



Università degli Studi di Trieste

DIPARTIMENTO DI FISICA

Laurea Magistrale in Fisica
Curriculum in Fisica Nucleare e Subnucleare

Associated production of the Higgs boson with a vector boson at the LHC: study of the CMS electromagnetic calorimeter and definition of the Run2 analysis selections.

Author

LUCIA RINCHIUSO

.....

Supervisor

Prof. GIUSEPPE DELLA RICCA

.....

co-Supervisor

Dott. FEDERICO FERRI

.....





Università degli Studi di Trieste

DIPARTIMENTO DI FISICA

Laurea Magistrale in Fisica
Curriculum in Fisica Nucleare e Subnucleare

**Produzione associata del bosone di Higgs con un bosone
vettore a LHC: studio del calorimetro elettromagnetico di
CMS e definizione delle selezioni dell'analisi del Run2.**

Laureando

LUCIA RINCHIUSO

.....

Relatore

Prof. GIUSEPPE DELLA RICCA

.....

co-Relatore

Dott. FEDERICO FERRI

.....

ANNO 2015/2016

*“Le papillon préfère se brûler à la torche
plutôt que vivre dans les ténèbres”.*

Saadi

Contents

Introduzione	II
Introduction	V
1 The Standard Model and the Higgs boson	1
1.1 Weak interaction	4
1.1.1 Charged weak interaction	4
1.1.2 Neutral weak interaction and Electroweak unification	7
1.2 Strong interaction	8
1.2.1 Deep inelastic scattering and parton model	8
1.2.2 Quantum Chromodynamics	11
1.2.3 Hadron collisions and Drell-Yan process	12
1.3 The Higgs boson	14
1.3.1 Spontaneous symmetry breaking	15
1.3.2 Higgs mechanism	16
1.3.3 Higgs boson in the Standard Model	17
1.3.4 Higgs boson production and decay channels	19
1.3.5 Higgs boson discovery in 2012	23
2 The LHC and CMS	25
2.1 The Large Hadron Collider (LHC): overview	26
2.2 Compact Muon Solenoid (CMS): overview	29
2.2.1 Superconducting solenoid	30
2.2.2 Tracker	31
2.2.3 Electromagnetic Calorimeter ECAL	34
2.2.4 Hadron Calorimeter HCAL	35
2.2.5 Muon detector	36
2.2.6 Trigger	39
2.3 Electromagnetic Calorimeter ECAL	40
2.3.1 Showering and calorimetry	40
2.3.2 Material and layout	43

2.3.3	Calibration	46
2.3.4	Resolution	49
3	Electrons and photons reconstruction for Higgs boson studies	53
3.1	Particle identification	54
3.2	Energy measurement and clustering	55
3.2.1	Clustering	56
3.2.2	Track assignment	57
3.3	Photons preselection for Higgs analysis	58
3.4	Photon identification for Higgs analysis	60
3.5	Pileup	61
3.6	Vertex identification for Higgs analysis	62
3.7	Diphoton selection for Higgs analysis	63
3.8	Tags and classification for Higgs analysis	65
3.9	Signal and background models for Higgs analysis	65
3.10	Systematic uncertainties for Higgs analysis	67
4	Detector studies	69
4.1	Study of the material budget in front of ECAL	69
4.2	Shower shape studies	78
5	Analysis of the Higgs boson	93
5.1	Monte Carlo studies	93
5.2	Higgs boson associated production with a vector boson and decay in two photons	99
	Conclusions	112
	Acronyms	114
	Bibliography	119
	Ringraziamenti	127

Introduzione

Uno dei quesiti che sono sempre stati parte della natura umana riguarda di cosa è fatto il mondo che ci circonda. Il Modello Standard, sviluppato nella seconda metà del secolo scorso, ha provato a rispondere parzialmente con una descrizione dell'Universo basata su 12 tasselli fondamentali (quark e leptoni) e 4 mediatori delle interazioni fondamentali (bosoni di gauge). Numerose conferme del modello sono giunte da diversi esperimenti. Gli unici aspetti del modello non confermati sino al 2012 erano l'esistenza del bosone di Higgs e il meccanismo attraverso il quale i bosoni di gauge massivi così come i quarks e i leptoni acquistano massa.

Per confermare queste previsioni teoriche sono state sviluppate tecnologie estremamente avanzate. Il Large Hadron Collider (LHC), l'acceleratore più potente mai costruito, è entrato in operazione nel 2008 con quattro esperimenti principali: CMS, ATLAS, LHCb e ALICE. Tra questi il Compact Muon Solenoid (CMS) è un rivelatore molto versatile impiegato per studiare le particelle che fuoriescono da collisioni protone-protone per testare il Modello Standard e cercare segni di nuova fisica oltre il Modello Standard. Uno dei fini ultimi di CMS era infatti osservare il bosone di Higgs predetto da Peter Higgs negli anni Sessanta e determinarne la massa. Lo scopo è stato raggiunto nel 2012 quando è stata osservata una particella neutra di spin 0 con massa circa 125 GeV, compatibile con le speculazioni sul bosone di Higgs all'interno del Modello Standard. Dopo l'upgrade di LHC per il Run2 a piena luminosità a 13 TeV, il bosone di Higgs deve essere nuovamente scoperto in un ambiente ancor più ostile, con una gran quantità di eventi di fondo. Inoltre, in questo nuovo intervallo energetico ci si può aspettare di scoprire fenomeni legati a nuova fisica.

Innanzitutto, per osservare il bosone di Higgs è necessario un lavoro preliminare per calibrare e mettere a punto il rivelatore e sviluppare o migliorare i generatori di eventi Monte Carlo che permettono di confrontare i dati osservati con le aspettative teoriche. Una serie di controlli deve essere eseguita prima d'iniziare l'analisi vera e propria per poter tener conto, ad esempio, degli effetti del rivelatore, come l'interazione delle particelle con il materiale del rivelatore stesso, che può influenzare la risoluzione della misura della massa del bosone di Higgs. In questa tesi vengono toccati diversi aspetti dell'analisi: studi del rivelatore, studi dei generatori di eventi, selezione degli eventi del decadimento del bosone di Higgs.

Nel capitolo 1 sono introdotti una breve descrizione del Modello Standard, il meccanismo di Higgs, i canali di produzione e decadimento del bosone di Higgs e la sua scoperta.

Nel capitolo 2 è presentata una panoramica dell'acceleratore e del rivelatore, con parti-

colare attenzione al calorimetro elettromagnetico (ECAL), fondamentale per la misura e l'identificazione dei fotoni dal decadimento del bosone di Higgs, e degli elettroni dal decadimento del bosone Z , che sono utilizzati per la calibrazione del metodo d'analisi e per stimare la risoluzione delle misure d'energia.

Nel capitolo 3 è descritto più in dettaglio il metodo di ricostruzione degli elettroni e fotoni: la misura dell'energia e l'associazione di una traccia ai fotoni convertiti. È inoltre descritta la selezione dei fotoni e dei candidati difotone promettenti come prodotti del decadimento del bosone di Higgs, l'assegnazione del vertice primario e la classificazione degli eventi associati al bosone di Higgs nei diversi canali.

Nel capitolo 4 è spiegato il lavoro legato agli studi del rivelatore. In particolare, si focalizza l'attenzione sul materiale posto di fronte a ECAL, dove gli elettroni perdono parte della propria energia per bremsstrahlung, e sulle variabili che caratterizzano lo sviluppo del profilo dello sciame elettromagnetico nel calorimetro e le loro distribuzioni in diverse condizioni: con e senza il campo magnetico e con due diversi algoritmi di ricostruzione del segnale. Il fine è quello di indagare alcune delle possibili fonti di discrepanza tra dati e simulazioni.

Nel capitolo 5 è descritto il lavoro sugli studi del bosone di Higgs: confronto dei principali generatori di eventi Monte Carlo impiegati a CMS e classificazione degli eventi che devono essere analizzati per la ricerca del bosone di Higgs nel canale di produzione in associazione con un bosone vettore e decadimento in due fotoni.

Introduction

One of the questions that have always been part of men's nature is about what the world around us is made of. The Standard Model, developed in the second half of the last century, tried to answer partially, with a description of the Universe based on 12 fundamental building blocks (quarks and leptons) and four mediators of the fundamental forces (gauge bosons). Many confirmations of the model came from several experiments. The only unconfirmed aspects of the model until 2012 were the existence of the Higgs boson and the mechanism through which it gives mass to the massive gauge bosons as well as quarks and leptons.

Advanced technologies were developed to confirm the speculations. The most powerful accelerator built so far, the Large Hadron Collider (LHC), became operational in 2008 with four main experiments: CMS, ATLAS, LHCb and ALICE. Among them the Compact Muon Solenoid (CMS) is a general-purpose detector employed to study the particles that emerge from proton-proton collisions in order to test the Standard Model and search for signatures of new physics beyond the Standard Model. Indeed, one of the ultimate goals of CMS was to observe the Higgs boson predicted by Peter Higgs in the Sixties and to determine its mass. The aim was achieved in 2012 when a new spin-0 neutral particle was observed with a mass of about 125 GeV compatible with the expectations for Standard Model Higgs boson. After the upgrade of LHC to Run 2 at full luminosity at 13 TeV, the Higgs boson must be discovered again in this harsh environment, with severe background conditions. Moreover, in this energetic range new physics phenomena can be expected.

First, to observe the Higgs boson a preliminary work is needed to calibrate and tune the detector and to develop or improve the Monte Carlo events generators that allow to compare the observed data to the theoretical expectations. A series of checks must be performed before the analysis itself to account, for example, for detector effects, as the interaction of the particles with the detector material which can affect the resolution of the measured mass of the Higgs boson. In this thesis diverse points of the analysis will be touched: the studies of the detector, the studies of event generators, the selection of events coming from the Higgs boson decay.

In chapter 1 a brief description of the Standard Model, the Higgs mechanism, the channels of production and decay of the Higgs boson and its discovery are introduced.

In chapter 2 an overview of the accelerator and the detector is presented, pointing the attention on the electromagnetic calorimeter (ECAL), crucial for the measurement and identification of the photons from the decay of the Higgs boson in the diphoton system and electrons from the

decay of the Z boson which are used for the calibration of the method of analysis and to estimate the resolution of the energy measurements.

Chapter 3 describes the method of reconstruction of electrons and photons in more details: the energy measurement and the association of a track to the converted photons. Furthermore, the selection of the photons and of the promising diphoton candidates from Higgs boson decay, the assignment of the primary vertex and the classification of the events associated to the Higgs boson in different channels are described.

Chapter 4 presents the work on the detector studies. In particular, the attention is focused on the tracking material in front of ECAL, where the electrons lose part of their energy by bremsstrahlung, and on the shower shape variables and their distributions in different conditions: with and without the magnetic field and with two different algorithms of signal reconstruction. The aim is to investigate some of the possible sources of discrepancy between data and simulations.

Chapter 5 shows the work on the studies of the Higgs boson: comparison of the main Monte Carlo events generators employed at CMS, and classification of the events to be analyzed for the Higgs boson search in the associated production channel with a vector boson and decay mode into two photons.

Chapter 1

The Standard Model and the Higgs boson

The Standard Model (SM) [1] was developed in the early Seventies to give an explanation to the numerous experimental discoveries of the last century in particle physics. It is aimed to describe the components of matter and three of the fundamental forces, and its “building blocks” are the elementary fermions and the gauge bosons. It is based on a Quantum Field Theory that interprets particles as excitations of quantum fields which interact with each other by exchanging the gauge bosons.

The *fermions* are six leptons and six quarks, divided in three generations of two leptons and two quarks. The four *gauge bosons* are the mediators of electromagnetic, strong and weak interaction, while gravity is not included in the SM. In addition there is another boson that gives mass to the other particles through a mechanism of spontaneous symmetry breaking: the *Higgs boson*.

Figure 1.1 shows a scheme of the SM. At the bottom, the leptons are shown: e is stable and builds the ordinary matter, while μ and τ are instable and decay into lighter particles. On the top are the quarks that clump together to create hadrons: mesons ($q\bar{q}$) and barions (qqq). On the right are the gauge bosons: the photon for the electromagnetic interaction, the gluon for the strong interaction and the $W^\pm$ and Z^0 bosons for the weak interaction.

In tables 1.1 and 1.2 the main properties of these particles are summarized [2].

gauge bosons				
interaction	electromagnetic	weak		strong
coupling (relative strength)	1	1/137		10^{-6}
particle	γ	$W^\pm$	Z^0	g
mass (GeV)	0	80.385 ± 0.015	91.1876 ± 0.0021	0
charge	0	± 1	0	0

Table 1.1: Properties of gauge bosons.

