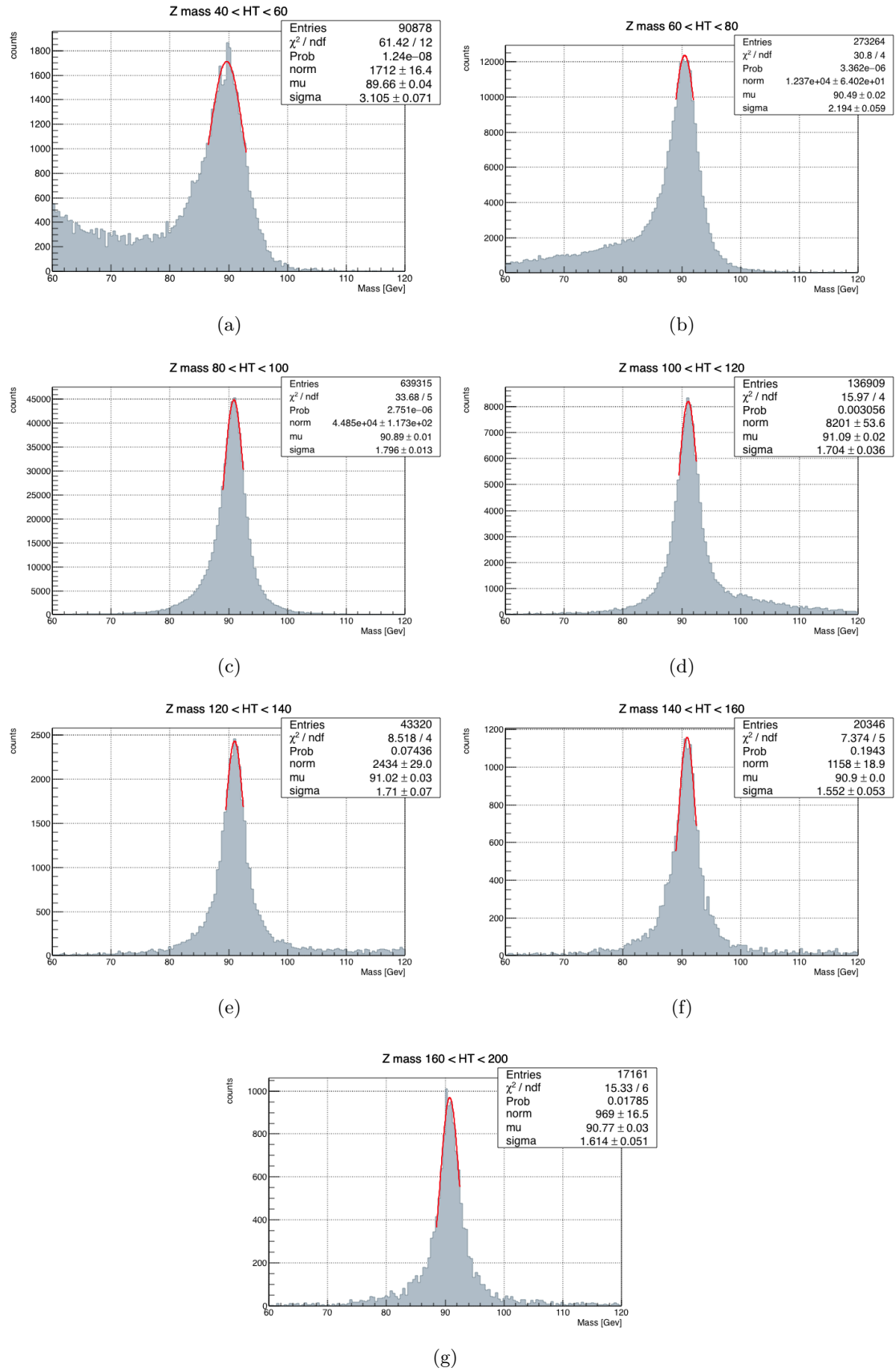


Figure A.4: Gaussian fits of the simulated Z-boson invariant mass distribution plotted as a function of the H_T of the two electrons for the 74X release in simulation.

APPENDIX A. FITS OF THE RESONANCES INVARIANT MASS DISTRIBUTION FOR THE 74X RELEASE

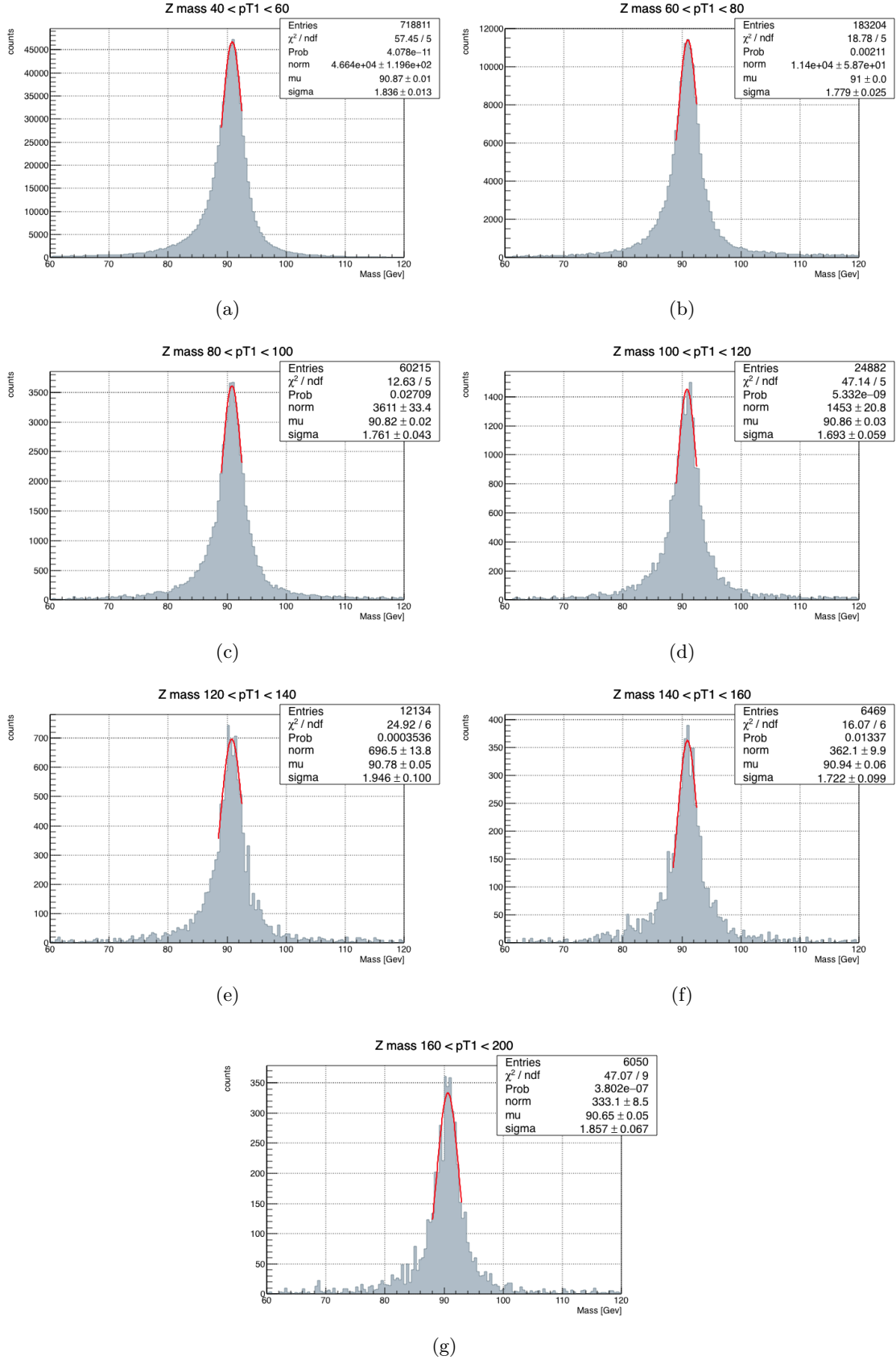
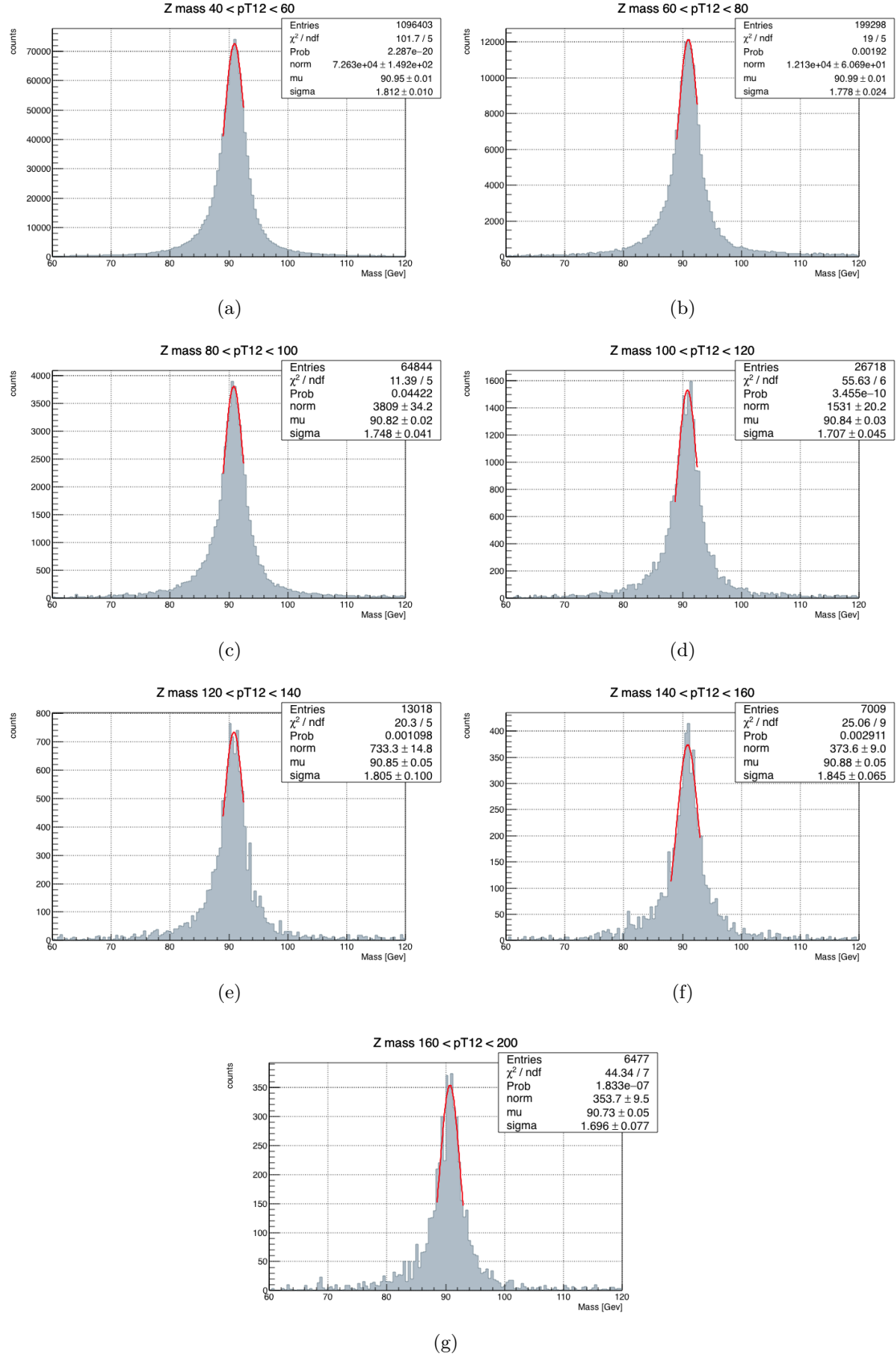


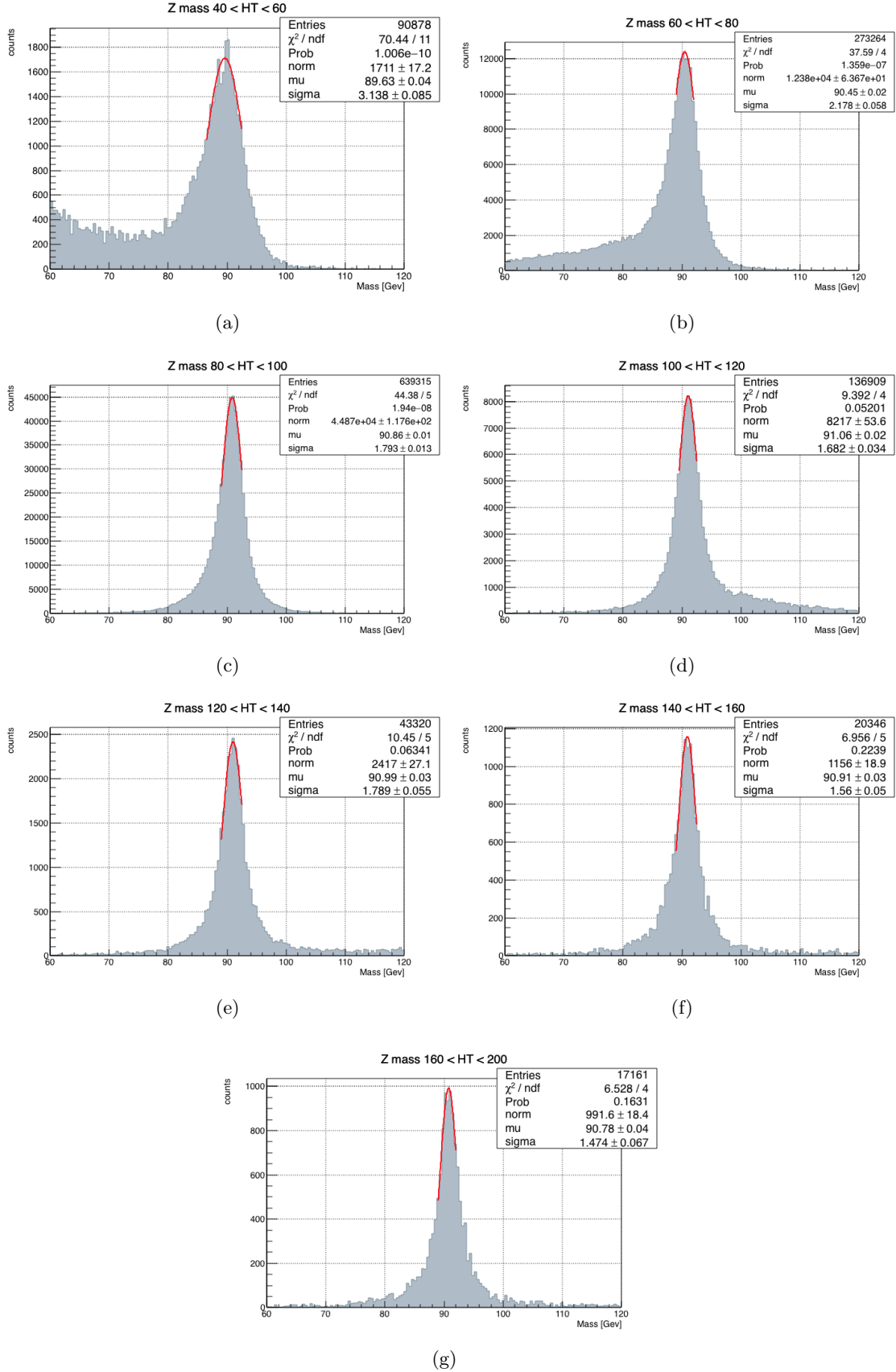
Figure A.5: Gaussian fits of the simulated Z-boson invariant mass distribution plotted as a function of the p_T of the first electron for the 74X release in simulation.



99

Figure A.6: Gaussian fits of the simulated Z-boson invariant mass distribution plotted as a function of the p_T of one of the two electrons for the 74X release in simulation.

APPENDIX A. FITS OF THE RESONANCES INVARIANT MASS DISTRIBUTION FOR THE 74X RELEASE



100

Figure A.7: Gaussian fits of the Z-boson invariant mass distribution in simulation plotted as a function of the H_T of the two electrons for the 74X release. The simulated electron energies are corrected with square root correction.