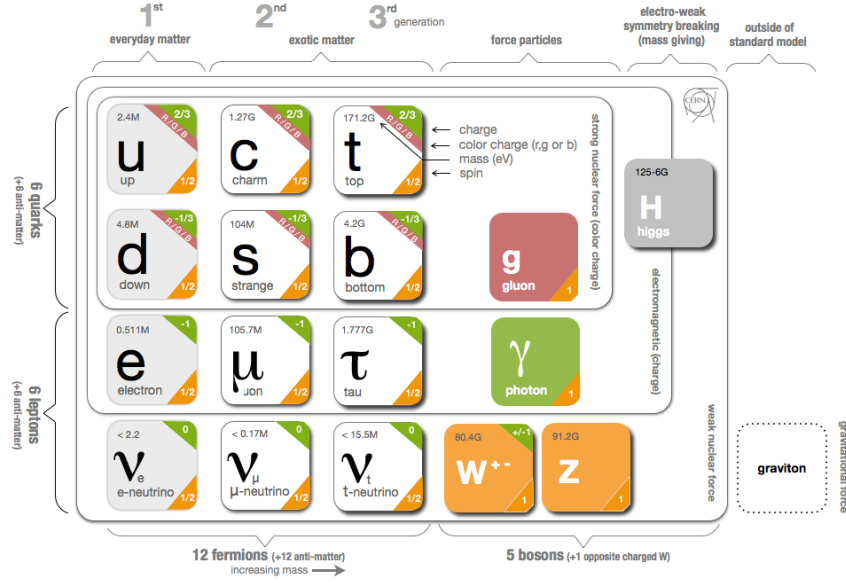


Figure 1.1: Scheme of the Standard Model.

For each particle an antiparticle exists, with same mass and lifetime (or decay rate) but opposite quantum numbers.

In nature there are some conservation laws (consequence of symmetries) that must be conserved by all or some of the forces.

Discrete symmetries, broken by weak interaction, are: *parity* (or space reflection) P such that an inversion of spacial coordinates $\vec{r} \rightarrow -\vec{r}$ takes place, *charge conjugation* C , that transforms a particle in its antiparticle and vice versa, and *time reversal* T $t \rightarrow -t$. The weak interaction breaks also the CP symmetry while the CPT symmetry must be conserved by all the forces (CPT theorem), in this way we assure that the mass and the lifetime of all particles are the same of those of their antiparticles.

Due to internal continuous symmetries all the fundamental interactions must respect electric charge conservation, baryon number (+1 for baryons, -1 for antibaryons, +1/3 for quarks and -1/3 for antiquarks) conservation and lepton number conservation where both the total and the single electronic, muonic and tauonic lepton numbers are conserved.

In this model each interaction, except gravity, is described by a quantum field theory. The first to be formulated is Quantum Electrodynamics (QED) for phenomena involving electrically charged particles that interact exchanging photons. These virtual photons mediate the momentum exchange between particles, that manifests as electromagnetic force. It is the simplest of the dynamical theories, that comes naturally out of the invariance under a local gauge transformation in $U(1)$, and a starting point for the others. In this context Feynman diagrams are introduced with the corresponding rules to compute the amplitude of a subatomic process.

leptons		
flavor	mass (MeV)	charge
electron	$0.510998928 \pm 0.000000011$	-1
muon	$105.6583715 \pm 0.0000035$	-1
tau	1776.82 ± 0.16	-1
$\nu_{e,(\mu,\tau)}$	$< 2 \times 10^{-6}$	0

quarks		
flavor	mass (MeV)	charge
up	$2.3^{+0.7}_{-0.5}$	$\frac{2}{3}$
down	$4.8^{+0.5}_{-0.3}$	$-\frac{1}{3}$
charm	$(1.275 \pm 0.025) \times 10^3$	$\frac{2}{3}$
strange	95 ± 5	$-\frac{1}{3}$
top	$(173.21 \pm 0.51 \pm 0.71) \times 10^3$	$\frac{2}{3}$
bottom	$(4.18 \pm 0.03) \times 10^3$	$-\frac{1}{3}$

Table 1.2: Properties of fermions.

Similarly to QED, quantum field theories are elaborated for the weak interaction, mediated by massive gauge bosons, and for the strong force, whose carriers are eight massless gauge bosons characterized by the color charge.

The development of the Standard Model is one of the most important achievements of modern physics, but it is still incomplete. There are some problems that can only be explained by theories beyond the SM:

- neutrino oscillations (change of flavor), possible only if they have non-null mass [3];
- dark matter (neutral, colorless, non-barionic and massive matter) and dark energy which form the 96% of the Universe [4];
- matter-antimatter asymmetry: it is still obscure the reason why the Universe that we know is made of matter if it was born with the same amount of matter and antimatter during the Big Bang [5];
- gravitation [6], with a relative strength of 10^{-39} , hypothetically mediated by the graviton, the only gauge boson with spin 2. This force starts to dominate at the Planck Scale, about 10^{19} GeV, where quantum field theory and general relativity are no longer reconcilable;
- unification theories: Grand Unified Theory (GUT) [7], that merges electromagnetic, strong and weak interaction at a scale of approximately 10^{14} GeV, and Theory of Everything (TOE) that includes also gravity.

1.1 Weak interaction

The interaction that is responsible for example of the radioactive decays is the weak interaction. Unlike the electromagnetic and nuclear forces the weak force is carried by massive gauge bosons. In addition to the charged W bosons responsible for the well known beta decay, also the neutral Z boson is predicted and its discovery revealed the common nature of electromagnetism and weak force which are unified in the electroweak interaction [8].

1.1.1 Charged weak interaction

The $W^\pm$ bosons are the vector bosons of the charged weak interaction. The massive propagator associated to those bosons of mass about 80 GeV is

$$-\frac{ig_{\mu\nu}}{q^2 - m_W^2}$$

which is approximated to $-\frac{ig_{\mu\nu}}{m_W^2}$ at $q^2 \ll m_W^2$. In this limit of low momentum exchange Fermi first described the weak interaction in 1934 [9] as a contact interaction

$$\mathcal{M}_{fi} = G_F g_{\mu\nu} [\bar{\psi}_3 \gamma^\mu \psi_1] [\bar{\psi}_4 \gamma^\nu \psi_2]. \quad (1.1)$$

But this force has a peculiarity that was discovered by Wu in 1957 [10]: it does not conserve parity. It was observed in the nuclear β -decay $^{60}\text{Co} \rightarrow ^{60}\text{Ni}^* + e^- + \bar{\nu}_e$, hence in the underlying weak decay $n \rightarrow p + e^- + \bar{\nu}_e$, that the electron is always emitted in a preferential direction opposite to the nuclear spin, as shown in figure 1.2. The fact that the electron is always observed

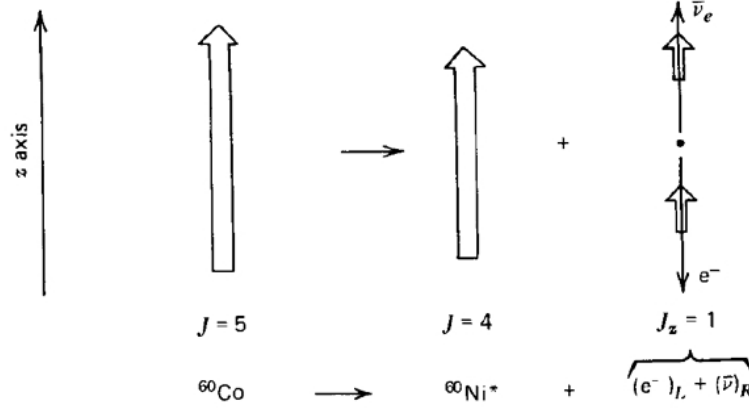


Figure 1.2: ^{60}Co experiment that lead to the discovery of the P violation by the weak interaction.

left-handed and the antineutrino right-handed, while the “mirror process” is absent, is a clear violation of parity conservation.

From this ascertainment, it is known that the weak processes need a formalism which differs from the “simple” vector interaction in QED.

In general the current is expressed as $\psi O \psi$ where the operator O can be

- a scalar $O = 1$, even under parity, associated to the exchange of a spin-0 boson;
- a pseudoscalar $O = \gamma^5$, odd under parity, associated to the exchange of a spin-0 boson;
- a vector $O = \gamma^\mu$, odd under parity, associated to the exchange of a spin-1 boson;
- an axial vector $O = \gamma^\mu \gamma^5$, even under parity, associated to the exchange of a spin-1 boson;
- a tensor $O = \sigma^{\mu\nu} = \frac{1}{2}[\gamma^\mu, \gamma^\nu]$, with a more complex behavior under parity, associated to the exchange of a spin-2 boson.

The most general interaction between a fermion and a spin-1 boson is a linear combination of the bilinear vector and axial vector covariants: **V-A interaction**.

The weak current is $j^\mu \propto \bar{\psi}(g_V \gamma^\mu + g_A \gamma^\mu \gamma^5)\psi = g_V j_V^\mu + g_A j_A^\mu$ and under parity it transforms like:

$$\begin{aligned} j_V^0 &\rightarrow +j_V^0, j_V^k \rightarrow -j_V^k \\ j_A^0 &\rightarrow -j_A^0, j_A^k \rightarrow +j_A^k \end{aligned}$$

and

$$j_A \cdot j_A \rightarrow j_A \cdot j_A, j_V \cdot j_V \rightarrow j_V \cdot j_V,$$

but the mixed scalar product is not invariant:

$$\vec{j}_V \cdot \vec{j}_A \rightarrow -\vec{j}_V \cdot \vec{j}_A.$$

Then, the vertex of weak V-A interaction is

$$-\frac{ig_W}{\sqrt{2}} \frac{1}{2} \gamma^\mu (1 - \gamma^5)$$

and the relative strength of the parity violation is $\frac{g_V g_A}{g_V^2 + g_A^2}$.

The weak interaction is associated with the local SU(2) symmetry $\psi(x) \rightarrow e^{ig_W \alpha(x) \cdot T} \psi(x)$. The *weak isospin* $T = \frac{1}{2} \sigma$ is introduced and it merges in multiplets the states that behave in the same way under the weak force. We distinguish particle states that are left-handed isodoublets of SU(2)_L and antiparticle states which are right handed singlets not affected by the local gauge transformation:

$$\begin{aligned} L_1 &= \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L & e_{R1} &= e_R^-, \bar{\nu}_{eR1} = \bar{\nu}_{eR} & Q_1 &= \begin{pmatrix} u \\ d \end{pmatrix}_L & u_{R1} &= u_R, d_{R1} = d_R \\ L_2 &= \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L & e_{R2} &= \mu_R^-, \bar{\nu}_{eR2} = \bar{\nu}_{\mu R} & Q_2 &= \begin{pmatrix} c \\ s \end{pmatrix}_L & u_{R2} &= c_R, d_{R2} = s_R \\ L_3 &= \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L & e_{R3} &= \tau_R^-, \bar{\nu}_{eR3} = \bar{\nu}_{\tau R} & Q_3 &= \begin{pmatrix} t \\ b \end{pmatrix}_L & u_{R3} &= t_R, d_{R3} = b_R \end{aligned}$$

After the discovery of the parity violation the equation 1.1 was changed into

$$\mathcal{M}_{fi} = \frac{g_W^2}{8m_W^2} g_{\mu\nu} [\bar{\psi}_3 \gamma^\mu (1 - \gamma^5) \psi_1] [\bar{\psi}_4 \gamma^\nu (1 - \gamma^5) \psi_2]$$

where the Fermi coupling constant and the charged weak coupling constant are linked by

$$\frac{G_F}{\sqrt{2}} = g_W^2 / 8m_W^2.$$

The value of G_F can be obtained for example from the decay of the muon or the tau and the result is the same: $G_F/(\hbar c)^3 = 1.1663787 \cdot 10^{-5} \text{ GeV}^{-2}$. On the other hand, if the measurement is performed using hadrons the coupling at the vertex depends on the flavor of the quarks. Some of the vertices are shown in figure 1.3. To explain this unexpected behavior Cabibbo

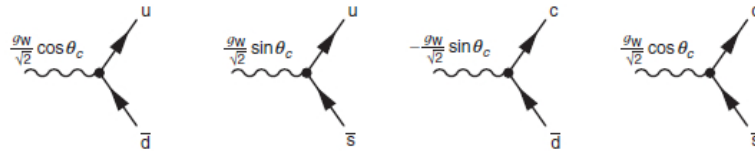


Figure 1.3: Some weak vertices of quarks.

developed a model where the weak eigenstates q' do not correspond to the flavor eigenstates q .

This hypothesis is extended with the introduction of the Cabibbo-Kobayashi-Maskawa matrix V_{CKM} such that

$$\begin{pmatrix} u' \\ c' \\ t' \end{pmatrix} = \begin{pmatrix} u \\ c \\ t \end{pmatrix}, \quad \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}.$$

The CKM matrix can be expressed in different forms:

- in terms of three rotation angles $\phi_{12}, \phi_{13}, \phi_{23}$ and a complex phase δ (linked to CP violation [11]):

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}$$

with $c_{ij} = \cos \phi_{ij}$ and $s_{ij} = \sin \phi_{ij}$;

- using Wolfenstein parameters [12] λ, ρ, η, A :

$$V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) - A\lambda^2 & 1 & \end{pmatrix} + \mathcal{O}(\lambda^4)$$

with $\lambda = 0.2253 \pm 0.0007$. So in terms of orders of λ it is easy to determine immediately which vertices are dominant and which suppressed.