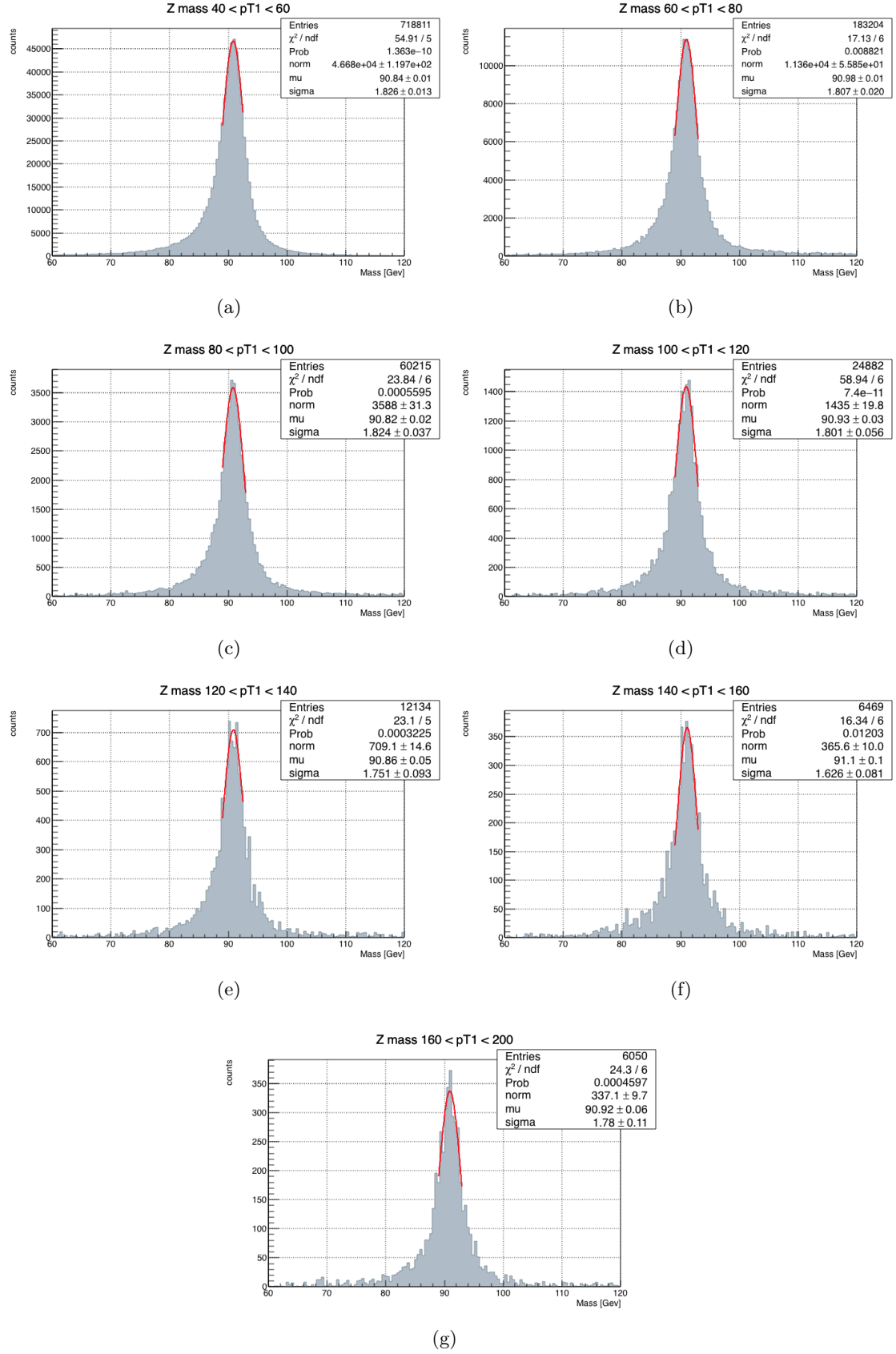


Figure A.8: Gaussian fits of the Z-boson invariant mass distribution in simulation plotted as a function of the p_T of the first electron for the 74X release. The simulated electron energies are corrected with square root correction.

APPENDIX A. FITS OF THE RESONANCES INVARIANT MASS DISTRIBUTION FOR THE 74X RELEASE

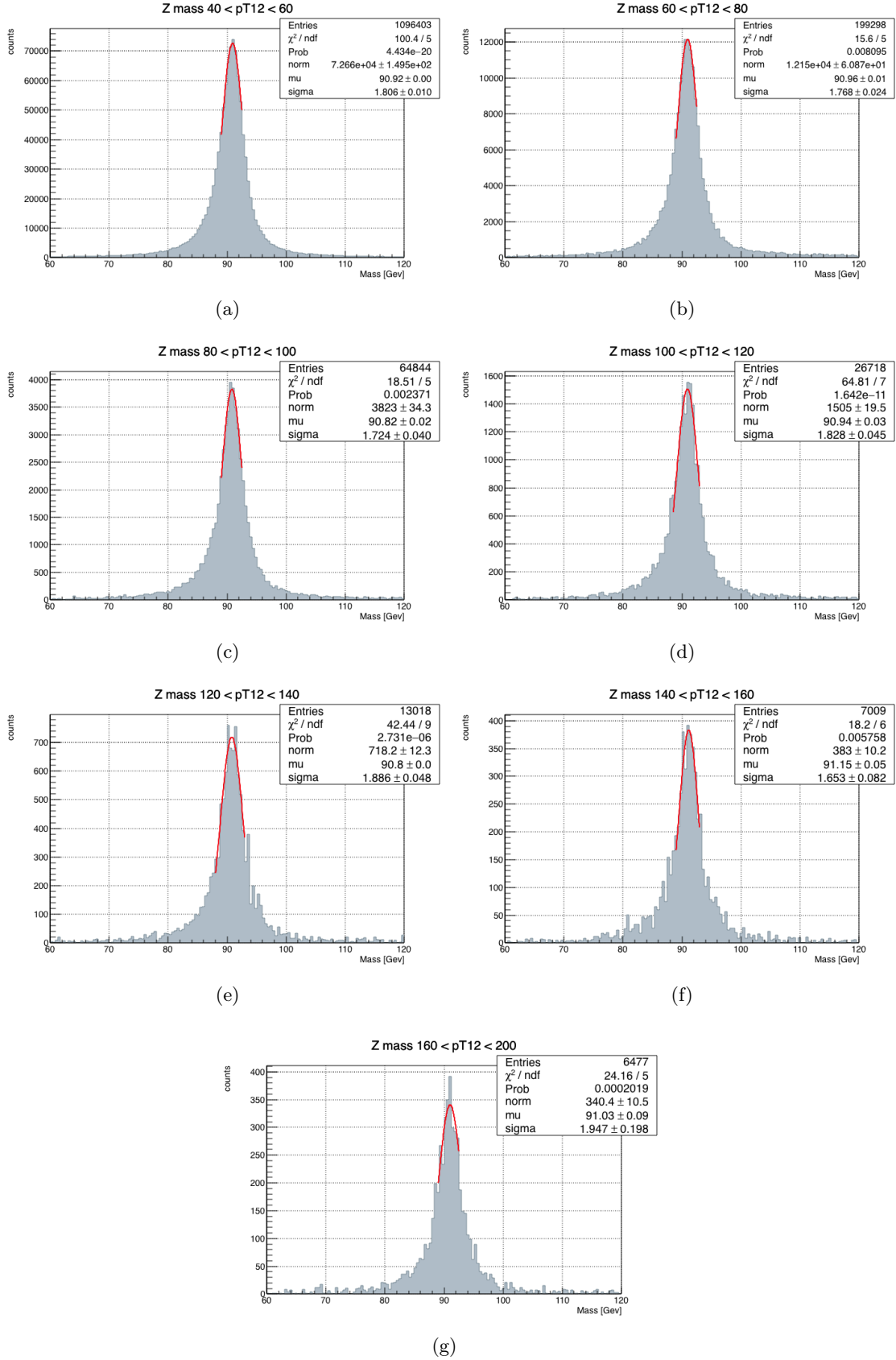


Figure A.9: Gaussian fits of the Z-boson invariant mass distribution in simulation plotted as a function of the p_T of one of the two electrons for the 74X release. The simulated electron energies are corrected with square root correction.

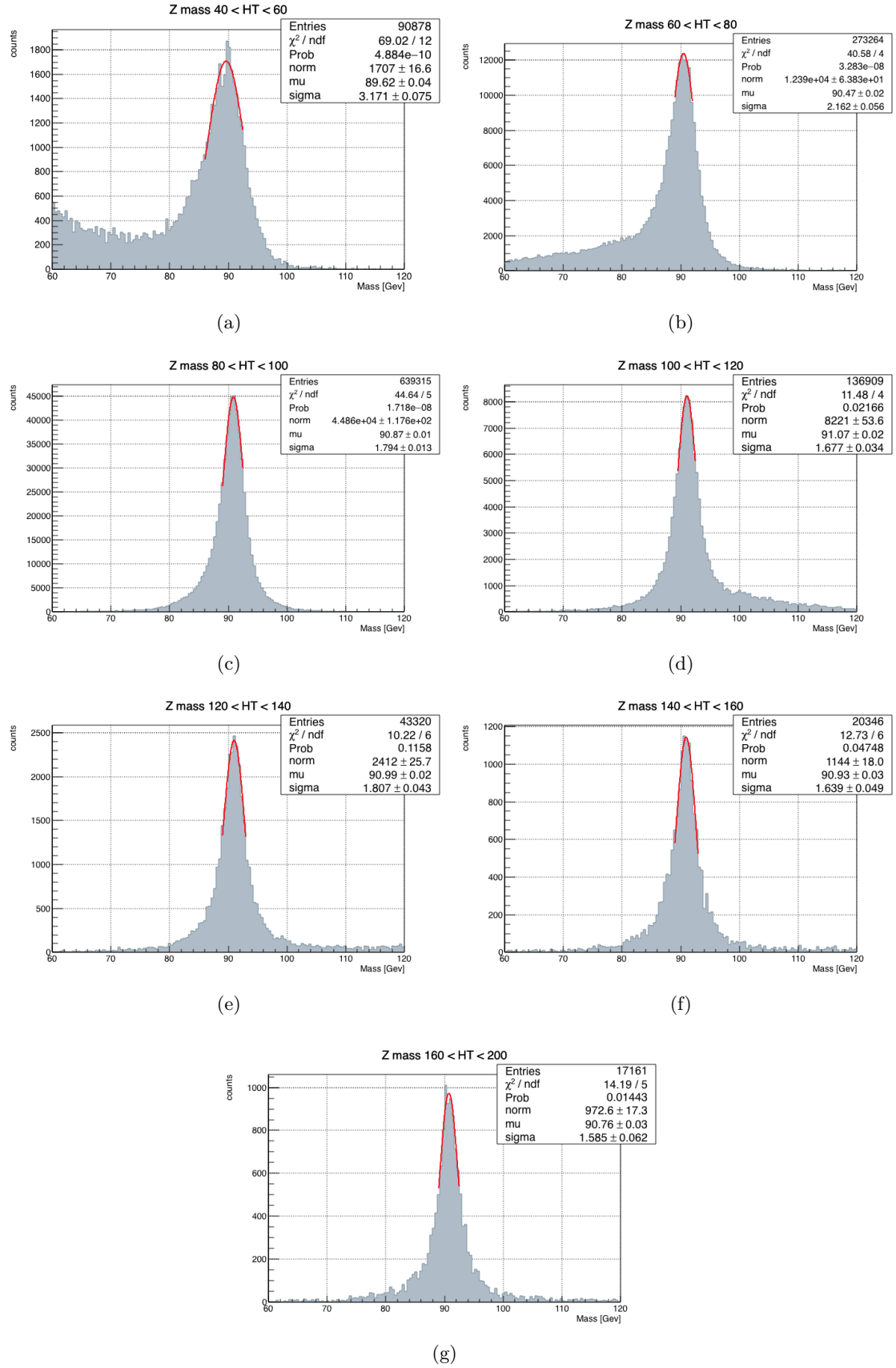


Figure A.10: Gaussian fits of the Z-boson invariant mass distribution in simulation plotted as a function of the H_T of the two electrons for the 74X release. The simulated electron energies are corrected with the $f_a(p_T)$ function extracted from fig.4.7.

APPENDIX A. FITS OF THE RESONANCES INVARIANT MASS DISTRIBUTION FOR THE 74X RELEASE

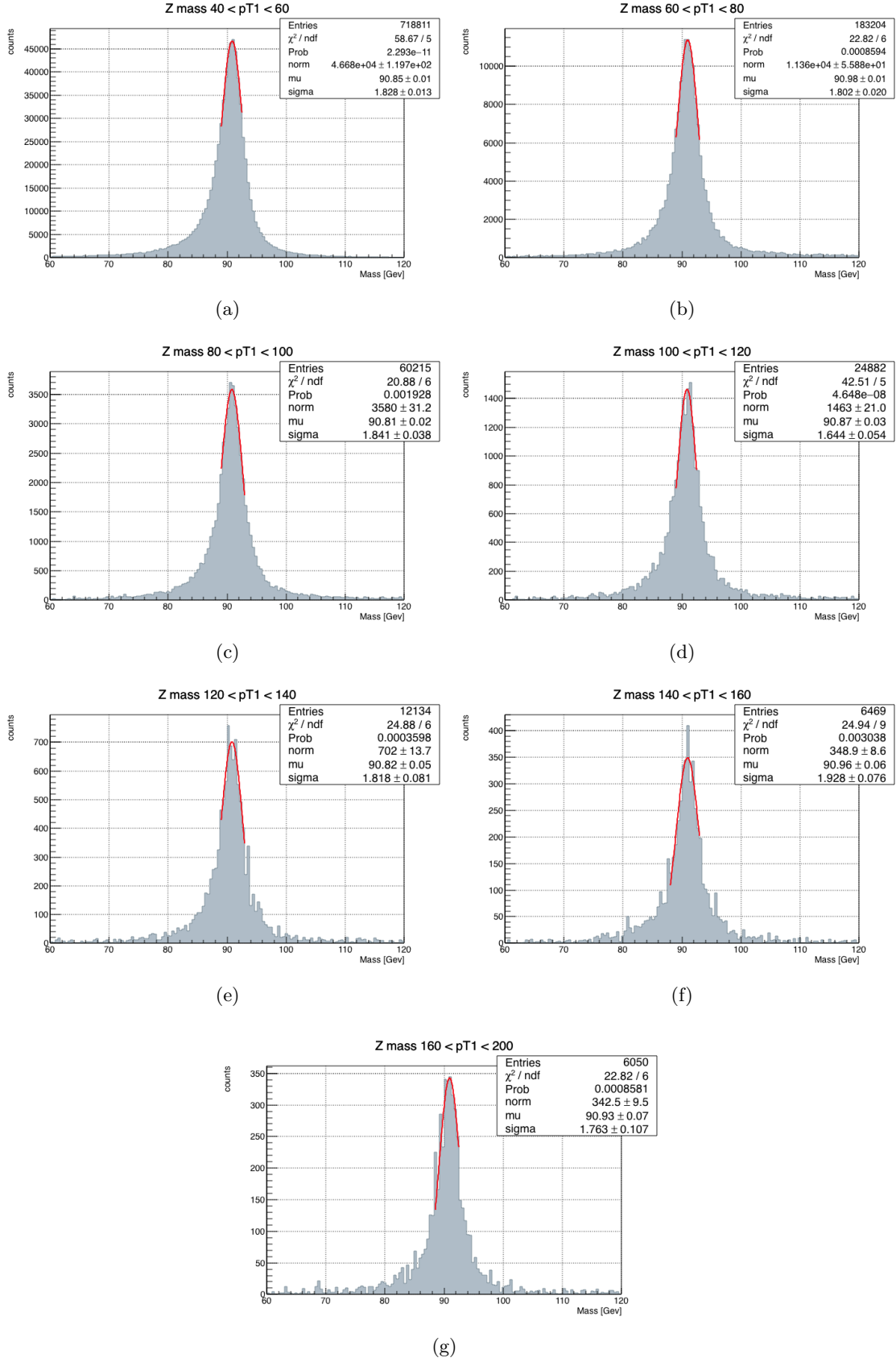


Figure A.11: Gaussian fits of the Z-boson invariant mass distribution in simulation plotted as a function of the p_T of the first electron for the 74X release. The simulated electron energies are corrected with the $f_a(p_T)$ function extracted from fig.4.7.