The CKM matrix is unitary and its elements are orthonormal.

1.1.2 Neutral weak interaction and Electroweak unification

In the discussion about the charged-weak interaction [8] only two gauge bosons $W^\pm$ were introduced, but the local gauge symmetry in $SU(2)_L$ needs three gauge fields $W_\mu^{(1)}, W_\mu^{(2)}$ and $W_\mu^{(3)}$. The rising/lowering operators, expressed in terms of Pauli matrices, allow to move to the physical bosons as linear combinations of the W components:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^{(1)} \pm W_\mu^{(2)}).$$

Furthermore, the third current j_3 implies the existence of a neutral-current interaction. Also in this case we cannot simply denote $W^{(3)}$ as the Z boson, but the issue is more sensitive.

Glashow, Salam and Weinberg received the Nobel Prize in 1979 for a model that merges together the weak and the electromagnetic interactions into a *unified electroweak theory* [13].

The gauge $U(1)$ symmetry of QED is replaced by a local gauge $U(1)_Y$ symmetry. From that the new gauge field B_μ originates, that couples to the weak hypercharge Y , defined as $Y = 2(Q - I_W^{(3)})$.

Now both the photon and the Z boson are described in terms of $B_\mu, W_\mu^{(3)}$ and the Weinberg angle (or weak mixing angle) θ_W :

$$A_\mu = B_\mu \cos \theta_W + W_\mu^{(3)} \sin \theta_W$$

$$Z_\mu = -B_\mu \sin \theta_W + W_\mu^{(3)} \cos \theta_W,$$

where $\sin^2 \theta_W \sim 0.23$.

The new local gauge symmetry of the electroweak interaction is $SU(2)_L \otimes U(1)_Y$ and the relation between the electromagnetic and the weak couplings is $e = g_W \sin \theta_W = g_Z \sin \theta_W \cos \theta_W$.

1.2 Strong interaction

The strong interaction [14, 15] is responsible for binding together the quarks in the hadrons and the nucleons (protons and neutrons) in the nuclei.

It is described by Quantum Chromodynamics (QCD) and carried by the *gluons*.

1.2.1 Deep inelastic scattering and parton model

It is well known that the structure of a target B can be studied using a low mass particle A as a projectile [14, 16]. In scattering with low momentum exchange q we talk about an elastic process $A + B \rightarrow A + B$, while when the probe loses a large amount of its momentum it can enter the target and break it. In this case we talk about inelastic scattering (see figure 1.4).

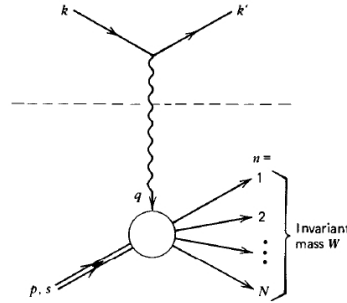


Figure 1.4: Inelastic scattering.

The simplest example of this method to study matter is the elastic $ep \rightarrow ep$ scattering. In the non relativistic limit (spinless) for a point-like proton with infinite mass the cross section, quantity that measures the probability for the process to take place, is just

$$\frac{d\sigma}{d\Omega}_{point} = \frac{\alpha^2}{4} \left(\frac{1}{mv^2} \right)^2 \frac{1}{\sin^4(\theta/2)}.$$

In the relativistic case Mott's cross section is used

$$\frac{d\sigma}{d\Omega}_{Mott} = \left(\frac{d\sigma}{d\Omega} \right)_{point} \cos^2(\theta/2),$$

where the cosine comes from the overlap between initial/final state electron wave-functions. When we consider a finite mass of the target the form factor $F(q)$ is introduced that describes the size of the proton

$$\frac{d\sigma}{d\Omega} = \left(\frac{d\sigma}{d\Omega} \right)_{point} |F(q)|^2.$$

If the spin and the finite mass are taken into account in a point-like model the following formula can be derived

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E^2 \sin^4(\theta/2)} \frac{E'}{E} \left(\cos^2(\theta/2) - \frac{q^2}{2M} \sin^2(\theta/2) \right).$$

But finally in Rosenbluth's equation spin, finite mass and extended charge distribution are all taken into account:

$$\frac{d\sigma}{d\Omega} = \frac{d\sigma}{d\Omega_{point}} \cos^2(\theta/2) \left(\frac{G_E^2 + \frac{q^2}{4M^2} G_M^2}{1 + \frac{q^2}{4M^2}} + \frac{q^2}{4M^2} 2G_M^2 \tan^2(\theta/2) \right).$$

G_E and G_M are called electric and magnetic form factors respectively and are the means to access the distribution of charge.

As show in figure 1.5, when the energy of the electron-proton collision is enough to probe

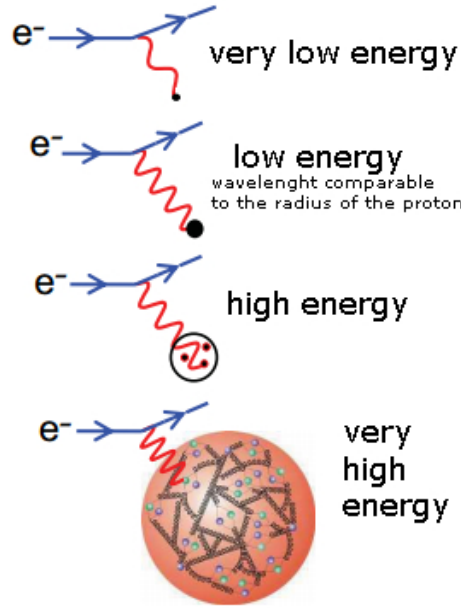


Figure 1.5: Representation of the electron-proton scattering at different energies.

the inside of the nucleus, its internal structure is revealed. This process at very high energy is called deep inelastic scattering (DIS) and it lead to the discovery of the elementary constituents of the nucleons and the development of the **parton model**.

Some point-like particles are resolved inside the proton and they are called *partons*. An interesting feature of the model is that the inelastic interaction of a particle with a proton can be

described as its elastic interaction with a point-like spin 1/2 free parton (later on identified as quark) inside the proton itself.

The *parton distribution functions* (PDF) f_i are introduced in the parton model to give the probability to find a specific quark that carries a fraction x of the longitudinal momentum. x is called *Bjorken variable* and it is also the fraction of the energy and the mass of the proton which is carried by the parton. The non interacting partons are called spectators.

The scattering is now described in terms of the independent variables $q^2 = -Q^2$ and $\nu = E - E'$ or alternatively $x = -\frac{q^2}{2M\nu}$ and $y = \frac{E-E'}{E}$.

The new form factors of the inelastic scattering, interpreted as the elastic scattering of the constituents, are $\nu W_2(\nu, Q^2) = \delta\left(1 - \frac{Q^2}{2M\nu}\right)$ and $MW_1(\nu, Q^2) = \frac{Q^2}{4M\nu}\delta\left(1 - \frac{Q^2}{2M\nu}\right)$ such that

$$\sigma_{DIS} = \sigma_{Mott}\left(W_2 + 2W_1 \tan(\theta/2)\right).$$

At high Q^2 , when the free partons are resolved, they tend to the dimensionless structure functions $F_2(x) = \sum_i e_i^2 f_i(x)$ and $F_1(x) = \frac{1}{2x}F_2(x)$. This is called *Bjorken scaling* and it is violated while adding higher orders in perturbation theory. The relation between F_1 and F_2 is the *Callan-Gross formula* and it says that the spin of the charged partons is 1/2 (otherwise for a spin-0 particle we would find $F_1/F_2 = 0$) [17].

In figure 1.6 different examples of models of the proton are compared. The measurements of F_2 show that the correct model is the last one. Indeed, it is observed that the proton contains not only the valence quarks that give the structure uud of the proton but also the sea quarks, couples $q\bar{q}$ that come from the vacuum. Moreover, it is observed that if we sum over all the partons momenta we do not reconstruct the whole momentum p of the proton but only $p - p_g$ where the missing momentum p_g is carried by neutral partons that will be recognized as the gluons. For the moment they are not identified as the carriers of the strong force but they will

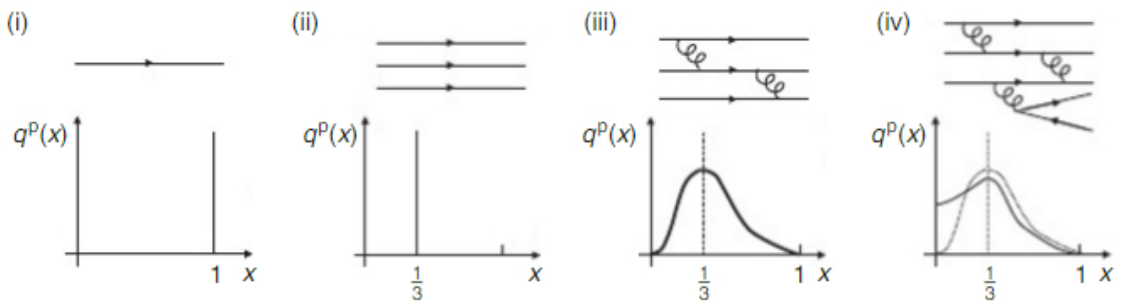


Figure 1.6: Structure function for different compositions of the proton. The last one describes well the observations.

be in QCD where it is known that there is a finite probability to emit gluons that can also create a couple $q\bar{q}$ from the vacuum (sea). These sea quarks dominate for $x \rightarrow 0$, when the probability to emit a gluon is higher.

1.2.2 Quantum Chromodynamics

The theory of strong interaction is extremely difficult and it is developed in comparison to the simpler electromagnetic interaction, outlined in Quantum Electrodynamics.

The strong force is described by the Quantum Chromodynamics (QCD) with its Lagrangian:

$$\mathcal{L}_{QCD} = -\frac{1}{4}F_{\mu\nu}^a F_a^{\mu\nu} + \sum_k [i\bar{\psi}_k \gamma^\mu \delta_\mu \psi + g_S \bar{\psi}_k T_a G_\mu^a \psi]. \quad (1.2)$$

The strength of the strong field is defined as

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a + g_S f^{abc} G_\mu^b G_\nu^c,$$

and $g_S = \sqrt{4\pi\alpha_S}$ is the strong coupling.

The states of strong interaction are characterized by the *color charge*: conventionally “red”, “green” and “blue”. It is the corresponding to the electric charge and it is introduced to assure that the quarks inside hadrons satisfy Pauli exclusion principle. All the hadrons that we find in nature are neutral for color charge and they are said to be “colorless” [1].

The QCD is a non abelian gauge theory invariant under $SU(3)_C$ and its generators are $T^a = \frac{1}{2}\lambda^a$. The λ^a are the Gell-Mann matrices, such that $[\lambda^\alpha, \lambda^\beta] = 2if^{\alpha\beta\gamma}\lambda^\gamma$, where $f^{\alpha\beta\gamma}$ are the completely antisymmetric structure functions. These matrices are used to compute the color factors and they are associated to the strong octet of gluon fields $G_\mu^{1,\dots,8}$.

The gluons, corresponding to the neutral particles found in the parton model, are the gauge bosons of the strong interaction. They are, for the strong force, the counterpart of the photons for the electromagnetic force. But unlike the photons that do not carry electric charge, the gluons carry color charge, hence, we can find vertices of multiple gluons.

Each of the eight gluons carries a color and an anticolor. On the other hand, the quarks are described by the Dirac spinor $u^s(p)$ multiplied by the vector column of the color c , by convention:

$$c = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \text{ for red, } \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ for blue, } \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \text{ for green.}$$

To compute the cross sections at leading order all the interfering Feynman diagrams are drawn and their amplitudes are evaluated applying QCD Feynman rule, similar to those of QED. Then $|M|^2 = |\sum_i M_i|^2$ is calculated and the phase space factors of the produced protons and the 4-momentum conservation are attached. At next-to-leading order ($\mathcal{O}(\alpha_S^3)$) 1-loop corrections and gluon emission must be accounted since they introduce logarithmic singularities.

The strong coupling constant α_S for $|q|^2 \gg \mu^2$ runs as

$$\alpha_S(|q|^2) = \frac{\alpha_S(\mu^2)}{1 + \frac{\alpha_S(\mu^2)}{12\pi}(11n - 2f) \ln\left(\frac{|q|^2}{\mu^2}\right)} \quad (1.3)$$

due to an effect of color screening that comes out in quark-antiquark bubbles. In the formula above, n is the number of colors (3 in the SM) and f the number of flavors (6 in the SM). The coupling decreases at high energy. In fact, for $q^2 \gg 1 \text{ GeV}$ (hence $r \rightarrow 0$) we are in the so called regime of **asymptotic freedom**, where the quarks behave like free particles. In this energy range perturbative QCD is used.

On the other hand, for $r \rightarrow \infty$ we are in the **confinement** regime, where the quarks, while pulled apart, produce $q\bar{q}$ couples that group together to form hadrons. Indeed, we do not observe isolated quarks. In these energy region lattice QCD is needed.

These conditions reflect the trend of the strong potential of the kind

$$V(r) = -k \frac{\alpha_S}{r} + k' r. \quad (1.4)$$

This behavior assures also that the assumption of free particles in the parton model (that works well at high q^2) is correct.

Equation 1.3 can also be put in terms of the scale of the strong interaction $\Lambda_{QCD} \sim 200 \text{ MeV}$ such that $\ln \Lambda^2 = \ln(\mu^2) - \frac{12\pi}{(11n-2f)\alpha_s(\mu^2)}$:

$$\alpha_S(|q|^2) = \frac{12\pi}{(11n-2f) \ln \left(\frac{|q|^2}{\Lambda^2} \right)}. \quad (1.5)$$

1.2.3 Hadron collisions and Drell-Yan process

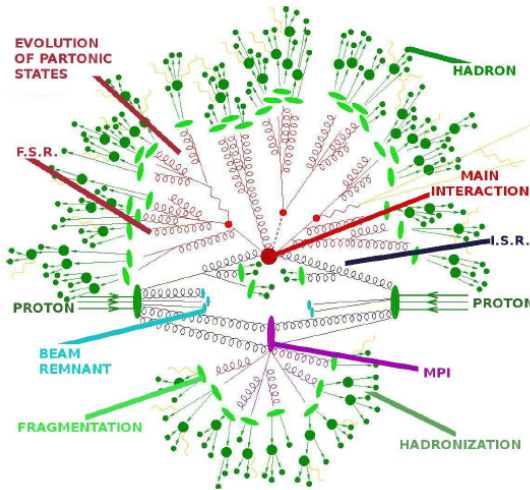


Figure 1.7: Representation of the proton-proton collision: hard process, initial and final state radiation, multiple parton interaction and hadronization.

When we have to deal with hadronic collisions we begin from the hard process: parton scattering characterized by high momentum transfer that causes production of particles with large transverse momentum p_T . The next steps are the parton shower formation and the hadronization, hence the creation of jets: cones of particles along the direction of the initial quark. What we detect in the end is the product of the decay of those particles.

To describe a realistic initial state the PDFs of its components are used. Then to have a reasonable final state the several partons that come out of the main interaction and their hadronization into jets should be considered. In addition, the quarks can be subject to gluon emission (both real and virtual): initial state radiation and final state radiation.

Furthermore, the quarks that collide in the main interaction are only part of the constituents of the proton, there are also the beam remnants that can in turn interact and give rise to secondary parton showers and jets: multiple parton interaction. These underlying events are on the other hand characterized by low transverse momentum so it is possible to separate them from the hard process of interest. Figure 1.7 pictures the evolution of showers from a pp interaction.

As already mentioned, when we study the interaction beyond the Born-level (leading order LO), at next-to-leading order (NLO) there are virtual loops and emission of gluons in the final and initial states that introduce infinities and make the cross section not-computable [18]. The real emission of the gluons that are observed gives rise to infrared (IR) divergences in the cross-section: soft divergence for emission of gluon with very small energy and collinear divergence for gluons emitted at very small angles with respect to the other parton. The emission of gluons that are not observed introduces virtual ultraviolet (UV) corrections. Due to the UV divergences at short-time, the coupling constant has to be redefined in order to stay finite. For the Kinoshita-Lee-Nauenberg theorem [19], the total NLO inclusive cross-section

$$\sigma_{inclusive}^{NLO} \propto |\mathcal{M}_{tot}|^2 = |\mathcal{M}_{Born} + \mathcal{M}_{virtual}|^2 + |\mathcal{M}_{real}|^2$$

is finite while summing virtual and real corrections in the final state that cancel each other.

The diagrams of some QCD corrections that respect this theorem are represented in figure 1.8.

On the other side, the exclusive observables, as the number of particles in a jet or the momen-

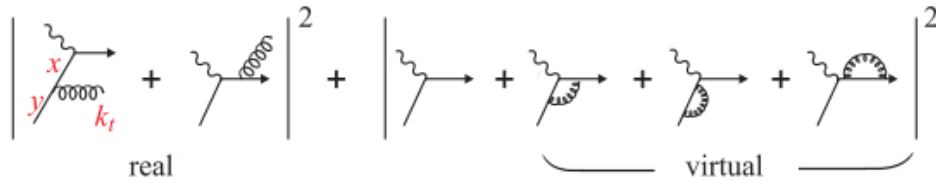


Figure 1.8: QCD corrections.

tum of a quark, are not independent on large-distance physics, hence they are not infrared safe. This makes the PDF infinite and the non-inclusive cross section divergent.

These large-distance and short-time effects must be accounted and taken under control while computing physical cross sections by introducing the proper arbitrary cut-offs of the order of the hard scale characterizing the process: factorization and renormalization scales [20].

The evolution of the PDFs is controlled by the probability P_{qq} that a quark with momentum fraction x comes from a parent quark with larger momentum y that lost part of its energy through gluon emission and the probability P_{qg} that a gluon of momentum y annihilated into a couple of quarks with momentum fraction x :

$$\frac{dq_i(x, Q^2)}{d \log Q^2} = \frac{\alpha_S}{2\pi} \int_x^1 \frac{dy}{y} \left(q_i(y, Q^2) P_{qq}\left(\frac{x}{y}\right) + g(y, Q^2) P_{qg}\left(\frac{x}{y}\right) \right) \quad (1.6)$$

Similarly for the evolution of gluon density:

$$\frac{dg(x, Q^2)}{d \log Q^2} = \frac{\alpha_S}{2\pi} \int_x^1 \frac{dy}{y} \left(\sum_i q_i(y, Q^2) P_{gq}\left(\frac{x}{y}\right) + g(y, Q^2) P_{gg}\left(\frac{x}{y}\right) \right). \quad (1.7)$$

The 1.6 and 1.7 are called Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [21] and they are used to move from a scale to another one.

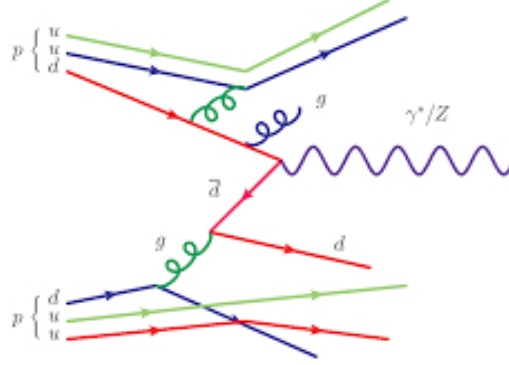


Figure 1.9: Drell-Yan scattering.

An example of hadron collision is the Drell-Yan process, consisting in the scattering of a quark of an hadron with the antiquark of another hadron that annihilate into a virtual photon or a Z boson. It is a hard process where the quarks behave like free particles and perturbative QCD can be used. In picture 1.9 the scheme of a Drell-Yan process in a proton-proton collision is shown.

The cross section σ of the pp interaction can be computed from the cross section at parton level $\hat{\sigma}_{p_1 p_2 \rightarrow C}$ of the process $p_1 p_2 \rightarrow C$:

$$\sigma = \sum_{p_1, p_2} \int \hat{\sigma}_{p_1 p_2 \rightarrow C}(\mu_F, \mu_R) f_{p_1}(x_1, \mu_F) f_{p_2}(x_2, \mu_F) dx_1 dx_2, \quad (1.8)$$

where the summation is over the partons that interact in the two protons, f_{p_i} are the PDFs, μ_F is the factorization scale and μ_R is the renormalization scale.

1.3 The Higgs boson

The Higgs boson is necessary to explain the mass of the particles and to give consistency to the Standard Model. It is known from *Goldstone theorem* [22] that if a continuous symmetry is spontaneously broken for each generator of the symmetry group a massless scalar particle appears. But there is a only exception: when the conserved currents of the group are coupled with gauge fields (gauged symmetry breaking) the Goldstone bosons corresponding to the broken generators are “absorbed” by the gauge bosons that gain mass.

In the SM the electroweak symmetry breaking [23, 8], without disrupting gauge invariance, is achieved through the **Brout-Englert-Higgs mechanism**, usually known just as Higgs mechanism.

1.3.1 Spontaneous symmetry breaking

The non-zero mass of fermions and gauge bosons is achieved by their interaction with the Higgs field, the Higgs mechanism, wherein the neutral component of an isodoublet scalar field acquires a non-zero vacuum expectation value.

It is shown in a simple scalar case how the mass terms arise from a broken symmetry. A scalar field ϕ is considered, with Lagrangian

$$\mathcal{L}(\phi) = K - V = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \left(\frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda\phi^4\right) \quad (1.9)$$

where $\lambda > 0$, the self-coupling, represents the intensity of the interaction, and the potential is symmetric under $\phi \rightarrow -\phi$.

The ground of the spectrum, state with lowest energy, is called vacuum state v and it is the minimum of the potential. If the mass term is $\mu^2 > 0$ the minimum is in $\phi = 0$, as shown in picture 1.10 a), hence the vacuum is zero and the Lagrangian represents a scalar particle with mass μ .

If $\mu^2 < 0$ there are two degenerate vacuum states in the local minima $\phi = \pm v = \pm\sqrt{\frac{-\mu^2}{\lambda}}$, as

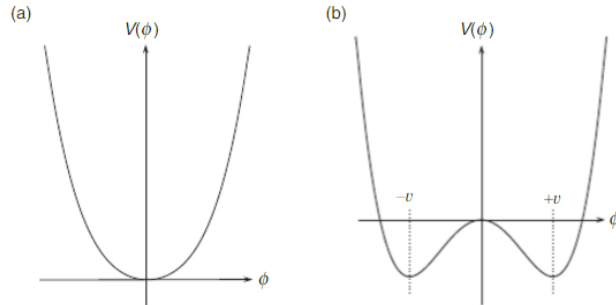


Figure 1.10: One dimensional potential $V(\phi) = \mu^2\phi^2/2 + \lambda\phi^4/4$ for $\lambda > 0$ in the cases $\mu^2 > 0$ on the left and $\mu^2 < 0$ on the right.

shown in figure 1.10 b). Now μ cannot be considered the mass of a scalar particle, and v has a non-null expectation value. The excitations (particles) are found as perturbations of the vacuum, hence expanding the field around the minimum, in this case one of the two minima. We have to choose between the two degenerate vacuum states to perform the expansion and in this way the symmetry is spontaneously broken. Taking the positive v we expand $\phi(x) = v + \eta(x)$ and we find the Lagrangian density in terms of η

$$\mathcal{L}(\eta) = \frac{1}{2}\partial_\mu\eta\partial^\mu\eta - \left(\lambda v^3\eta^2 + \lambda v\eta^3 + \frac{1}{4}\lambda\eta^4\right). \quad (1.10)$$

Here the second term gives the mass $m_\mu^2 = 2\lambda v^2 = -2\mu^2$ and as expected from the **spontaneous symmetry breaking**, the new Lagrangian density in η is not symmetric anymore under reflection, even if the descriptions in terms of $\mathcal{L}(\phi)$ and $\mathcal{L}(\eta)$ are equivalent.

The same idea can be applied to a complex scalar field $\phi = \frac{\phi_1 + i\phi_2}{\sqrt{2}}$ described by the Lagrangian

$$\mathcal{L}(\phi_1, \phi_2) = \frac{1}{2}\partial_\mu\phi_1\partial^\mu\phi_2 + \frac{1}{2}\partial_\mu\phi_1\partial^\mu\phi_2 - \frac{1}{2}\mu^2(\phi_1^2 + \phi_2^2) - \frac{1}{4}\lambda(\phi_1^2 + \phi_2^2)^2 \quad (1.11)$$

symmetric under $\phi \rightarrow e^{i\chi}\phi$. In this case, for $\mu^2 > 0$ the minimum is in $\phi_1 = \phi_2 = 0$, while for $\mu^2 < 0$ there is an infinite set of minima that lie on the circumference of radius v : $\phi_1^2 + \phi_2^2 = -\frac{\mu^2}{\lambda} = v^2$. By choosing $\phi_1, \phi_2 = (v, 0)$ we break the symmetry and from the expansion $\phi(x) = \frac{v + \eta(x) + i\rho(x)}{\sqrt{2}}$ we derive the Lagrangian density

$$\mathcal{L}(\eta, \rho) = \frac{1}{2}\partial_\mu\rho\partial^\mu\rho + \frac{1}{2}\partial_\mu\eta\partial^\mu\eta + \mu^2\eta^2 - \lambda v(\eta\rho^2 + \eta^3) - \frac{1}{2}\lambda\eta^2\rho^2 - \frac{1}{4}\lambda\eta^3 - \frac{1}{4}\lambda\rho^4 + \text{const} \quad (1.12)$$

and find $m_\eta = \sqrt{2\lambda v^2}$, while ρ is massless (it is the Goldstone boson).