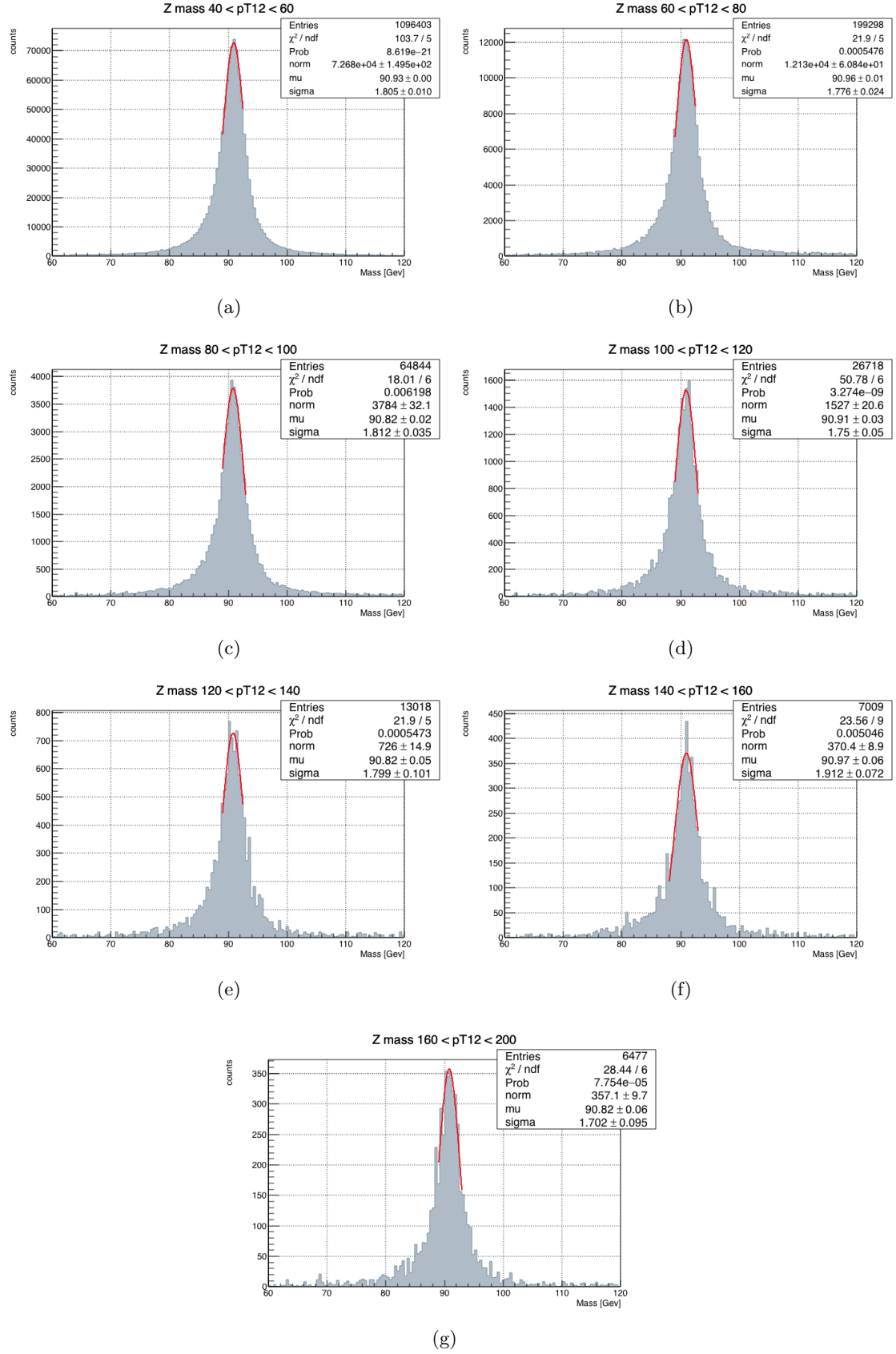


Figure A.12: Gaussian fits of the Z-boson invariant mass distribution in simulation plotted as a function of the p_T of one of the two electrons for the 74X release. The simulated electron energies are corrected with the $f_a(p_T)$ function extracted from fig.4.7.

APPENDIX A. FITS OF THE RESONANCES INVARIANT MASS DISTRIBUTION FOR THE 74X RELEASE

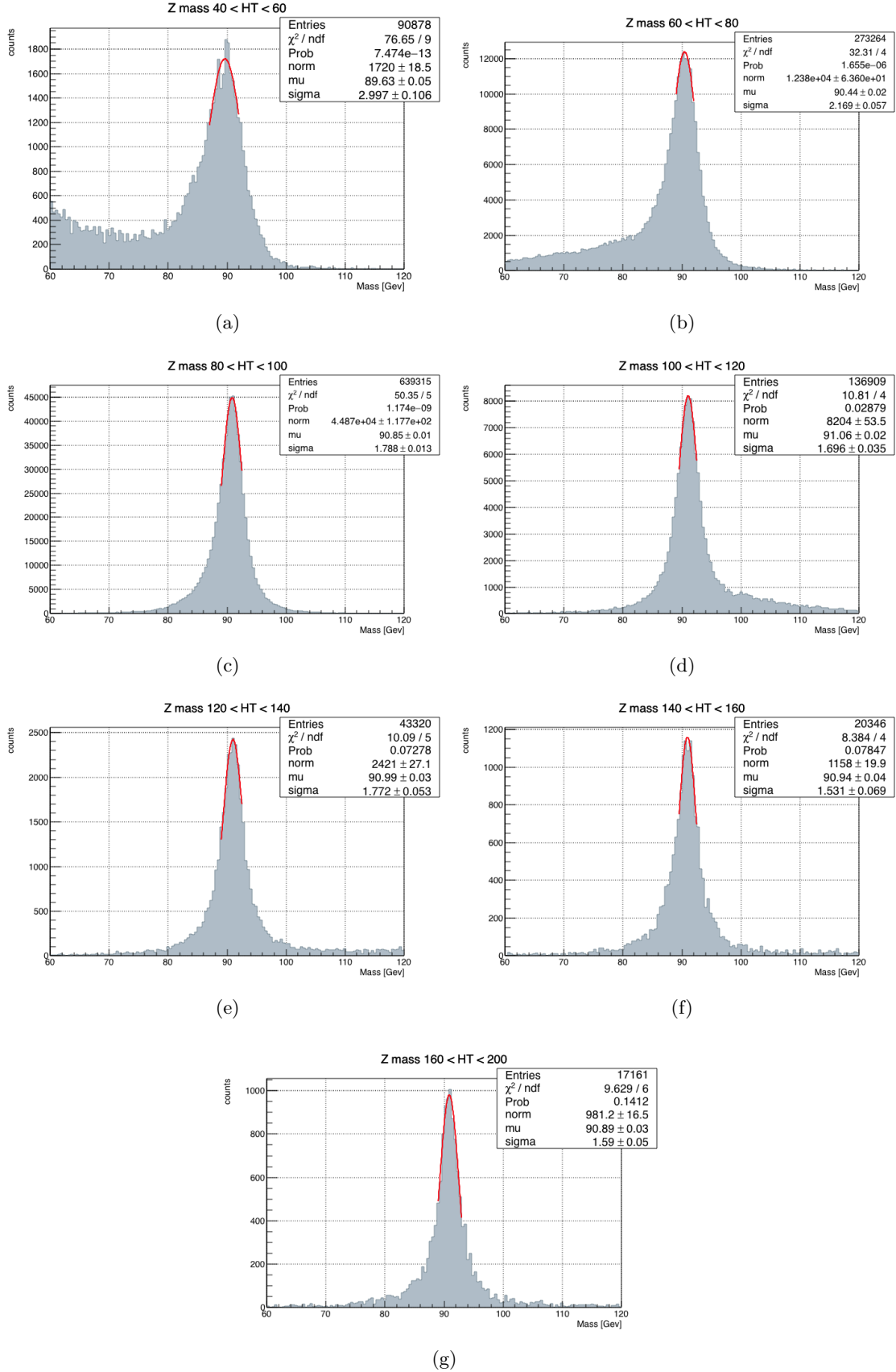


Figure A.13: Gaussian fits of the Z-boson invariant mass distribution in simulation plotted as a function of the H_T of the two electrons for the 74X release. The simulated electron energies are corrected with the $f_b(p_T)$ function extracted from fig.4.7.

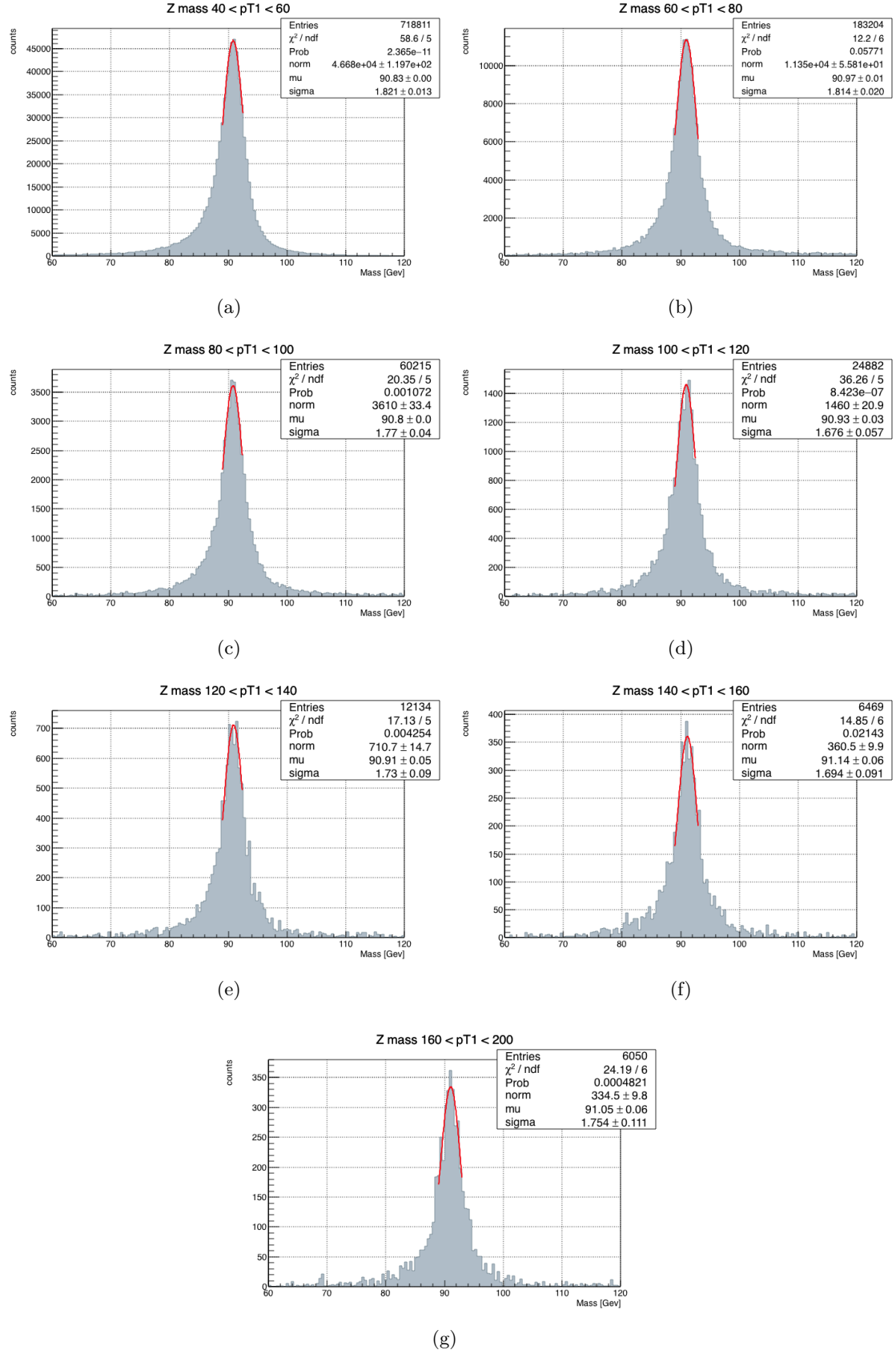


Figure A.14: Gaussian fits of the Z-boson invariant mass distribution in simulation plotted as a function of the p_T of the first electron for the 74X release. The simulated electron energies are corrected with the $f_b(p_T)$ function extracted from fig.4.7.

APPENDIX A. FITS OF THE RESONANCES INVARIANT MASS DISTRIBUTION FOR THE 74X RELEASE

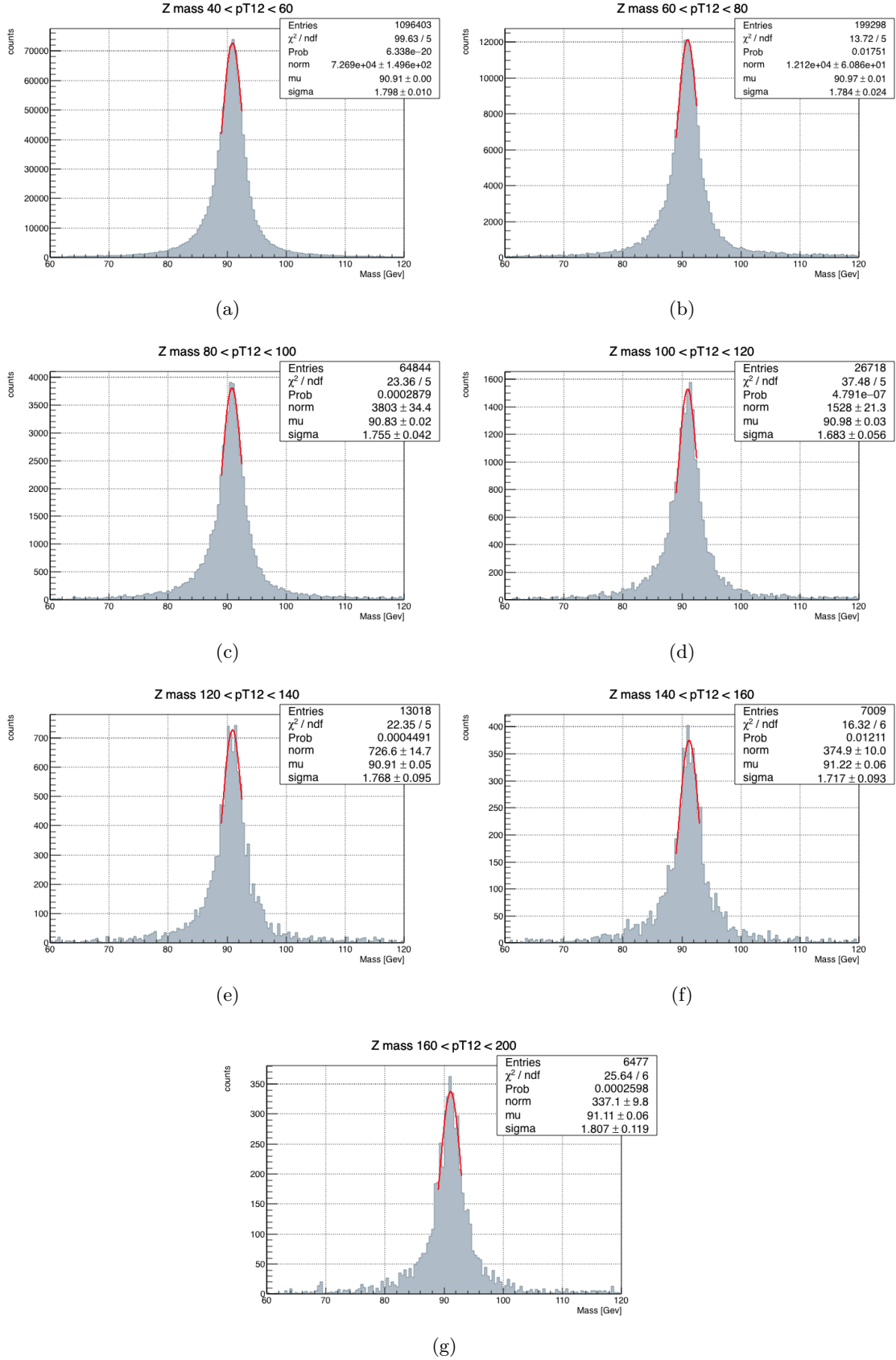


Figure A.15: Gaussian fits of the Z-boson invariant mass distribution in simulation plotted as a function of the p_T of one of the two electrons for the 74X release. The simulated electron energies are corrected with the $f_b(p_T)$ function extracted from fig.4.7.

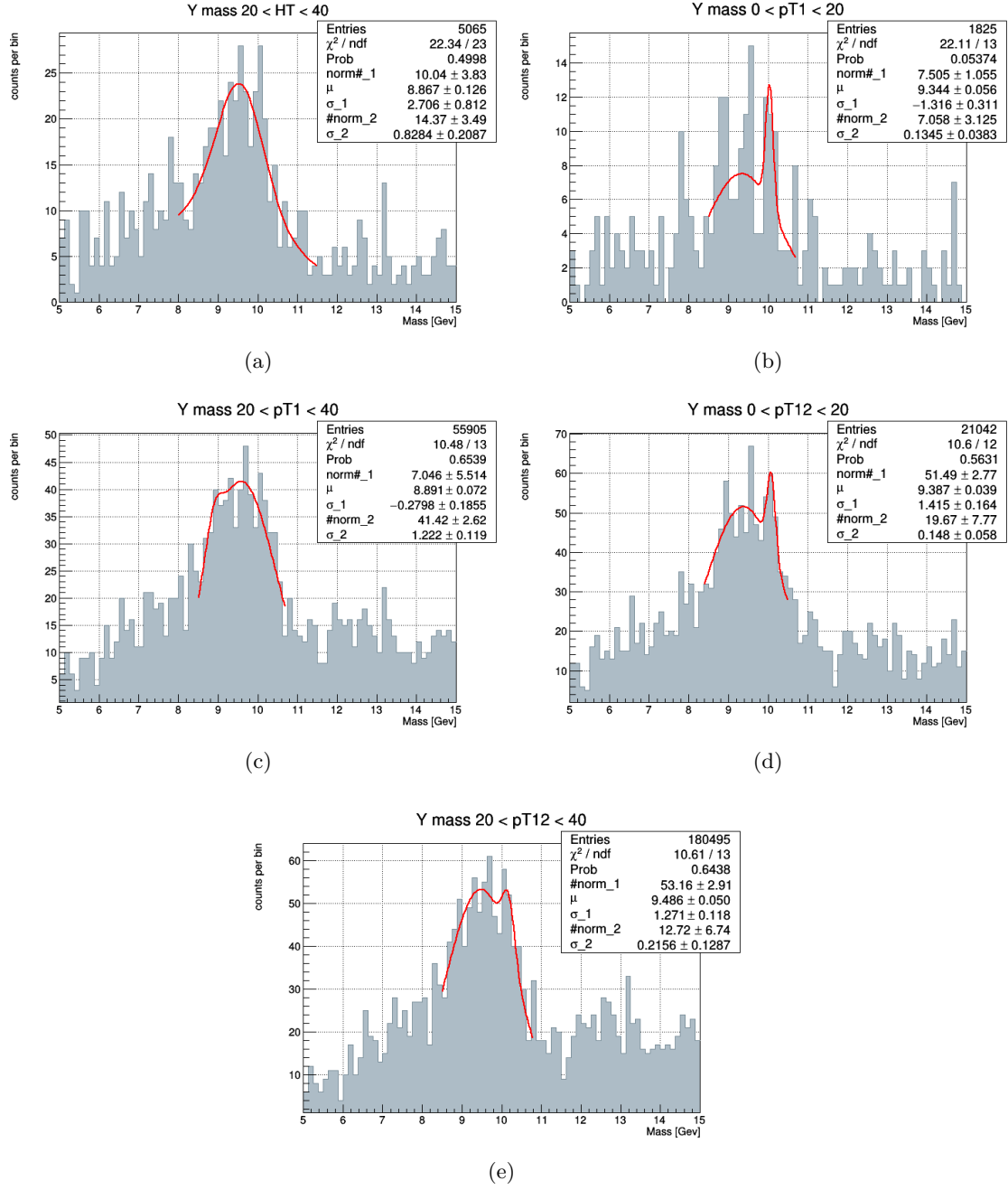


Figure A.16: Gaussian fits of the Y-meson invariant mass distribution plotted using different kinematical variables for the 74X release in data.

B | Fits of the resonances invariant mass distribution for the 76X release

The Z-boson invariant mass distribution, fitted using a Gaussian function, for the 76X release. The H_T set of histograms of the invariant mass for data and MC are shown in fig. B.1 and fig. B.4, the p_{T1} set of histograms in fig. B.2 and fig. B.5 and the p_{T12} set of histograms in fig. B.3 and fig. B.6 respectively.

The histograms in fig. B.7, B.8, and B.9 contain the invariant mass of the Z-boson for the H_T , p_{T1} and p_{T12} scans, with the energy correction applied to the decay electrons.

The histograms in fig. B.10, B.11 and B.12 contain the invariant mass of the Z-boson for the H_T , p_{T1} and p_{T12} scans, with the energy correction applied to the decay electrons corresponding to the lower error band of fig. 4.10 (green). In figs. B.13, B.14 and B.15 are shown the histograms with the energy correction applied to the decay electrons corresponding to the upper error band of fig. 4.10 (blue).

The Y-meson invariant mass distribution is fitted in data using the sum of two Gaussian function. The H_T set of histograms of the invariant mass is shown in fig. A.16(a) only for the $20 \text{ GeV} < H_T < 40 \text{ GeV}$ due to the lack of statistics in the $0 \text{ GeV} < H_T < 20 \text{ GeV}$ energy interval. The p_{T1} set of histograms is shown only in and fig. A.16(c) for lack of statistics in the $20 \text{ GeV} < p_{T1} < 40 \text{ GeV}$ region and the p_{T12} set of histograms are shown in fig. A.16(d) and fig. A.16(e).

The J/ψ -meson invariant mass distribution cannot be fitted due to the lack of statistics.

APPENDIX B. FITS OF THE RESONANCES INVARIANT MASS DISTRIBUTION FOR THE 76X RELEASE

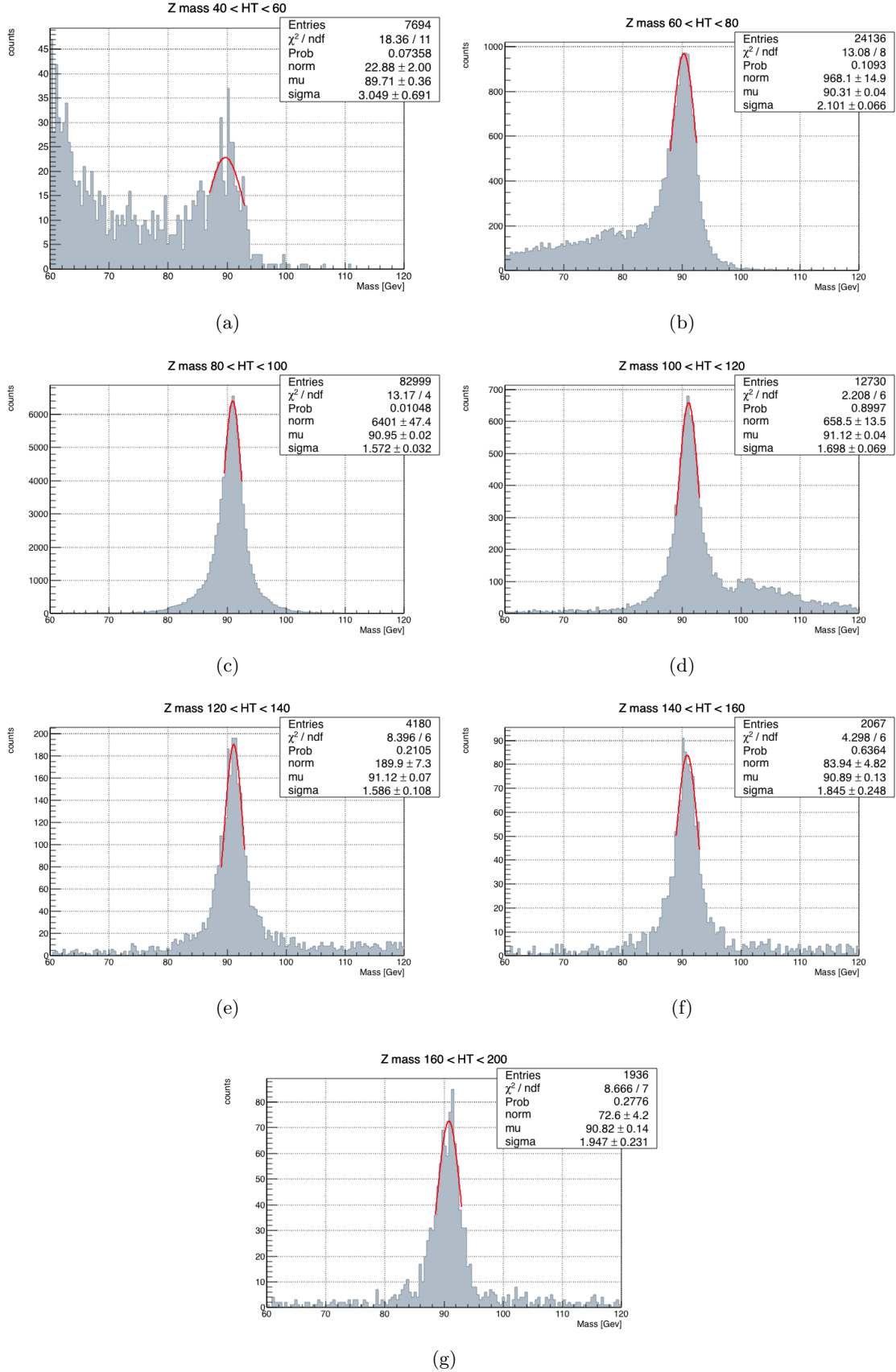


Figure B.1: Gaussian fits of the Z-boson invariant mass distribution plotted as a function of the H_T of the two electrons for the 76X release in data.

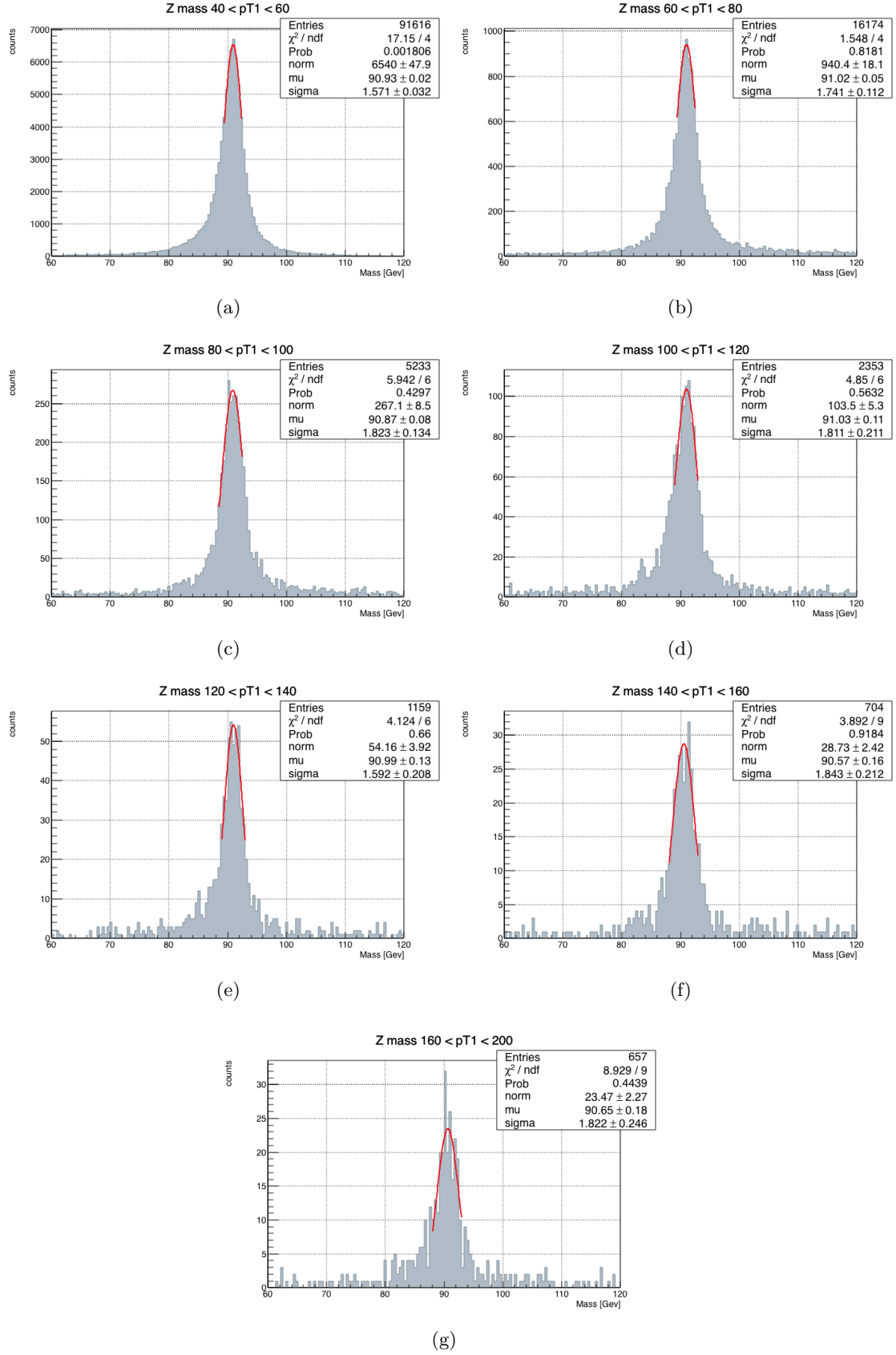


Figure B.2: Gaussian fits of the Z-boson invariant mass distribution plotted as a function of the p_T of the first electron for the 76X release in data.