1.3.2 Higgs mechanism

In the Higgs mechanism the local gauge symmetry of the Higgs complex scalar field $V(\phi) = \mu^2\phi^2 + \lambda\phi^4$ is spontaneously broken. This complex field is characterized by the peculiar “Mexican hat” shape, shown in figure 1.11. In addition to the U(1) local gauge transforma-

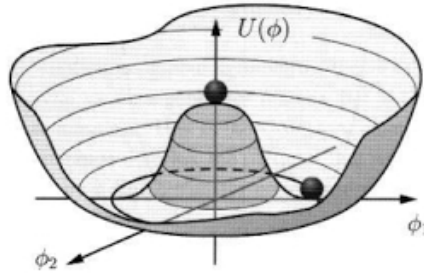


Figure 1.11: Mexican hat shaped Higgs potential.

tion $\phi(x) \rightarrow e^{ig\chi(x)}\phi(x)$, it is required to replace the derivatives with the covariant derivatives $D_\mu = \partial_\mu + igB_\mu$, where B_μ is a new gauge boson that transforms like $B_\mu \rightarrow B_\mu - \partial_\mu\chi(x)$. The combined Lagrangian for ϕ and B is

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + (D_\mu\phi)^*(D^\mu\phi) - \mu^2\phi^2 - \lambda\phi^4 \quad (1.13)$$

with $F^{\mu\nu} = \partial^\mu B^\nu - \partial^\nu B^\mu$.

As before, we break the symmetry by choosing the vacuum state $v = \phi_1 + \phi_2$ and expanding $\phi(x) = \frac{1}{\sqrt{2}}(v + \eta(x) + i\xi(x))$. Now the Lagrangian becomes

$$\mathcal{L} = \frac{1}{2}\partial_\mu\eta\partial^\mu\eta - \lambda v^2\eta^2 + \frac{1}{2}\partial_\mu\xi\partial^\mu\xi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}g^2v^2B_\mu B^\mu - V_{int}(\eta, \xi, B) + gvB_\mu\partial^\mu\xi. \quad (1.14)$$

As in the previous case, we find a scalar field η and a Goldstone boson ξ , but there is also a new term that gives mass to the gauge boson of the gauge local symmetry B , $m_B = gv$. But it seems that now there is an additional degree of freedom and moreover the spin-1 Goldstone field ξ couples to the gauge field B . To solve these problems we can make a gauge transformation $B_\mu(x) \rightarrow B_\mu + \frac{1}{gv}\partial_\mu\xi(x)$. Then we get rid of ξ in the Lagrangian choosing $\chi(x) = -\xi(x)/gv$, hence $\phi(x) \rightarrow e^{-i\xi(x)/v}\phi(x)$. This choice is called *Unitary Gauge* and when we expand $\phi(x) = \frac{1}{\sqrt{2}}(v + \eta(x)) = \frac{1}{\sqrt{2}}(v + h(x))$ and call $h(x)$ the Higgs field, the new Lagrangian density

$$\mathcal{L} = \frac{1}{2}\partial_\mu h\partial^\mu h - \lambda v^2 h^2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}g^2 v^2 B_\mu B^\mu + g^2 b B_\mu B^\mu h + \frac{1}{2}g^2 B_\mu B^\mu h^2 - \lambda v h^3 - \frac{1}{4}\lambda h^4 + \frac{1}{4}\lambda v^4 \quad (1.15)$$

shows the terms that give mass both to the gauge boson associated with the U(1) local gauge symmetry and to the scalar Higgs boson $m_H = \sqrt{2\lambda}v$ associated to the field h . There are also terms of interaction of the Higgs with B and self-interaction of the Higgs. It is important to note that λ is an unknown parameter so the mass of the Higgs is not fixed.

1.3.3 Higgs boson in the Standard Model

In the Standard Model the Higgs field [24] is introduced as a self-interacting complex doublet of scalar fields in SU(2): $\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$. It inserts in the Lagrangian a term $V(\phi) = -\mu^2(\Phi^\dagger\Phi) - \lambda(\Phi^\dagger\Phi)^2$ and if $\mu^2 < 0$, as in the previous sections, the neutral component of Φ develops a non-null vacuum expectation value (VEV): $\begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}$ where $v = \sqrt{\frac{-\mu^2}{\lambda}}$. Φ has four degrees of freedom, so we can express it in terms of four fields $\theta_{1,2,3}(x), H(x)$:

$$\Phi(x) = \begin{pmatrix} \theta_2 + i\theta_1 \\ \frac{1}{\sqrt{2}}(v + H) - i\theta_3 \end{pmatrix} = e^{i\theta_a(x)\tau^a(x)/v} \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(v + H(x)) \end{pmatrix}. \quad (1.16)$$

The choice to move to the Unitary Gauge, defined by the transformation

$$\Phi(x) \rightarrow e^{-i\theta_a(x)\tau^a(x)/v}\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix},$$

breaks the $SU(2)_L \times U(1)_Y$ symmetry into $U(1)_Q$. Three of the degrees of freedom take out of the vacuum three Goldstone bosons that mix with the gauge fields corresponding to the broken generators and become the longitudinal component of the weak gauge bosons giving them mass.

In the new Lagrangian the fields $W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iX_\mu^2)$, $Z_\mu = \frac{g_2 W_\mu^3 - g_1 B_\mu}{\sqrt{g_2^2 + g_1^2}}$ and $A_\mu = \frac{g_2 W_\mu^3 + g_1 B_\mu}{\sqrt{g_2^2 + g_1^2}}$ are defined and their masses come from the term:

$$M_W^2 X_\mu^+ X_\mu^- + \frac{1}{2}M_Z^2 Z_\mu Z^\mu + \frac{1}{2}M_A^2 A_\mu A^\mu.$$

The weak bosons acquire $M_W = \frac{1}{2}vg_2$ and $M_Z = \frac{1}{2}v\sqrt{g_2^2 + g_1^2}$, while the photon remains massless $M_A = 0$ since the $U(1)_Q$ symmetry is unbroken.

Moreover, in a similar way, the fermions couple to the Higgs boson through the Yukawa interaction $\mathcal{L}_F = -\lambda_e \bar{L}\Phi e_R - \lambda_d \bar{Q}\Phi d_R - \lambda_u \bar{Q}\tilde{\Phi} u_R + h.c.$ and gain mass $m_e = \frac{\lambda_e v}{\sqrt{2}}, m_d = \frac{\lambda_d v}{\sqrt{2}}$ and $m_u = \frac{\lambda_u v}{\sqrt{2}}$.

The fourth degree of freedom $H(x)$ represents the physical Higgs boson. Its Lagrangian term is $\mathcal{L}_H = \frac{1}{2}(\partial^\mu H)^2 - \lambda v^2 H^2 - \lambda v H^3 - \frac{1}{4}\lambda H^4$ and its mass $M_H = \sqrt{2\lambda v^2}$ is not defined because the parameter λ is not known. It is crucial to find out the mass of the Higgs from measurements.

In figure 1.12 the vertices of interaction of the bosons and fermions with the Higgs and of the Higgs with itself are shown.

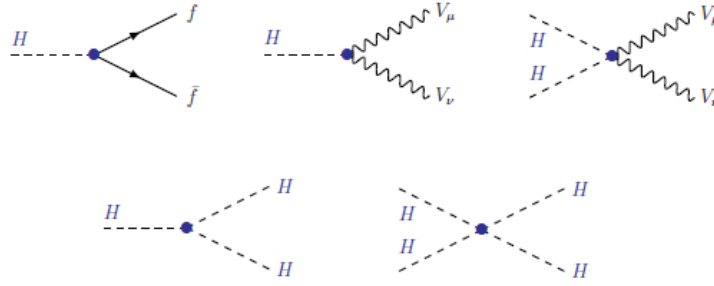


Figure 1.12: Feynman diagrams of the coupling between bosons of fermions and the Higgs and vertices of self-interaction of the Higgs boson.

The couplings are given by

$$g_{Hf\bar{f}} = \frac{m_f}{v}, g_{HVV} = \frac{2m_V^2}{v}, g_{HHVV} = \frac{2m_V^2}{v^2},$$

$$g_{HHH} = \frac{3m_H^2}{v}, g_{HHHH} = \frac{3m_H^2}{v^2}.$$

Finally, after the introduction of the Higgs boson in the theory, the electroweak Lagrangian is expressed as:

$$\begin{aligned} \mathcal{L}_F = & \sum_i \bar{\psi}_i \left(i\not{\partial} - m_i - \frac{m_i H}{v} \right) \psi_i + && \text{Yukawa coupling of the Higgs to the fermions} \\ & - \frac{g}{2\sqrt{2}} \sum_i \bar{\psi}_i \gamma^\mu (\gamma^5) (T^+ W_\mu^+ + T^- + W_\mu^-) \psi_i + && \text{charged current weak interaction} \\ & - e \sum_i Q_i \bar{\psi}_i \gamma^\mu \psi_i A_\mu + && \text{QED} \\ & - \frac{g}{2 \cos \theta_W} \sum_i \bar{\psi}_i \gamma^\mu (g_V^i - g_A^i \gamma^5) \psi_i Z_\mu. && \text{neutral current weak interaction} \end{aligned} \quad (1.17)$$

where $\theta_W = \tan^{-1}(g'/g)$ is Weinberg (or weak mixing) angle introduced in the previous section, A_μ is the photon field, Z_μ and $W_\mu^\pm$ are the weak boson fields, v is the vacuum expectation value and $T^+(T^-)$ is the isospin raising (lowering) operator.

The bosons masses are accessed by

$$M_H = \lambda v$$

$$M_W = \frac{1}{2}gv = \frac{ev}{2\sin\theta_W}$$

$$M_Z = \frac{1}{2}\sqrt{g^s + g'^2}v = \frac{ev}{s\sin\theta_W\cos\theta_W} = \frac{M_W}{\cos\theta_W}$$

$$M_\gamma = 0.$$

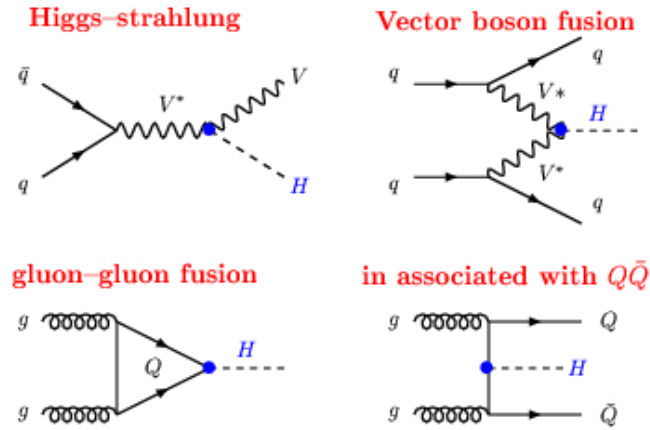


Figure 1.13: Feynman diagrams of the mechanism of production of the Higgs boson in pp collisions.

1.3.4 Higgs boson production and decay channels

The Higgs boson can be produced in many ways in hadron collisions [24, 25]. The Feynman diagrams of the four main channels are drawn in figure 1.13 and their cross sections at LHC at 7 TeV are plot in figure 1.14.

The gluon gluon fusion **ggH** [26, 27] is the dominant mode of production of the Higgs boson (with mass below 1 TeV) at proton colliders. It is mediated by a top quark triangular loop. The contribution to the cross section from bottom quark loop is less than 10%, since the coupling depends on the mass of the quarks. In this channel the QCD corrections are not negligible and they are computed by means of the K-factor, the ratio between the cross sections at an higher order and the lowest order. Since all the quarks except the top are much lighter than the Higgs boson their masses are often considered negligible in the computations of the cross section. The cross section of this channel decreases with m_H as the PDFs of gluons decrease

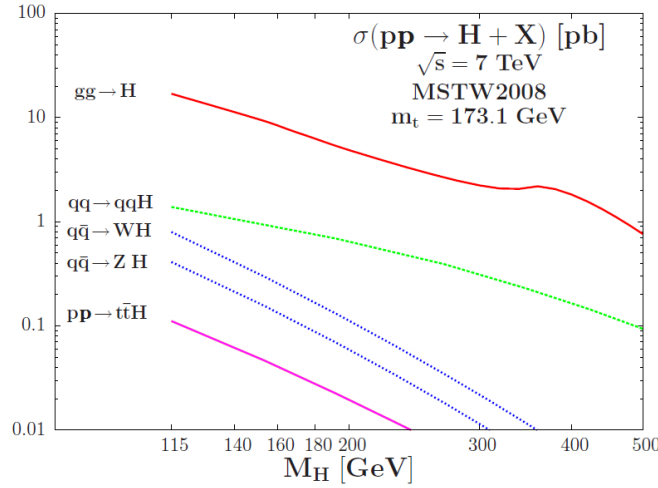


Figure 1.14: Cross section of the different mechanisms of production of the Higgs boson at LHC at 7 TeV.

with increasing Bjorken x variable.