APPENDIX B. FITS OF THE RESONANCES INVARIANT MASS DISTRIBUTION FOR THE 76X RELEASE

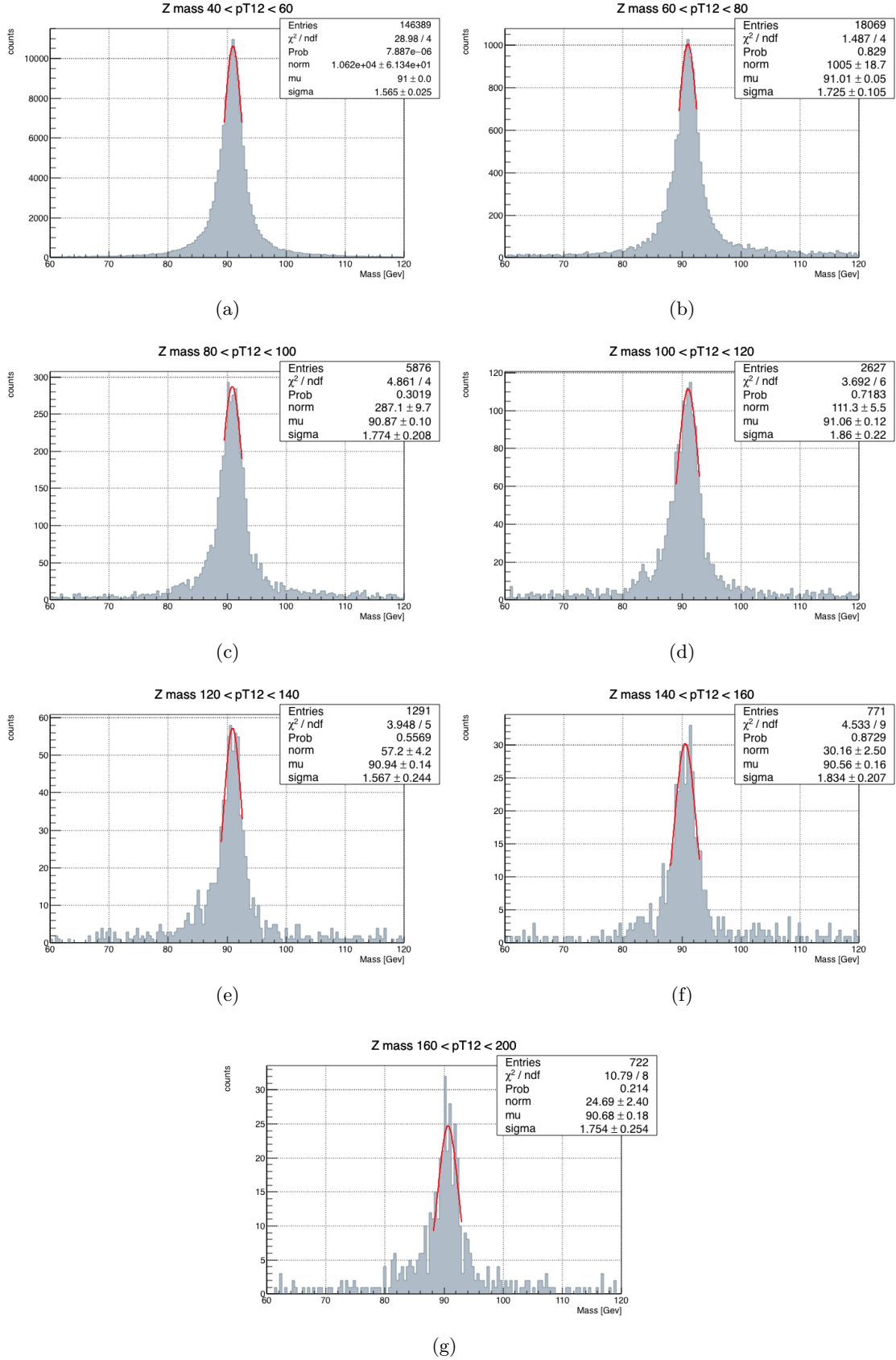
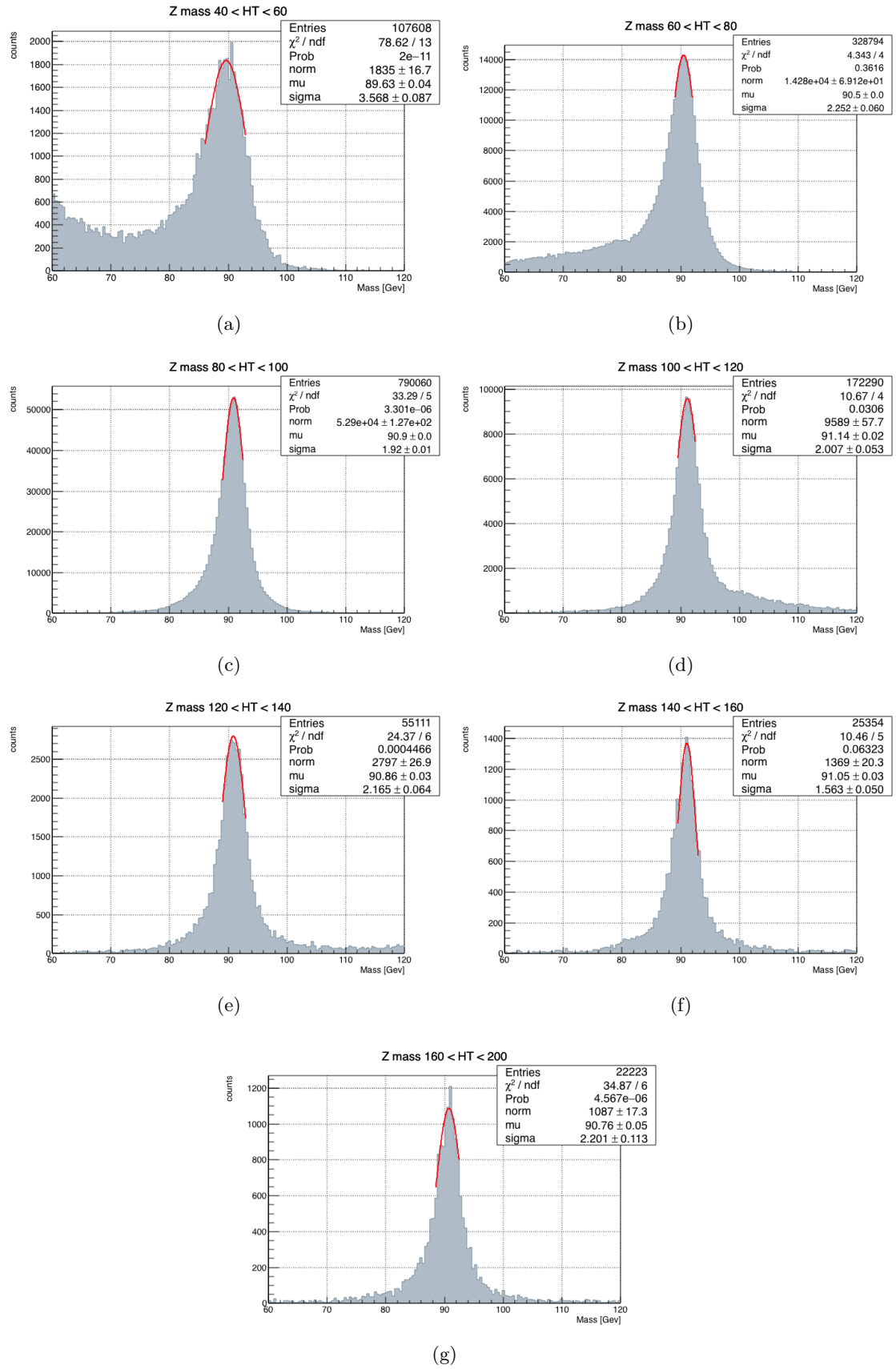


Figure B.3: Gaussian fits of the Z-boson invariant mass distribution plotted as a function of the p_T of one of the two electrons for the 76X release in data.



115

Figure B.4: Gaussian fits of the simulated Z-boson invariant mass distribution plotted as a function of the H_T of the two electrons for the 76X release in simulation.

APPENDIX B. FITS OF THE RESONANCES INVARIANT MASS DISTRIBUTION FOR THE 76X RELEASE

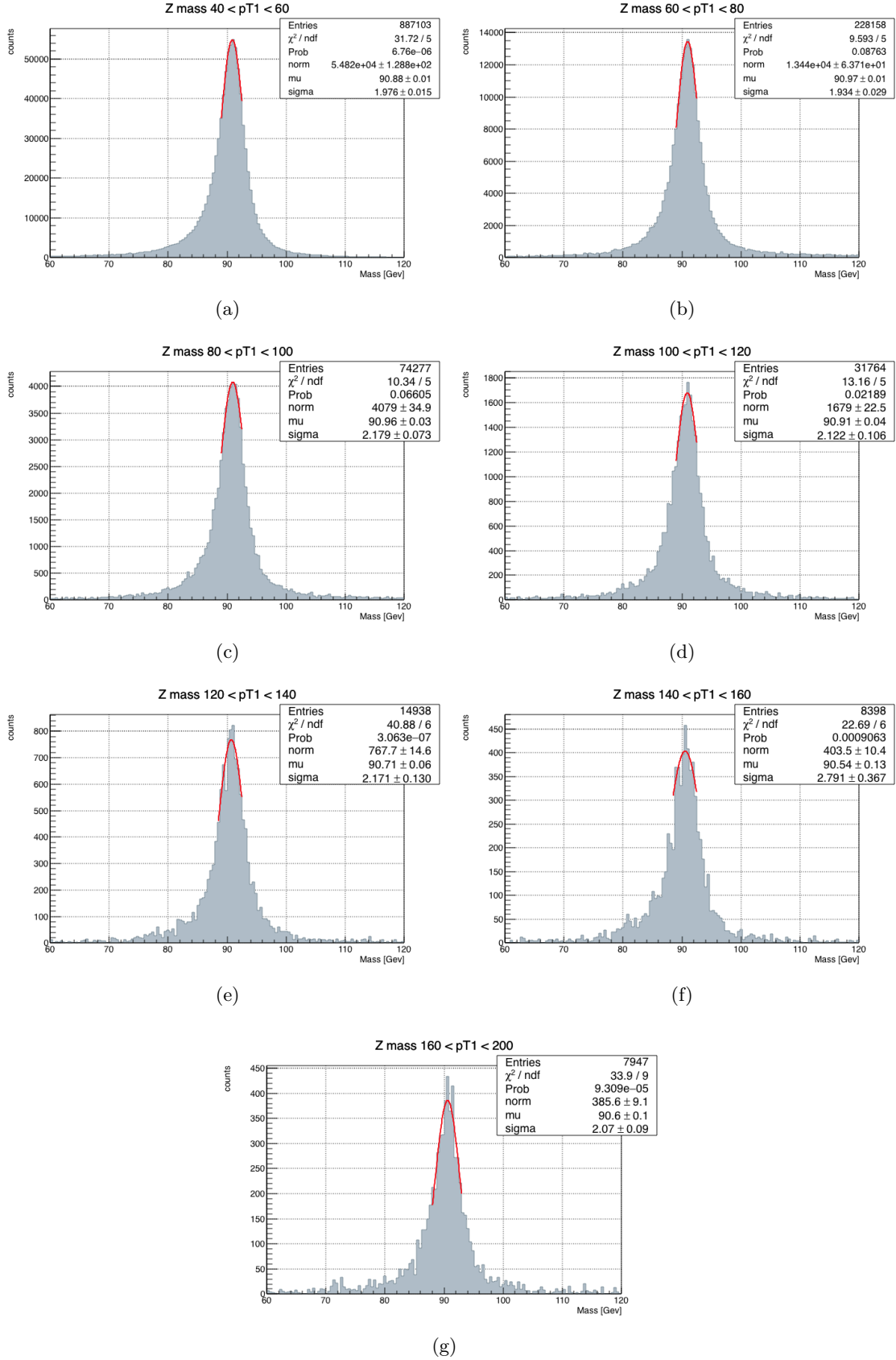


Figure B.5: Gaussian fits of the simulated Z-boson invariant mass distribution plotted as a function of the p_T of the first electron for the 76X release in simulation.

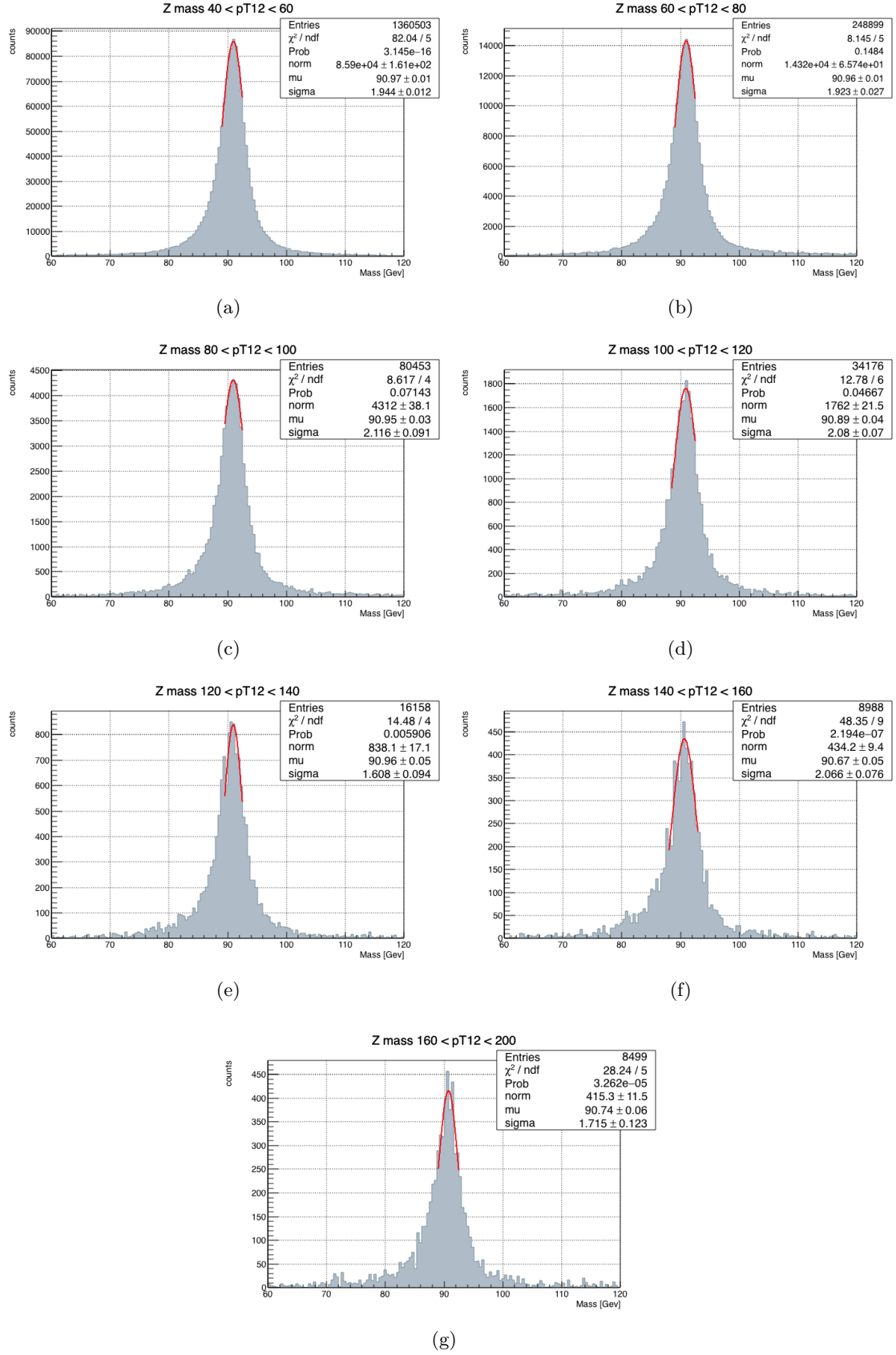


Figure B.6: Gaussian fits of the simulated Z-boson invariant mass distribution plotted as a function of the p_T of one of the two electrons for the 76X release in simulation.

APPENDIX B. FITS OF THE RESONANCES INVARIANT MASS DISTRIBUTION FOR THE 76X RELEASE

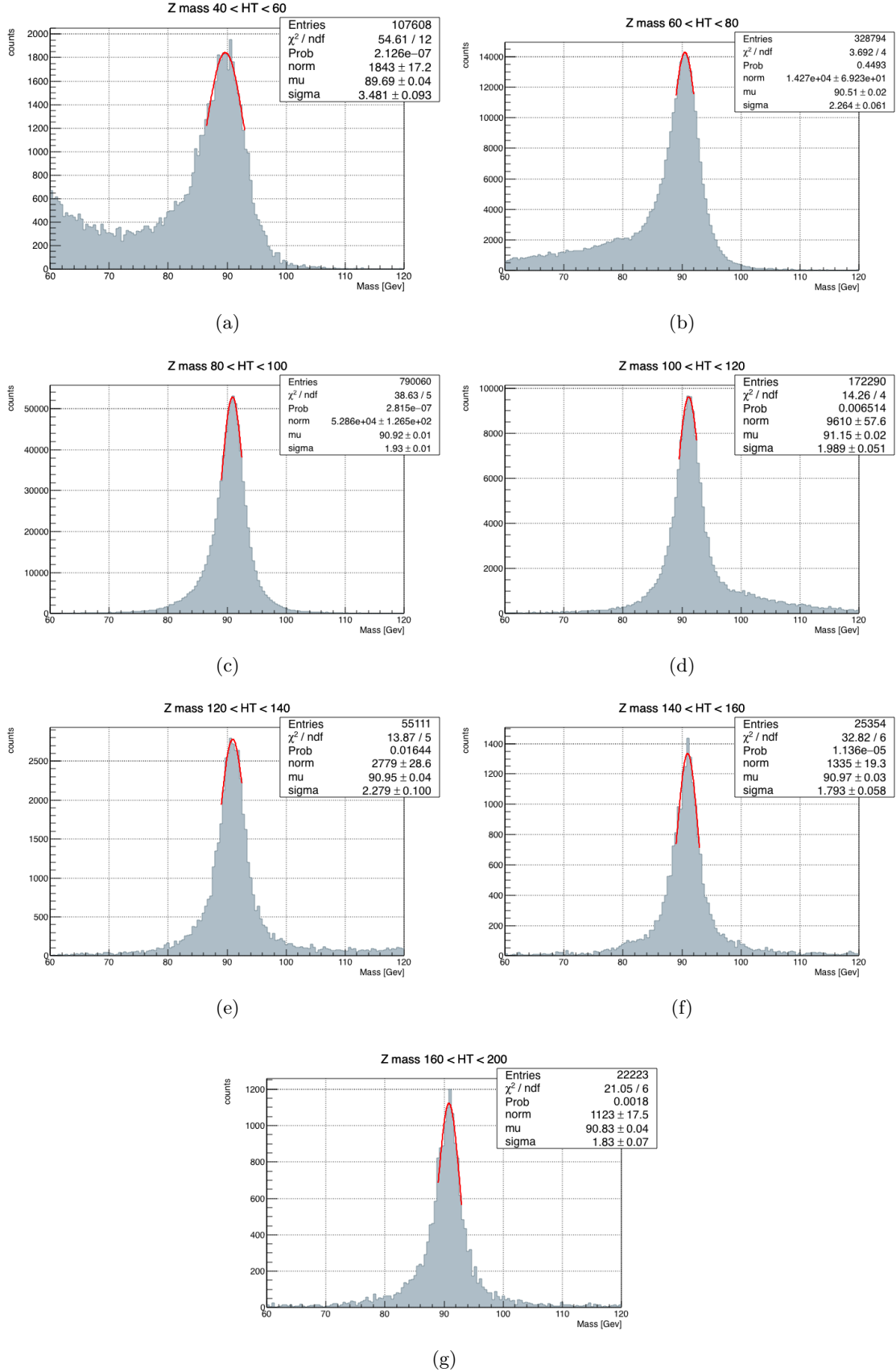


Figure B.7: Gaussian fits of the Z-boson invariant mass distribution in simulation plotted as a function of the H_T of the two electrons for the 76X release. The simulated electron energies are corrected with square root correction.