Through the interaction of two gluons, but without a quark loop, the Higgs boson can be produced also in association with a couple of top quarks ttH [28]. In this channel it is possible to observe $H \rightarrow b\bar{b}$ decays and to measure directly the top-quark Yukawa coupling.

Another associated production channel is the Higgsstrahlung VH [29, 30], where the Higgs boson comes out with a vector boson $V = W, Z$. This channel opens for $\sqrt{s} \geq m_V + m_H$. This mechanism has a very low production cross section since an anti-quark from the proton sea is involved. On the other hand, this mode is very clean when the Higgs decays leptonically: it is simple to reduce the QCD background requiring a high-transverse-momentum lepton.

The Higgs is often produced in association with heavy particles because its coupling is proportional to the mass of fermions and the mass squared of bosons.

The Vector Boson Fusion VBF [31, 32] is the second strongest production mechanism and becomes dominant in the theories of very massive Higgs bosons. It has a clear signature: two almost back-to-back hard tagging jets are expected with the products of the Higgs decay confined at low rapidity.

The selection of events from the different modes by means of their peculiar final states is described in section 3.8, in the diphoton decay channel.

The main decay channels of the Higgs boson [33], shown in figure 1.15, are listed below.

The $H \rightarrow \gamma\gamma$ [34] is the clearest channel, despite the low branching ratio (it is suppressed due to the higher order contribution of the loop). The background is estimated from data in $100 < m_{\gamma\gamma} < 180$ GeV. The irreducible background comes from QCD production of diphotons (prompt photons) while the reducible background is due to γ +jet or jet+jet events with jets misidentified as photons. A more detailed description of the analysis in this channel is given

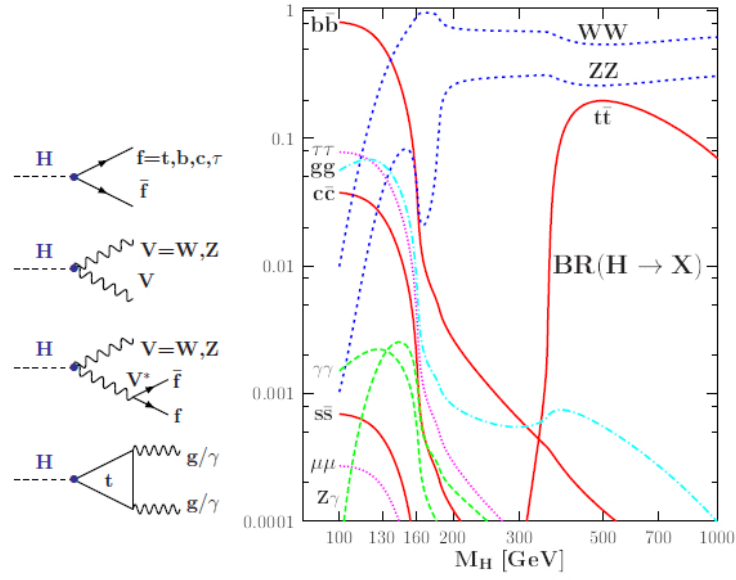


Figure 1.15: Decays of the Higgs boson. Feynman diagrams on the left. Plot of the Branching ratios for the different channels on the right.

in chapter 3.

The $H \rightarrow ZZ^{(*)}$ [35] has small continuum background that grants excellent momentum resolution. Two pairs of leptons of same flavor are needed, the one with invariant mass closer to m_Z must have $40 < m_{\ell\ell} < 120$ GeV and the other $12 < m_{\ell\ell} < 120$ GeV. The events are split in three subchannels: $4e$, 4μ and $2e2\mu$. The dominating irreducible background comes from the production of two Z in $q\bar{q}$ (Drell-Yan) and gg processes, while the reducible background from Z +jets or $t\bar{t}$ production with misidentified jets or leptons from decay of heavy quarks. Seven kinematic variable with strong discriminating power are the five angles and the two invariant masses of the leptonic pairs that describe the process in its center of mass frame. Furthermore, to select the signal the two leptons of each pair are required to appear from the same vertex. Electrons are selected in $|\eta| < 2.5$ with $p_T > 7$ GeV, and muons in $|\eta| < 2.4$ with $p_T > 5$ GeV. The final state radiation from leptons is recovered: photons are collected in $|\eta| < 2.4$ with $\Delta R < 0.07$ from a muon and $p_T^\gamma > 2$ GeV or $0.07 < \Delta R < 0.5$ and $p_T^\gamma < 4$ GeV, isolated in $\Delta R = 0.3$.

The $H \rightarrow WW^{(*)} \rightarrow 4\ell$ [36] shows as signature of two W bosons (at least one off-shell) two isolated leptons with opposite charge and large missing energy (from the neutrinos matched in flavor with their leptons). The presence of the neutrinos makes the analysis difficult, hence the mass resolution poor. The events are classified in categories based on the number of jets (0, 1 or 2) and the flavors (ee , $\mu\mu$ and $e\mu$). Events with more than two jets are rejected. The best sensitivity is expected in the subchannel $e\mu$ with 0 jets where the irreducible background is given by non resonant WW production and the reducible by misidentified jets in W +jets.

Subchannels $e\mu$ with 1 jet and ee or $\mu\mu$ without jets contribute with 10% to the sensitivity and they have large background from top decay and Drell-Yan production. The subchannels with two jets have the highest signal to background ratio (S/B) but lower cross section hence they give very poor contribution to the sensitivity. The leading lepton is required to have $p_T^{\ell_1} > 20$ GeV and the subleading $p_T^{\ell_2} > 10$ GeV. A cut on $m_{\ell\ell} > 12$ GeV is used to select the signal above the background from low-mass resonance. The direction of the missing energy is used to reject Drell-Yan events where it is aligned with the transverse momentum of the lepton. Also the events with dilepton mass closer than 15 GeV to m_Z are rejected. The t quark background is selected using “top tagging” techniques.

The $H \rightarrow \tau^+\tau^-$ [37] channel gives access to the coupling of the Higgs to τ . Different subchannels can be identified depending on the decay of the tau: $e\mu$, $\mu\mu$, $e\tau_h$ and $\mu\tau_h$, where τ_h refers to hadronic decays. These events are split in categories based on number of jets and momentum of the τ . The high- p_T class contains events with $p_T^{\tau_h} > 40$ GeV for the hadronic decay and $p_T^\mu > 35(30)$ GeV for $e\mu(\mu\mu)$. The main sources of background in the totally leptonic subchannels are Drell-Yan processes and $t\bar{t}$ production, while in the hadronic subchannels they are W or Z plus misidentified τ_h candidate. Another important (reducible) background comes from $Z \rightarrow \tau\tau$, W +jets and QCD multijet production. This leptonic channel is dominant in comparison to $H \rightarrow \mu\mu$ because at leading-order the coupling of the Higgs with the fermion goes like m_f^2 .

The $H \rightarrow b\bar{b}$ [38], that gives direct access to the coupling of the Higgs with the bottom, is the channel with the highest branching ratio at $m_H < 135$ GeV, but what makes difficult to observe the Higgs is that the signal is overbeared by QCD production of bottom quarks $p\bar{p} \rightarrow b\bar{b} + X$. To suppress this background the Higgs is searched for in events with a high- p_T resonance associated to a Z or W boson and by putting an upper limit to the additional activity. The $b\bar{b}$ production has been studied in five independent channels: $Z(\ell\ell)H$, $Z(\nu\nu)H$, $W(\ell\nu)H$ which in turn are divided in p_T classes. The two jets must be reconstructed as b-jets and a cut is set at 120 GeV for $W(\ell\nu)$ and at 160 GeV for $Z(\nu\nu)H$. Other backgrounds sources are the production of vector bosons with jets, t quarks and WW or ZZ or WZ with one of the bosons that decays hadronically. The $b\bar{b}$ mode is dominant with respect to the $c\bar{c}$ due to the increase of the coupling with the square of the mass.

The signal strength [39] of the $i \rightarrow H \rightarrow f$ mode, $\mu_i^f = \mu_i \cdot \mu^f$, is the ratio between the measured Higgs boson rate and its theoretical prediction in SM. Given the cross section σ_i of the $i \rightarrow H$ production channel and the branching fraction B^f of the $H \rightarrow f$ decay mode, the strengths are experimentally measured:

$$\mu_i = \frac{\sigma_i}{(\sigma_i)_{SM}} \text{ and } \mu^f = \frac{B^f}{(B^f)_{SM}}.$$

1.3.5 Higgs boson discovery in 2012

In 2012 both the experiments CMS [33] and ATLAS [40] at LHC observed an excess of events compatible with the hypothesis of SM Higgs boson.

Data from the pp collisions at 7 and 8 TeV, with an integrated luminosity 5.1 and 5.3 fb⁻¹ respectively, have been analyzed to search for the Higgs boson in the five main channels: $\gamma\gamma$, $ZZ \rightarrow 4\ell$, $W^+W^- \rightarrow 2\ell 2\nu$, $\tau^+\tau^-$ and $b\bar{b}$.

A lower and an upper limits were set on the Higgs mass at a 95% confidence level by previous measurements at LEP (Large Electron-Positron collider) and precision electroweak measurements: $114.4 < m_H < 152$ GeV. Thanks to the analysis of the dataset at 7 TeV other intervals were excluded: 127-600 GeV by CMS, 111.4-116.6 GeV, 119.4-122.1 GeV and 129.2-541 GeV by ATLAS. Moreover, both the LHC experiments had already observed an excess around 125 GeV but far below the expected significance.

The LHC upgrade at 8 TeV enhanced the sensitivity of the search. The analysis of the complete data set was performed in the range 110-160 GeV. The cross section of the Higgs production was estimated to vary between 23(29) and 10(14) pb at 7(8) TeV.

The significance of the observation can be evaluated by means of the local p-value, the probability for a background fluctuation to be at least as large as the observed maximum excess, and the global p-value for an excess everywhere. The latter being larger for the look elsewhere effect (for multiple tests) [41].

The peak around 125 GeV, whose width is dominated by instrumental mass resolution, was observed in the $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ \rightarrow 4\ell$ and $H \rightarrow WW \rightarrow 2\ell 2\nu$ channels. In figure 1.16 the peak over the background in the distribution of the invariant mass of the diphoton system is shown for ATLAS (left) and CMS (right). In figure 1.16a each event is weighted by the $S/(S+B)$ value of its category, being S the signal and B the background.

The combined analysis by CMS of $\gamma\gamma$ and ZZ channels, whose p-values are shown in figure 1.17a, resulted in the observation in 115-130(110-145) GeV of an excess with significance 4.6(4.5) σ . This peak denotes a particle with mass around 125 GeV that decays in ZZ and $\gamma\gamma$ and must then be a spin-1 boson, hence compatible with the SM Higgs hypothesis. The combined best-fit mass of this new particle, confidently recognized as the Higgs boson, is $m_H = 125.3 \pm 0.4(stat.) \pm 0.5(syst.)$ GeV. The discovery is very reliable since the same excess was observed by ATLAS in 122-131(110-600) GeV with significance 5.9(5.1) σ . The p-value of the combined analysis of the channels $H \rightarrow ZZ \rightarrow 4\ell$, $H \rightarrow \gamma\gamma$ and $H \rightarrow WW \rightarrow e\nu\mu\nu$ is shown in figure 1.17b. The fit of the invariant mass gave $m_H = 126.0 \pm 0.4(stat.) \pm 0.4(syst.)$ GeV.

The combined [42] result of CMS and ATLAS in the decay channels $H \rightarrow \gamma\gamma$ and $H \rightarrow 4\ell$ is $m_H = 125.09 \pm 0.21(stat.) \pm 0.11(syst.)$ GeV.