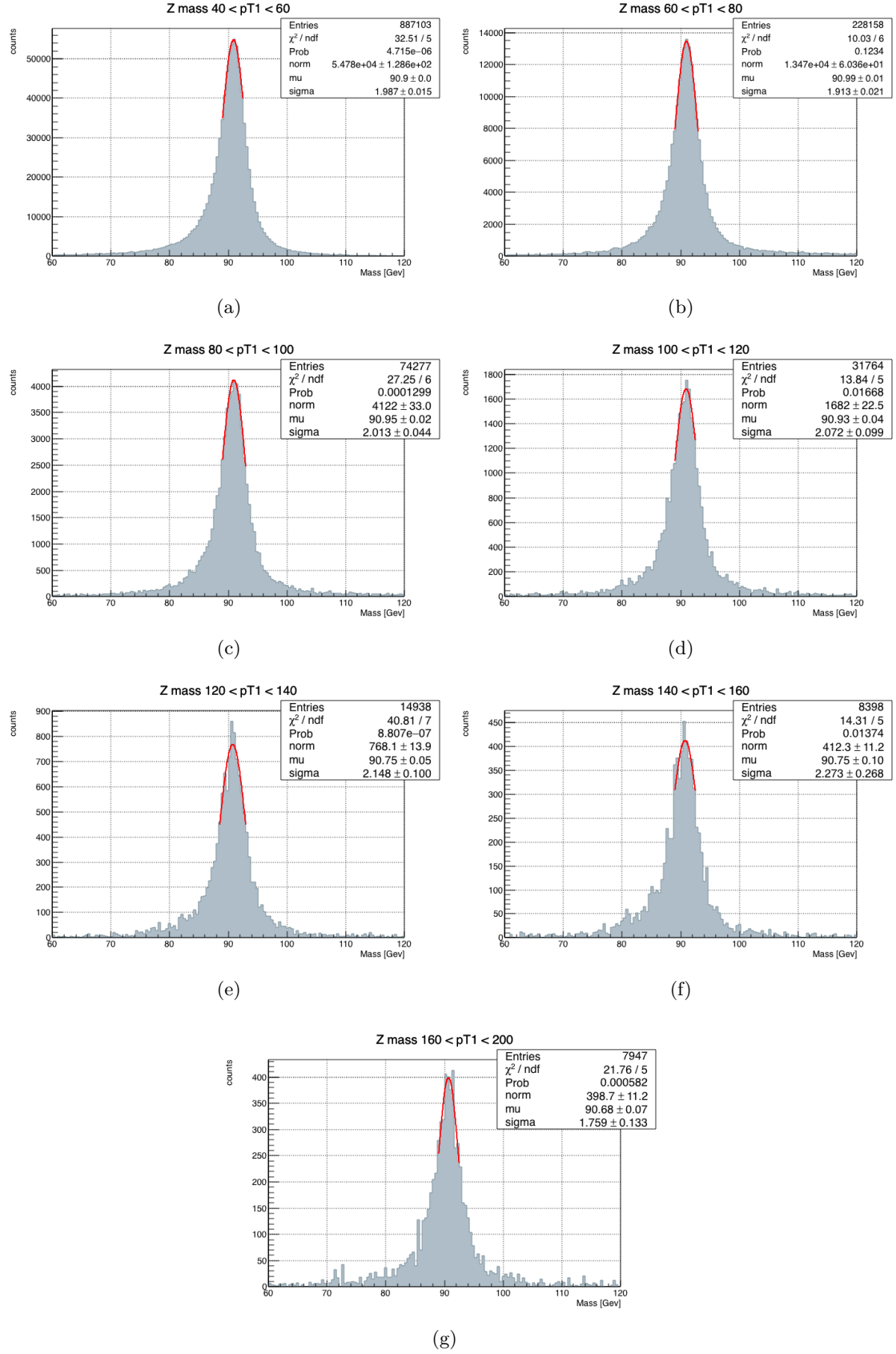
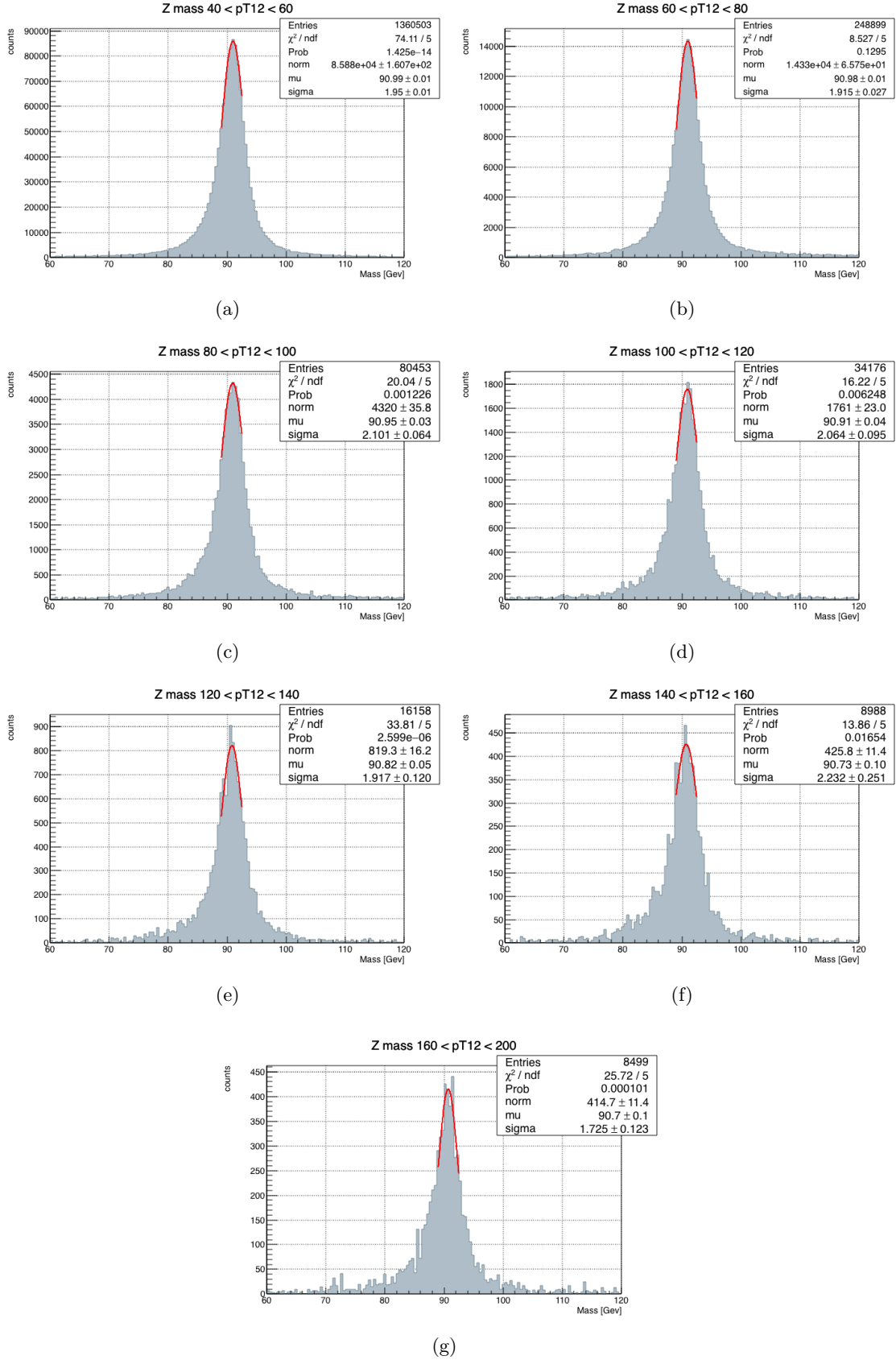


Figure B.8: Gaussian fits of the Z-boson invariant mass distribution in simulation plotted as a function of the p_T of the first electron for the 76X release. The simulated electron energies are corrected with square root correction.

APPENDIX B. FITS OF THE RESONANCES INVARIANT MASS DISTRIBUTION FOR THE 76X RELEASE



120

Figure B.9: Gaussian fits of the Z-boson invariant mass distribution in simulation plotted as a function of the p_T of one of the two electrons for the 76X release. The simulated electron energies are corrected with square root correction.

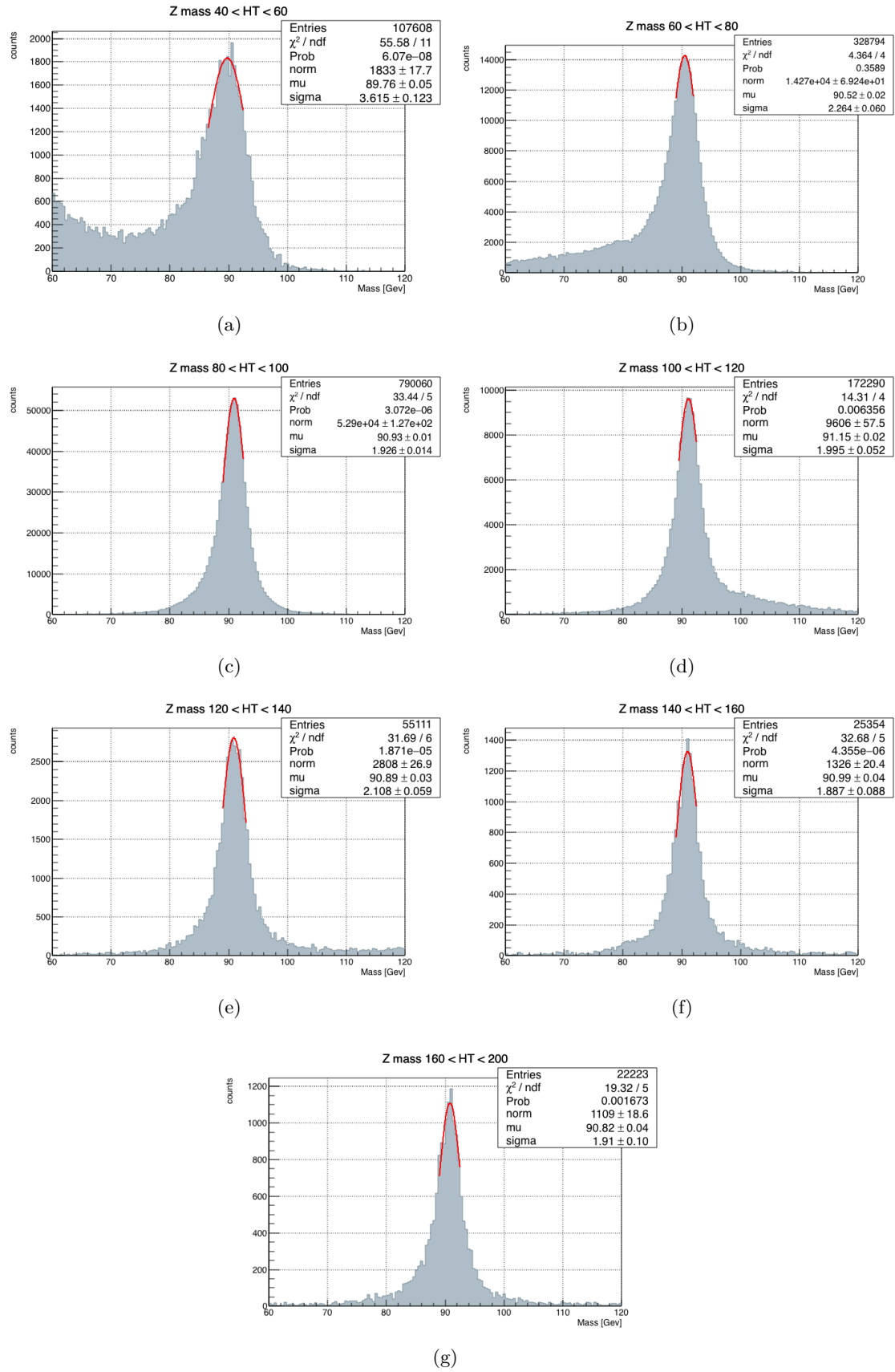


Figure B.10: Gaussian fits of the Z-boson invariant mass distribution in simulation plotted as a function of the H_T of the two electrons for the 76X release. The simulated electron energies are corrected with the $f_a(p_T)$ function extracted from fig.4.10.

APPENDIX B. FITS OF THE RESONANCES INVARIANT MASS DISTRIBUTION FOR THE 76X RELEASE

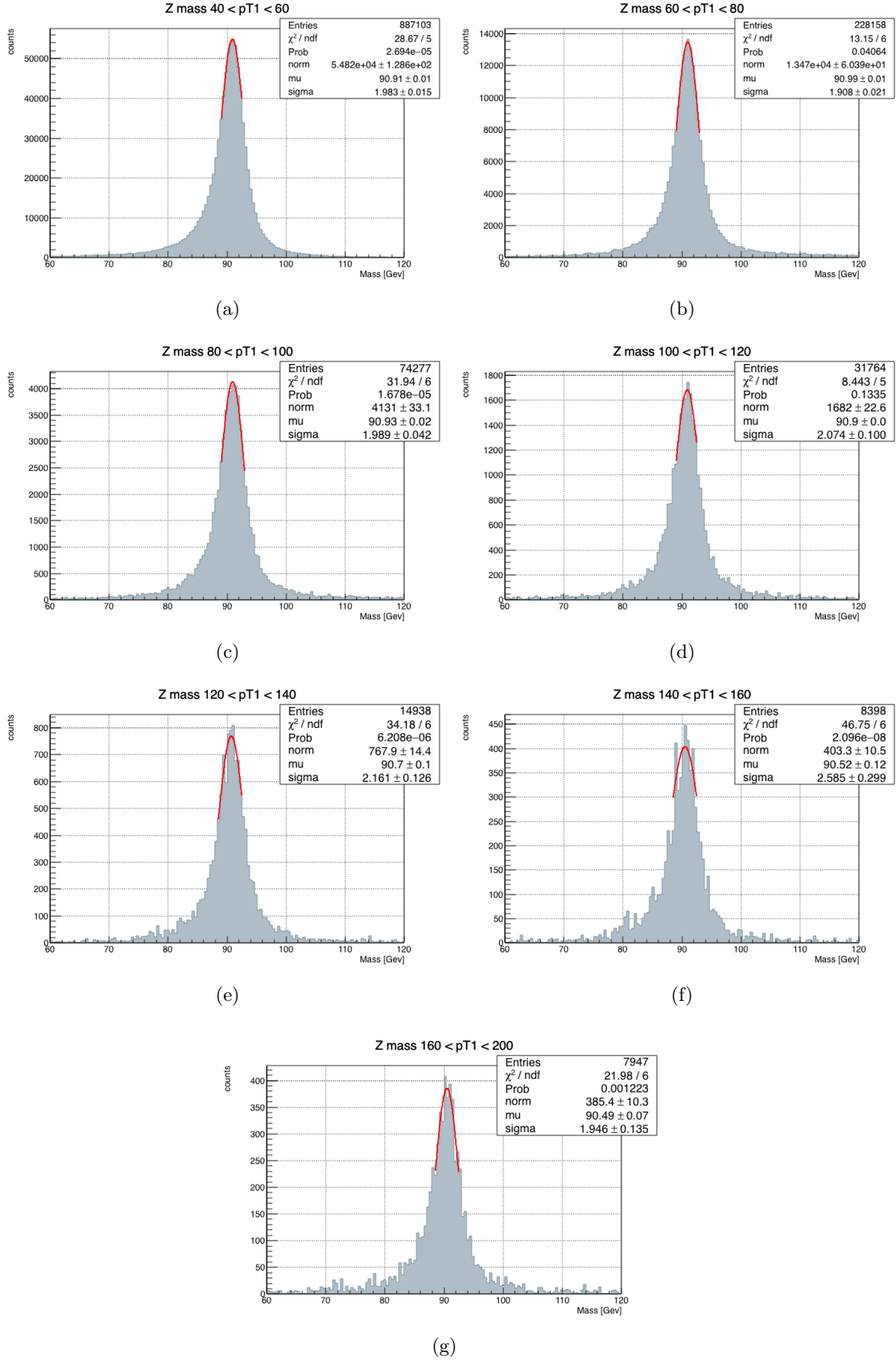


Figure B.11: Gaussian fits of the Z-boson invariant mass distribution in simulation plotted as a function of the p_T of the first electron for the 76X release. The simulated electron energies are corrected with the $f_a(p_T)$ function extracted from fig.4.10.