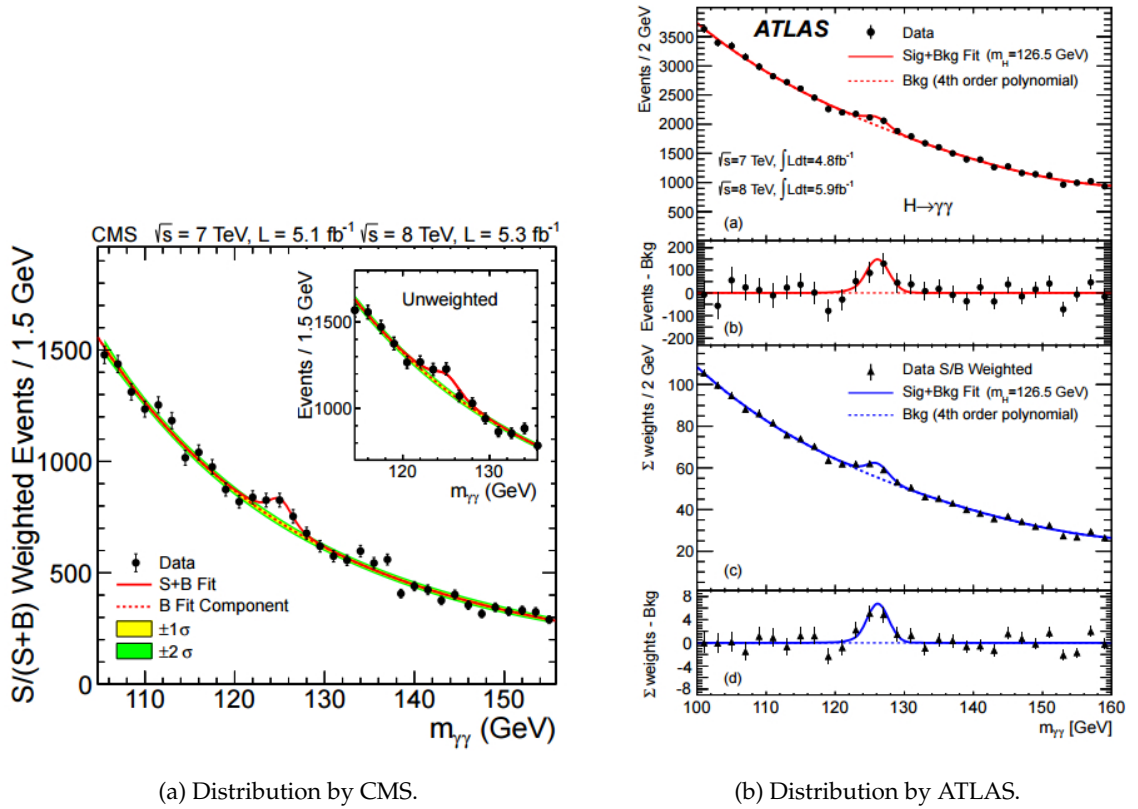
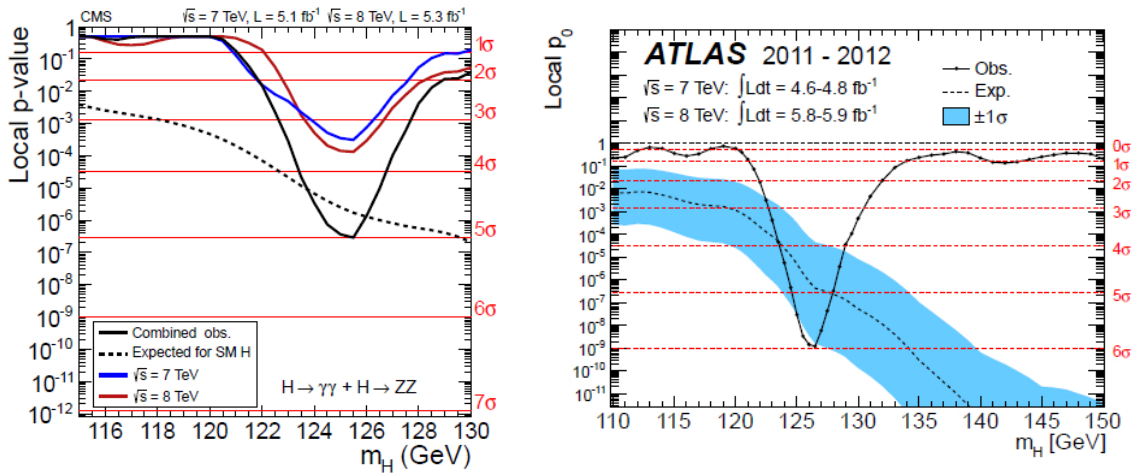


Figure 1.16: Detection of the Higgs boson over the background in the diphoton channel.



(a) Combined $\gamma\gamma$ and $ZZ \rightarrow 4\ell$ channels at CMS. (b) Combined $\gamma\gamma$, $ZZ \rightarrow 4\ell$ and $WW \rightarrow e\nu\mu\nu$ channels at ATLAS.

Figure 1.17: Observed local p-value of the combined analysis in the most sensitive decay channels of the Higgs boson.

Chapter 2

The LHC and CMS

The Large Hadron Collider (LHC) is the largest and most powerful particle accelerator ever realized. It was built by the European Organization for Nuclear Research (CERN) between 1998 and 2008 and it is located near Geneva on the border between Switzerland and France. Throughout its 26.7 km of circumference four particle detectors are hosted: the Compact Muon Solenoid (CMS) [43], A Toroidal LHC ApparatuS (ATLAS) [44], A Large Ion Collider Experiment (ALICE) [45] and the Large Hadron Collider beauty (LHCb) experiment [46]. Figure 2.1 displays their position along the ring.

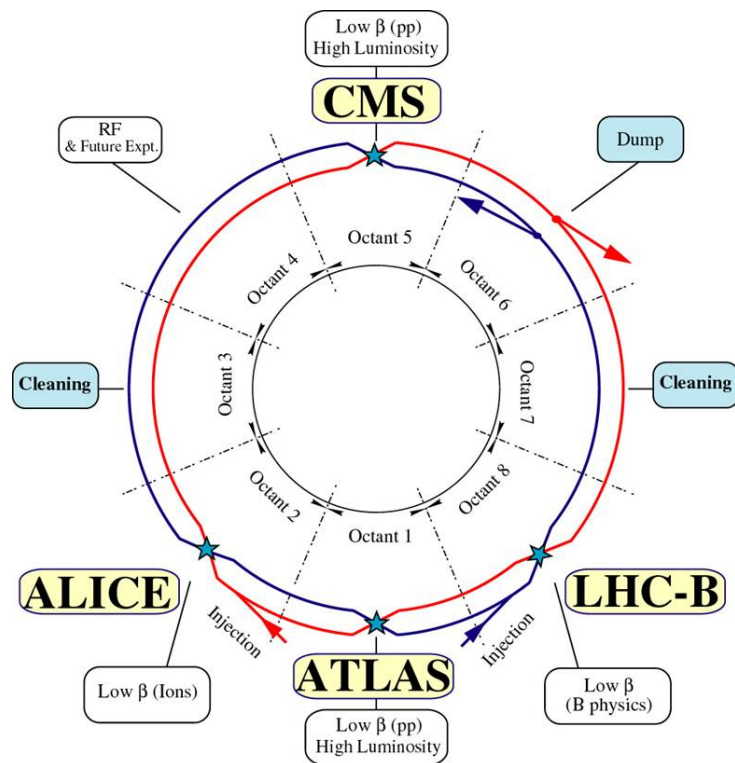


Figure 2.1: Position of the four experiments CMS, ATLAS, ALICE and LHCb along the ring of LHC.

2.1 The Large Hadron Collider (LHC): overview

The LHC [47, 48] project was approved in 1994 and the first conceptual design was delivered in 1995. It was built in the site of LEP (Large Electron Positron collider) that was stopped in 2000 and still is the most powerful electron-positron collider ever realized.

The lattice [49] of LHC reflects the one of LEP. It is composed by 8 bending arcs interleaved with 8 insertions (IR). Each arc hosts a lattice of regular half-cells made of 3 dipoles (for bending) and 1 quadrupole (for focusing). Sextupoles are used for corrections. The magnets are kept at 1.9 K by the cryogenic system.

At the intersections between the arcs two dispersion suppressors (DS) are installed, one at each end, to reduce the horizontal dispersion of the beam. Between the DS a straight section is aimed to locate experimental or infrastructure spots (P1-P8) [50].

The interaction points (IP), which were 8 at LEP, have been reduced to 4, located in the points P1, P2, P5 and P8, to avoid unnecessary disruptions of the beam that might drop the luminosity of the experiments. P2 and P8 host ALICE and LHCb in the pre-existing caverns, while new caverns have been digged in the opposite points P1 and P5 for ATLAS and CMS. In P2 and P8 there are also the injection systems for Beam1 (clockwise) and Beam2 (anticlockwise) respectively. P3 and P7 contain 2 collimators each. P4 hosts two RF (radiofrequency) cavities (systems for accelerating and bunching around the ideal in-time particle) that work independently for the two beams at 400.8 Hz. In P3 the beam dump insertion (in the region of extraction of the beam) is located.

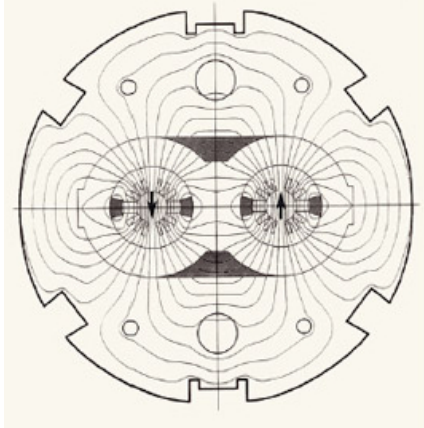


Figure 2.2: Layout of the LHC magnet design.

The two beams travel at a distance of 194 mm in two ultrahigh vacuum rings compressed in a tunnel with a 3.7 m inner radius, hence, due to spacial limitations, the twin-bore (or two-in-one) magnet design [51] was chosen (see figure 2.2), where the magnets are separate but the cryogenic system and the support structure are in common. One of the most important parameters of LHC, beside the center of mass energy, is the luminosity. The instantaneous luminosity, measured in $\text{cm}^{-2}\text{s}^{-1}$, for round beams is defined as

$$\mathcal{L} = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta}, \quad (2.1)$$

where N_b is the number of particles per bunch, n_b the number of bunches per beam, f_{rev} the frequency of revolution (hence collision), γ_r the relativistic factor, $\epsilon_n = \gamma_r \epsilon$ the normalized emittance (that quantifies the average spread of the beam particles in the phase space) and β the beta function (related to the transverse size of the particle beam) at the collision point. To quan-

tify the dimension of the beam the transverse beam size $\sigma = \sqrt{\epsilon\beta}$ is used. The luminosity is linked to the rate of generated events for a process with cross section σ :

$$N_{events} = \mathcal{L}\sigma.$$

ATLAS and CMS are general purpose detectors aimed to work at high luminosity, about $\mathcal{L} = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$, while, due to limitations of the bandwidth of the system of data acquisition in presence of a large amount of events, LHCb, specialized in B-physics, needs only $\mathcal{L} = 10^{32} \text{ cm}^{-2}\text{s}^{-1}$ and ALICE, a dedicated heavy-ion detector, works at low luminosity about $2 \cdot 10^{27} \text{ cm}^{-2}\text{s}^{-1}$.

In table 2.1 the main LHC parameters are summarized [52] .

Parameter	protons	heavy ions
center-of-mass energy [TeV]	13	2.8/n
field in bending magnet [T]	8.4	8.4
luminosity per collision region [$\text{cm}^{-2}\text{s}^{-1}$]	10^{34}	$2 \cdot 10^{27}$
number of collision regions	2	1
number of bunches per beam	2835	608
bunch spacing [ns (m)]	25 (7.5)	124.8 (37.4)
number of particles per bunch	10^{11}	$9.4 \cdot 10^7$
β parameter at interaction point [m]	0.5	0.5
RMS beam radius at collision point [μm]	16	15
beam crossing angle [μrad]	200	≤ 100

Table 2.1: Main LHC parameters.

Due to the short bunch spacing, the ideal head-on crossing would introduce a large number of parasitic collisions that can cause loss of particles, hence a little crossing angle is imposed, but there are still long-range collisions (see figure 2.3), about 15 on each side of the IP. These interactions can be minimized by alternating horizontal and vertical collisions between an IP and the subsequent. In addition, a large crossing angle would drop the luminosity by decreasing the effective area of the bunch that interact.

Other relevant characteristics of the accelerator [53] are the bunch radius $\sigma_x = \sigma_y = 16 \mu\text{m}$ (for round bunches), the bunch length $\sigma_s = 75 \mu\text{m}$ and the beam-beam tune shift parameter $\xi = 0.0034$.

The beam-beam nonlinear forces [54] are responsible for the shift of the tune, that can fall in regions of resonance and cause instability or in the worse case disruption of the beam. They are measured by $\xi = \frac{3N_b r}{2\epsilon_n}$, where r is the classical radius. The beam-beam tune shift parameter is independent on the energy (γ_r) and β .

Before entering the LHC, the beams are accelerated in several steps. The pre-existent acceleration chain (represented in figure 2.4 is used for the injection [55] into LHC of beams at 450 GeV. The injection chain for protons consists in: LINear ACcelerator 2 (LINAC2) that accelerates up

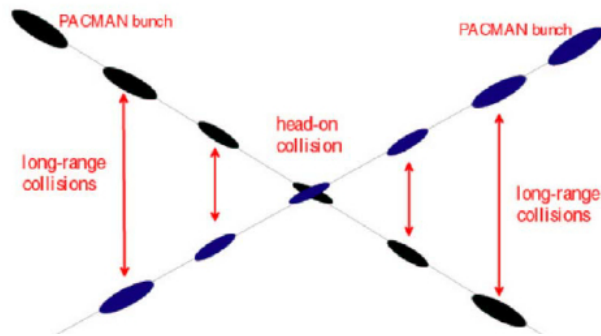


Figure 2.3: Parasitic collisions.