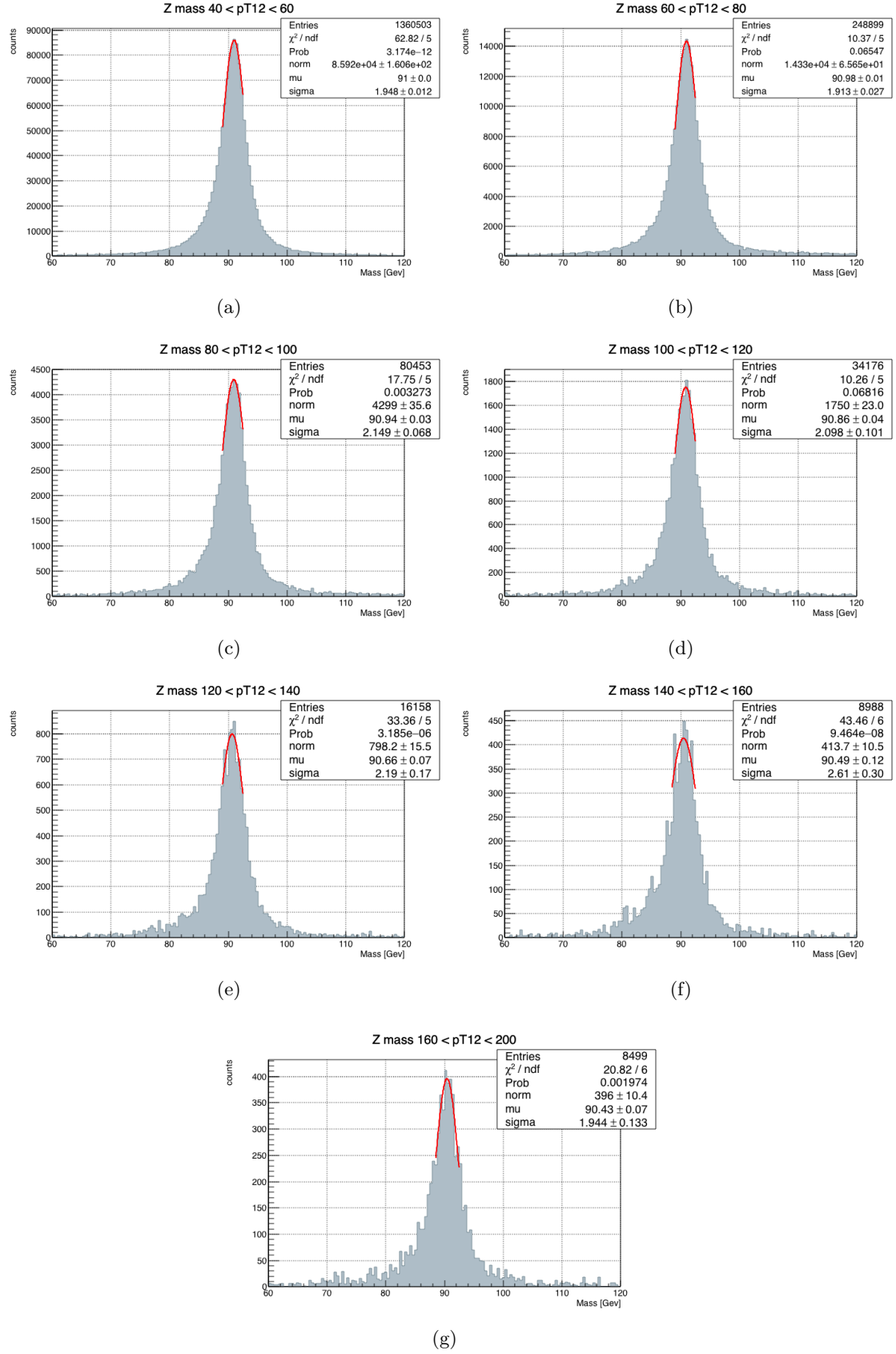


Figure B.12: Gaussian fits of the Z-boson invariant mass distribution in simulation plotted as a function of the p_T of one of the two electrons for the 76X release. The simulated electron energies are corrected with the $f_a(p_T)$ function extracted from fig.4.10.

APPENDIX B. FITS OF THE RESONANCES INVARIANT MASS DISTRIBUTION FOR THE 76X RELEASE

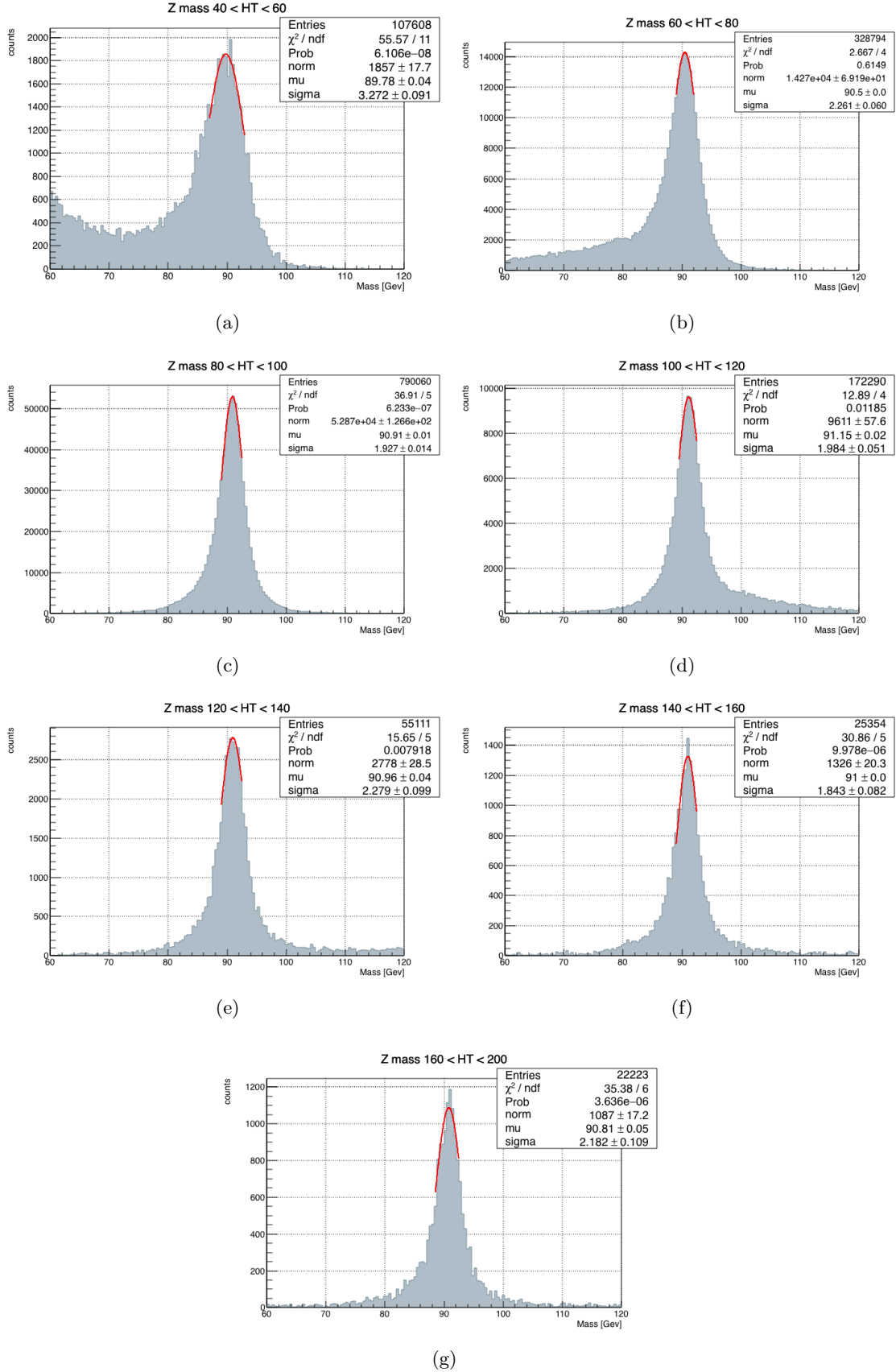


Figure B.13: Gaussian fits of the Z-boson invariant mass distribution in simulation plotted as a function of the H_T of the two electrons for the 76X release. The simulated electron energies are corrected with the $f_b(p_T)$ function extracted from fig.4.10.

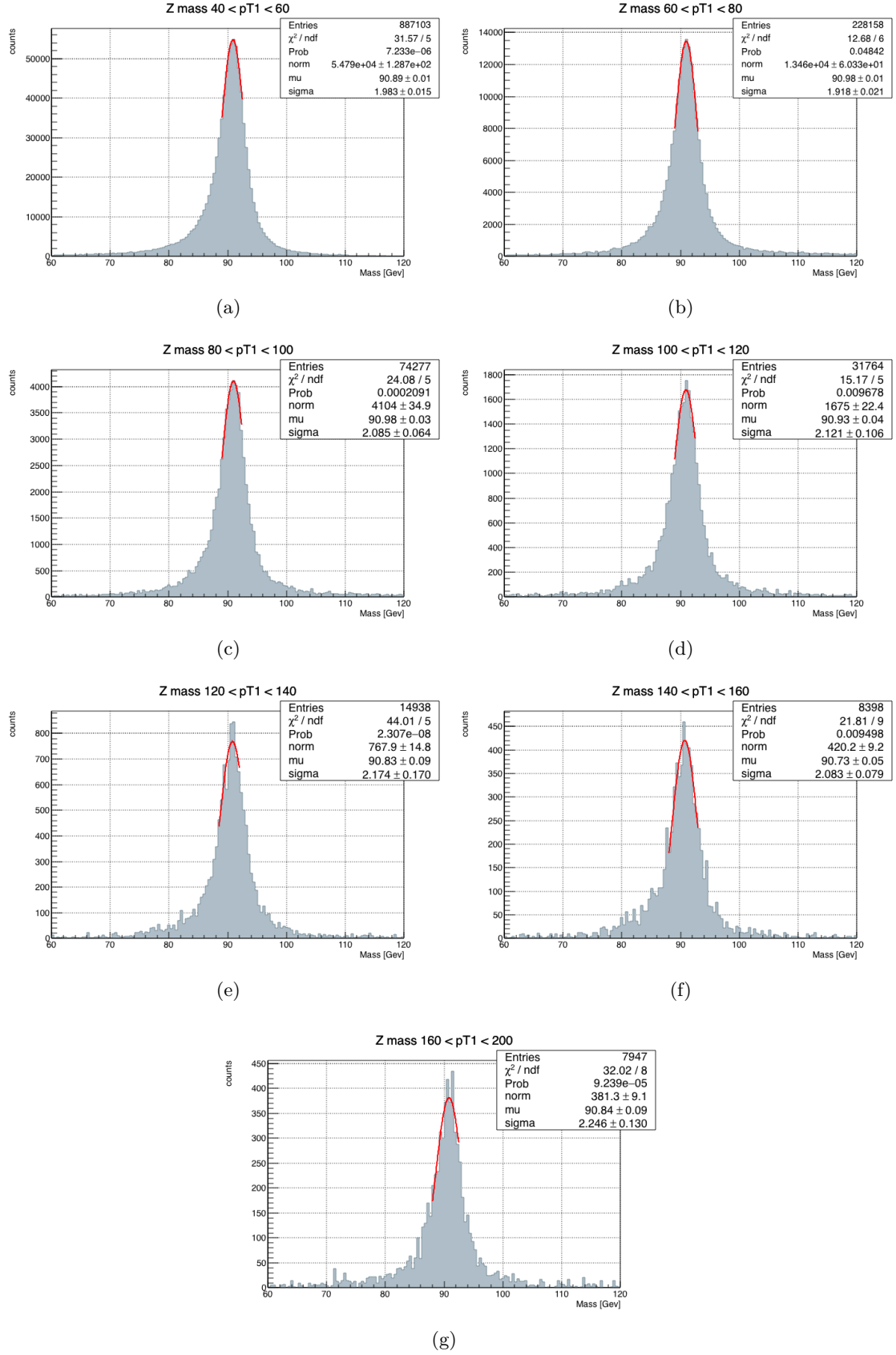


Figure B.14: Gaussian fits of the Z-boson invariant mass distribution in simulation plotted as a function of the p_T of the first electron for the 76X release. The simulated electron energies are corrected with the $f_a(p_T)$ function extracted from fig.4.10.

APPENDIX B. FITS OF THE RESONANCES INVARIANT MASS DISTRIBUTION FOR THE 76X RELEASE

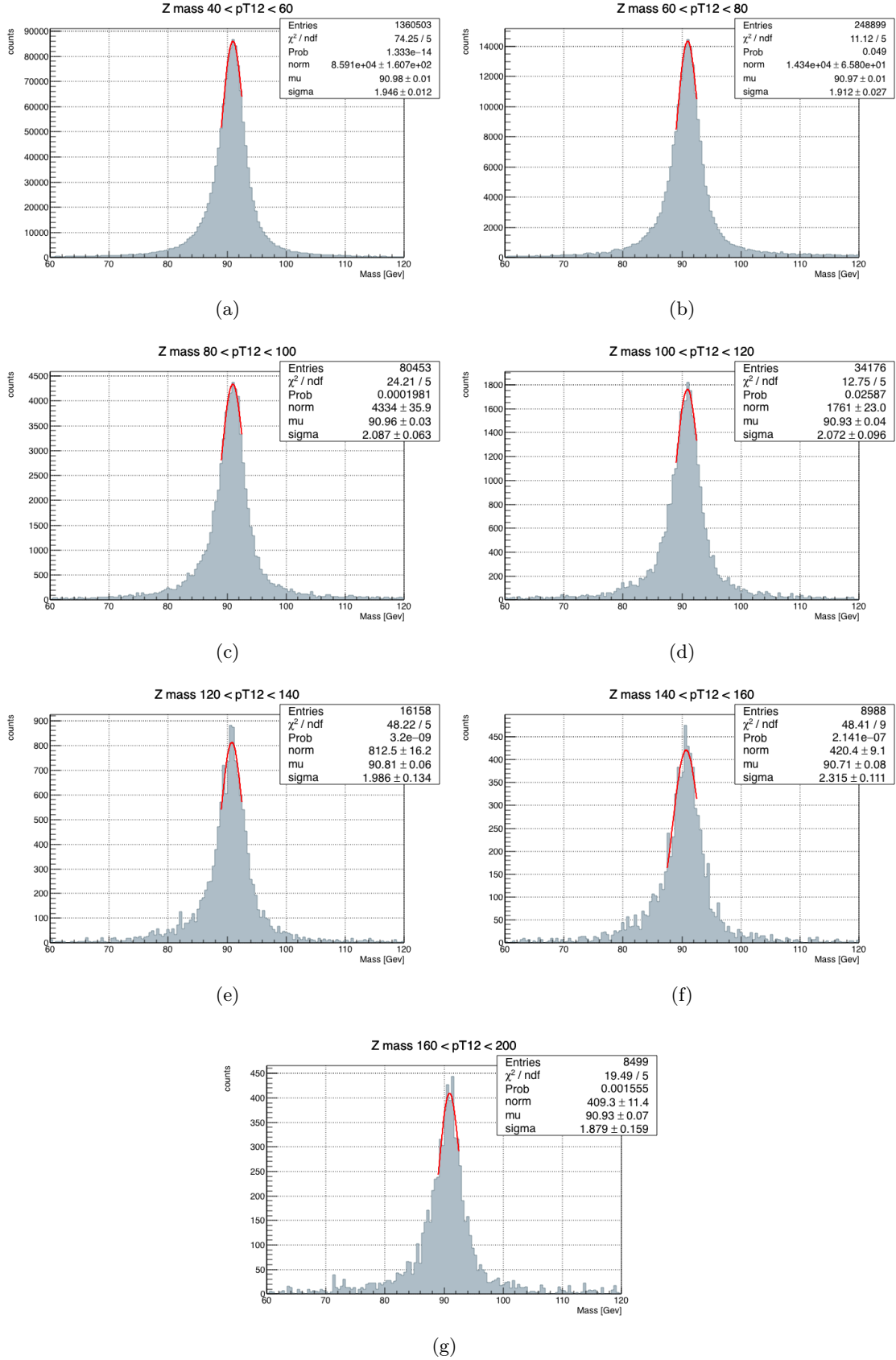


Figure B.15: Gaussian fits of the Z-boson invariant mass distribution in simulation plotted as a function of the p_T of one of the two electrons for the 76X release. The simulated electron energies are corrected with the $f_b(p_T)$ function extracted from fig.4.10.

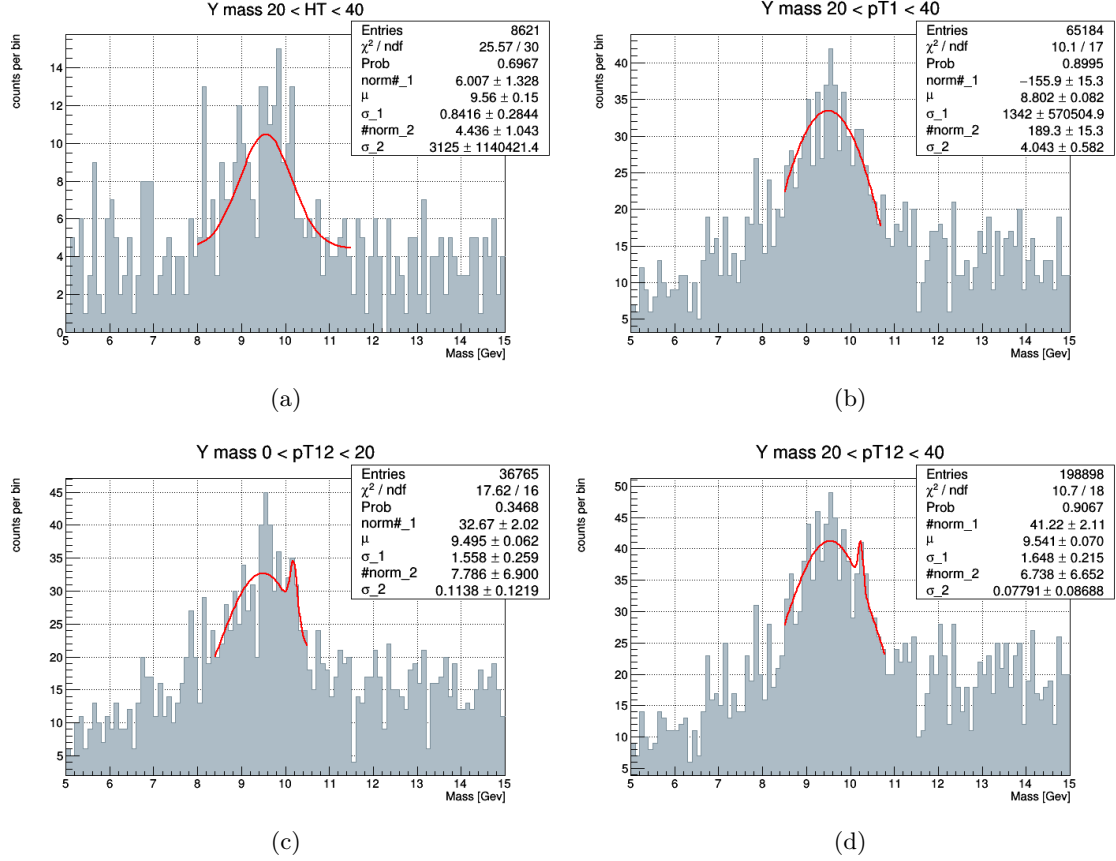


Figure B.16: Gaussian fits of the Y-meson invariant mass distribution plotted using different kinematical variables for the 76X release. The histograms corresponding to the $0 \text{ GeV} < H_T < 20 \text{ GeV}$ and $0 \text{ GeV} < p_{T1} < 20 \text{ GeV}$ are not displayed to the lack of statistics.

Bibliography

- [1] I. Brock and T. Schörner-Sadenius. *Physics at the Terascale*. Wiley, 2011.
- [2] M.D. Schwartz. *Quantum Field Theory and the Standard Model*. Quantum Field Theory and the Standard Model. Cambridge University Press, 2014.
- [3] Steven Weinberg. The Making of the standard model. *Eur. Phys. J.*, C34:5–13, 2004. [,99(2005)].
- [4] G. Hooft. The making of the standard model. *nature*, 448:271–273, jul 2007.
- [5] MissMJ and Nasfarley88. Standard Model of Elementary Particles. http://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg.
- [6] Particle Data Group. Z boson mass. <http://pdg.lbl.gov/2012/listings/rpp2012-list-z-boson.pdf>.
- [7] Particle Data Group. W boson mass. <http://pdg.lbl.gov/2012/listings/rpp2012-list-w-boson.pdf>.
- [8] R. D. Peccei. Discrete and global symmetries in particle physics. *Lect. Notes Phys.*, 521:1–50, 1999.
- [9] I.J.R. Aitchison and A.J.G. Hey. *Gauge Theories in Particle Physics: Volume I: From Relativistic Quantum Mechanics to QED, Third Edition*. Graduate Student Series in Physics. CRC Press, 2002.
- [10] Abdus Salam. Weak and Electromagnetic Interactions. *Conf. Proc.*, C680519:367–377, 1968.
- [11] Ansgar Denner. Techniques for calculation of electroweak radiative corrections at the one loop level and results for W physics at LEP-200. *Fortsch. Phys.*, 41:307–420, 1993.
- [12] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.
- [13] Abdelhak Djouadi. The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model. *Phys. Rept.*, 457:1–216, 2008.

- [14] S. Dittmaier et al. Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables. 2011.
- [15] S. Dawson. Radiative corrections to Higgs boson production. *Nucl. Phys.*, B359:283–300, 1991.
- [16] A. Djouadi, M. Spira, and P. M. Zerwas. Production of Higgs bosons in proton colliders: QCD corrections. *Phys. Lett.*, B264:440–446, 1991.
- [17] T. Figy, C. Oleari, and D. Zeppenfeld. Next-to-leading order jet distributions for Higgs boson production via weak boson fusion. *Phys. Rev.*, D68:073005, 2003.
- [18] Paolo Bolzoni, Marco Zaro, Fabio Maltoni, and Sven-Olaf Moch. Higgs production at NNLO in QCD: The VBF channel. *Nucl. Phys. Proc. Suppl.*, 205-206:314–319, 2010.
- [19] T. Han and S. Willenbrock. Qcd correction to the $pp \rightarrow wh$ and zh total cross sections. *Physics Letters, Section B: Nuclear, Elementary Particle and High-Energy Physics*, 273(1-2):167–172, 12 1991.
- [20] W. Beenakker, S. Dittmaier, M. Kramer, et al. NLO QCD corrections to t anti- t H production in hadron collisions. *Nucl. Phys.*, B653:151–203, 2003.
- [21] Serguei Chatrchyan et al. Search for the standard model Higgs boson decaying into two photons in pp collisions at $\sqrt{s} = 7$ TeV. *Phys. Lett.*, B710:403–425, 2012.
- [22] Serguei Chatrchyan et al. Measurement of the properties of a Higgs boson in the four-lepton final state. *Phys. Rev.*, D89(9):092007, 2014.
- [23] Serguei Chatrchyan et al. Search for the standard model Higgs boson decaying to WW in the fully leptonic final state in pp collisions at $\sqrt{s} = 7$ TeV. *Phys. Lett.*, B710:91–113, 2012.
- [24] Serguei Chatrchyan et al. Search for neutral Higgs bosons decaying to tau pairs in pp collisions at $\sqrt{s} = 7$ TeV. *Phys. Lett.*, B713:68–90, 2012.
- [25] Serguei Chatrchyan et al. Search for the standard model Higgs boson decaying to bottom quarks in pp collisions at $\sqrt{s} = 7$ TeV. *Phys. Lett.*, B710:284–306, 2012.
- [26] G. Aad, T. Abajyan, B. Abbott, et al. Observation of a new particle in the search for the standard model higgs boson with the {ATLAS} detector at the {LHC}. *Physics Letters B*, 716(1):1 – 29, 2012.
- [27] S. Chatrchyan, V. Khachatryan, A.M. Sirunyan, et al. Observation of a new boson at a mass of 125 gev with the {CMS} experiment at the {LHC}. *Physics Letters B*, 716(1):30 – 61, 2012.
- [28] R. Barate et al. Search for the standard model Higgs boson at LEP. *Phys. Lett.*, B565:61–75, 2003.

-
- [29] S. Chatrchyan, V. Khachatryan, A. M. Sirunyan, et al. Measurement of the properties of a higgs boson in the four-lepton final state. *Phys. Rev. D*, 89:092007, May 2014.
- [30] Georges Aad et al. Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments. *Phys. Rev. Lett.*, 114:191803, 2015.
- [31] Lyndon Evans and Philip Bryant. LHC Machine. *JINST*, 3:S08001, 2008.
- [32] *LEP design report*. CERN, Geneva, 1983. By the LEP Injector Study Group.
- [33] S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008.
- [34] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.
- [35] A. Augusto Alves, Jr. et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008.
- [36] K. Aamodt et al. The ALICE experiment at the CERN LHC. *JINST*, 3:S08002, 2008.
- [37] TE-EPC-LPC in LHC. CERN’s accelerator complex. <http://te-epc-lpc.web.cern.ch/te-epc-lpc/machines/lhc/general.stm>, October 2016.
- [38] Taking a closer look at LHC. Lhc layout. https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.lhc_layout.
- [39] D. Decamp et al. ALEPH: A detector for electron-positron annihilations at LEP. *Nucl. Instrum. Meth.*, A294:121–178, 1990. [Erratum: Nucl. Instrum. Meth.A303,393(1991)].
- [40] The CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004, 2008.
- [41] Compact Muon Solenoid at CERN. CMS Collaboration. <http://cms.web.cern.ch/content/cms-collaboration>.
- [42] S Chatrchyan et al. Performance of the CMS Drift-Tube Local Trigger with Cosmic Rays. *JINST*, 5:T03003, 2010.
- [43] Compact Muon Solenoid at CERN. CMS Detector Design. <http://cms.web.cern.ch/news/cms-detector-design>.
- [44] The CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *Journal of Instrumentation*, 9(10):P10009, 2014.
- [45] Bora Isildak. *Measurement of the differential dijet production cross section in proton-proton collisions at $\sqrt{s} = 7$ TeV*. PhD thesis, Bogazici U., 2011.

- [46] *The CMS electromagnetic calorimeter project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.
- [47] K. Deiters, M. Diemoz, N. Godinovic, et al. Investigation of the avalanche photodiodes for the CMS electromagnetic calorimeter operated at high gain. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 461:574 – 576, 2001. 8th Pisa Meeting on Advanced Detectors.
- [48] K.W. Bell, R.M. Brown, D.J.A. Cockerill, et al. Vacuum phototriodes for the CMS electromagnetic calorimeter endcap. *IEEE Transactions on Nuclear Science*, 51(5):2284 – 2287, 2004.
- [49] *The CMS hadron calorimeter project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.
- [50] Compact Muon Solenoid at CERN. XDAQ - CMS Online Software project page. <http://xdaq.web.cern.ch/xdaq/setup/images/HCAL.png>.
- [51] *The CMS magnet project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.
- [52] *The CMS muon project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.
- [53] G L Bayatyan, N Grigorian, V G Khachatryan, et al. *CMS TriDAS project: Technical Design Report, Volume 1: The Trigger Systems*. Technical Design Report CMS.
- [54] Sergio Cittolin, Attila Racz, and Paris Sphicas. *CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project*. Technical Design Report CMS. CERN, Geneva, 2002.
- [55] The CMS high level trigger. *The European Physical Journal C - Particles and Fields*, 46(3):605–667, 2006.
- [56] S Dittmaier, C Mariotti, G Passarino, et al. Handbook of LHC Higgs Cross Sections: 2. Differential Distributions. Technical Report arXiv:1201.3084. CERN-2012-002, Geneva, 2012. Comments: 275 pages, 136 figures, to be submitted to CERN Report. Working Group web page: <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections>.
- [57] S. Baffioni, C. Charlot, F. Ferri, et al. Electron Reconstruction in CMS. *CMS Note 2006/40*, 2006.
- [58] E. Meschi, T. Monteiro, C. Seez, and P. Vikas. Electron Reconstruction in the CMS Electromagnetic Calorimeter. *CMS Note 2001/034*, 2001.
- [59] M Anderson, A Askew, A.F. Barfuss, et al. Review of clustering algorithms and energy corrections in ECAL. *CMS Internal Note 2010/008*, 2010.

-
- [60] Vardan Khachatryan et al. Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV. *JINST*, 10(06):P06005, 2015.
- [61] W. Adam, R. Frühwirth, A. Strandlie, and T. Todor. Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC. Technical Report CMS-NOTE-2005-001, CERN, Geneva, Jan 2005.
- [62] Serguei Chatrchyan et al. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *JINST*, 9(10):P10009, 2014.
- [63] W. Adam et others. Electron Reconstruction in CMS. (CMS-NOTE-2009-164), Nov 2009.
- [64] M. Anfreville, D. Bailleux, J.P. Bard, et al. Laser monitoring system for the {CMS} lead tungstate crystal calorimeter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 594(2):292 – 320, 2008.
- [65] ECAL DPG. ECAL Laser monitoring till end of 2016 and ECAL phi-symmetry. https://twiki.cern.ch/twiki/pub//CMSPublic/EcalDPGResultsCMSDPS2017003/histories_2011-2012-2015-2016_161212.png.
- [66] Serguei Chatrchyan et al. Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp Collisions at $\sqrt{s} = 7$ TeV. *JINST*, 8:P09009, 2013. [JINST8,9009(2013)].
- [67] ECAL DPG. 2015 ECAL detector performance plots. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/EcalDPGResultsCMSDPS2015057>.
- [68] Properties of the Higgs-like boson in the decay H to ZZ to $4l$ in pp collisions at $\sqrt{s} = 7$ and 8 TeV. Technical Report CMS-PAS-HIG-13-002, CERN, Geneva, 2013.
- [69] Measurements of properties of the Higgs boson and search for an additional resonance in the four-lepton final state at $\sqrt{s} = 13$ TeV. Technical Report CMS-PAS-HIG-16-033, CERN, Geneva, 2016.
- [70] Ernesto Migliore and Giacomo Sguazzoni. Altered scenarios of the CMS Tracker material for systematic uncertainties studies. 2010.
- [71] Studies of Tracker Material. Technical Report CMS-PAS-TRK-10-003, 2010.
- [72] N. Amapane et al. Measurement of the production and decay of a Higgs boson in the four-lepton final state. Technical Report CMS-NOTE-2013-108, CERN, Geneva, Nov 2013.
- [73] Particle Data Group. The review of Particle Physics (2016). <http://www-pdg.lbl.gov>.

- [74] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching nlo qcd computations with parton shower simulations: the powheg method. *Journal of High Energy Physics*, 2007(11):070, 2007.
- [75] J. Alwall, R. Frederix, S. Frixione, et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.
- [76] Torbjorn Sjostrand, Stefan Ask, Jesper R. Christiansen, et al. An Introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015.
- [77] S. Agostinelli, J. Allison, K. Amako, et al. Geant4: a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250 – 303, 2003.
- [78] Particle Data Group. Y meson mass. <http://pdg.lbl.gov/2015/tables/rpp2015-tab-mesons-b-bbar.pdf>.
- [79] Particle Data Group. $J\psi$ meson mass. <http://pdg.lbl.gov/2015/tables/rpp2015-tab-mesons-c-cbar.pdf>.

The most exciting phrase to hear in science, the one that heralds new discoveries, is not 'Eureka!' but 'That's funny...'

ISAAC ASIMOV (1920-1992)
Russian born, American Writer and Biochemist