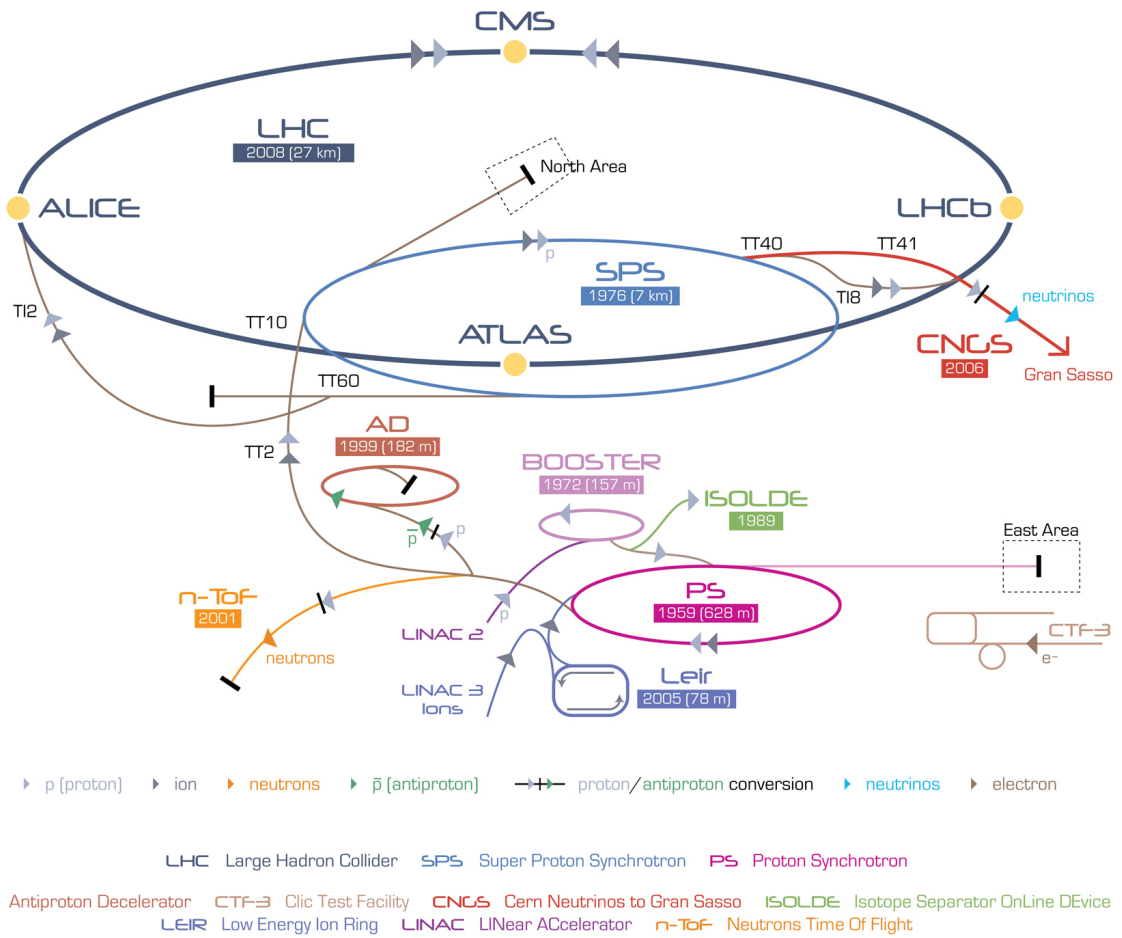


Figure 2.4: LHC injection chain.

to 50 MeV followed by Proton Synchrotron Booster (PSB) 1.4 GeV and Proton Synchrotron (PS) 25 GeV, in turn followed by Super Proton Synchrotron (SPS) 450 GeV, then the beam enters the LHC where it reaches 7 TeV.

The lead injector chain [56] is: LINear ACcelerator 3 (LINAC 3) up to 4.2 MeV/n followed by Low Energy Ion Ring (LEIR) 72 MeV/n and by PS 25 GeV/n and by SPS 450 GeV/n, and they enter the LHC where they are accelerated at 2.8 TeV/n. The LHC is filled in 12 SPS cycles, that in turn is filled in 3 or 4 PS cycles.

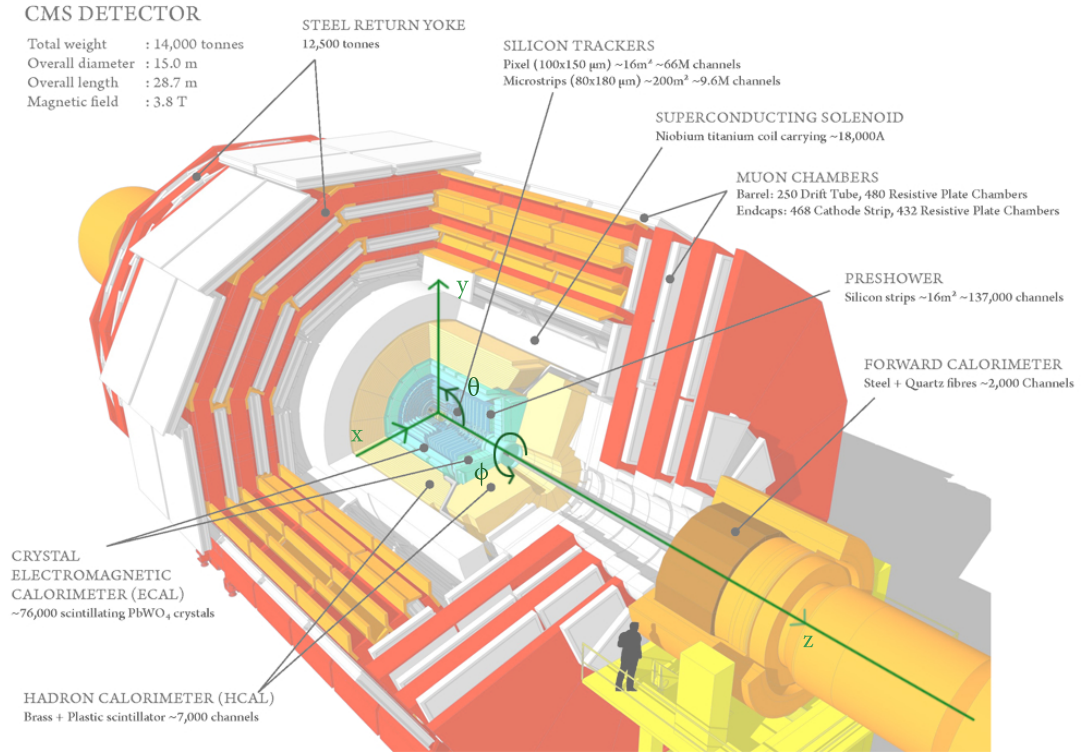


Figure 2.5: Overview of the whole CMS detector.

2.2 Compact Muon Solenoid (CMS): overview

The coordinates, depicted in figure 2.5 that represents the whole CMS with its subdetectors, have origin in the nominal interaction point. The z-axis points along the beam direction. The y-axis points vertically upward and the x-axis radially inward towards the centre of LHC. The azimuthal angle ϕ is measured from the x-axis in the x-y plane. In this plane the radial coordinate is called r . The polar angle θ is measured from the z-axis and in its place usually the pseudorapidity $\eta = \ln \tan(\theta/2)$ is employed.

Some important kinematical variables are defined as following: the transverse momentum $p_T = \sqrt{p_x^2 + p_y^2}$ equal to the transverse energy E_T in case of massless particles, the

transverse mass $m_T = \sqrt{m^2 + p_T^2}$, the rapidity $y = \frac{1}{2} \ln \frac{E+p_z}{E-p_z} = \ln(\tan(\theta/2))$ and the distance $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

The main requirements of CMS are [57]:

- large bending power for a precise measure of the momentum (section 2.2.1);
- good charged-particle momentum resolution and reconstruction efficiency (section 2.2.2);
- good electromagnetic energy resolution, good diphoton and dielectron mass resolution (section 2.3);
- good missing-transverse-energy and dijet-mass resolution (section 2.2.4);
- good muon identification, good momentum and muon-mass resolution (section 2.2.5).

2.2.1 Superconducting solenoid

The central feature of CMS is the magnet [57, 58], designed to reach a magnetic field of up to 4 T. To maximize its longevity, the actual operating field is 3.8 T.

It measures 6 m of diameter and 12.5 m of length. An overview of the device is in figure 2.6.

It is a superconducting solenoid, hence, a thin coil of superconducting wires: the 4-layers

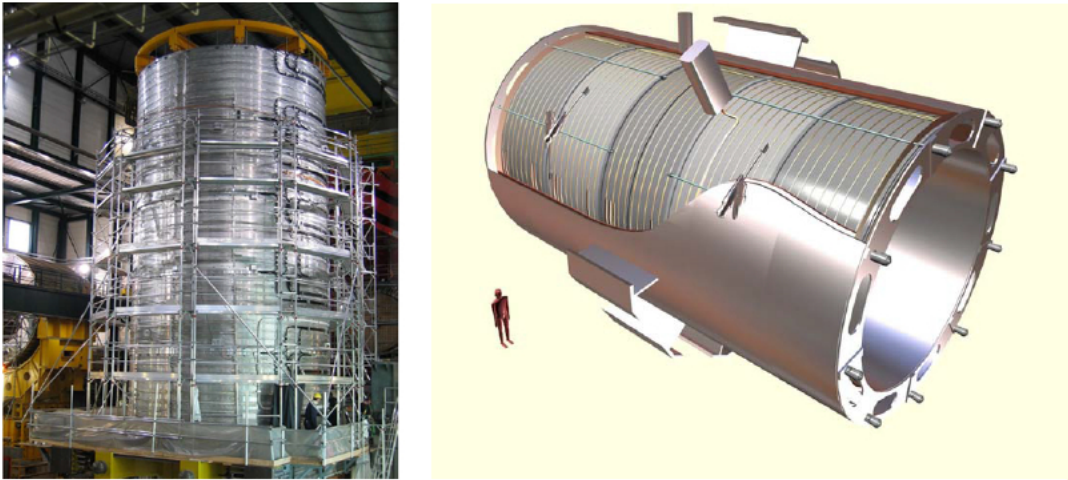


Figure 2.6: Overview of the superconducting solenoid.

winding is made of NbTi conductor. It is self-supporting since the structural material is included in the conductor.

To maintain its property of null resistance in presence of electric current the superconductor must be kept at cryogenic temperatures (below its critical temperature). For this reason the cold mass is refrigerated with liquid helium, down to 4.6 K with a margin of 1.8 T. A safety system prevents the occurrence of local overheating.

The calorimeters and tracker are housed inside the magnet while the muon chamber is interleaved with the iron return yoke made of 5 wheels and 2 endcaps, composed of 3 disks each. It is important the presence of a strong uniform magnetic field precisely mapped [59] in combination with high-precision position measurements in the tracker to assess the momentum of the particles. In fact, the bending in the field varies with the momentum as $\frac{1}{\rho} = \frac{eB}{pc}$.

2.2.2 Tracker

The main goal of the tracker [57, 60, 61] is to measure precisely the tracks of charged particles and reconstruct secondary vertices. The tracks are also used to evaluate the momentum. Immersed in the magnetic field of the solenoid, the tracker measures 5.8 m of length and 2.5 m of diameter and it is the closest to the interaction point among the subdetectors. A scheme of the tracker with its r, z, η coordinates is shown in figure 2.7.

For a proper attribution of the track to each particle fast response and high granularity are needed since several particles come out of each pp bunch crossing.

On the other hand, phenomena like multiple scattering, nuclear interaction, bremsstrahlung and photon conversion can degrade the position resolution, worsen the track assignment for hadrons and affect also the measurements in the calorimeter. Moreover, the intense particle flow can provoke radiation damage and aging, hence radiation hardness is crucial. These are some of the reasons for the choice of a silicon detector, able to work in such a harsh environment. Nevertheless, positively charged holes can be generated on the surface of the detector, in the silicon oxide layer, by ionizing particles, and non-ionizing radiation can modify the crystal lattice of the bulk, creating additional energy levels and increasing the temperature (hence the leakage current) and the depletion voltage (cause of reduction of the signal).

The operating temperature is about -10°C and the silicon sensor should never go above 0°C . The tracker is composed by an inner pixel tracker and an outer silicon strip detector to cover overall $|\eta| > 2.5$ (see figure 2.8).

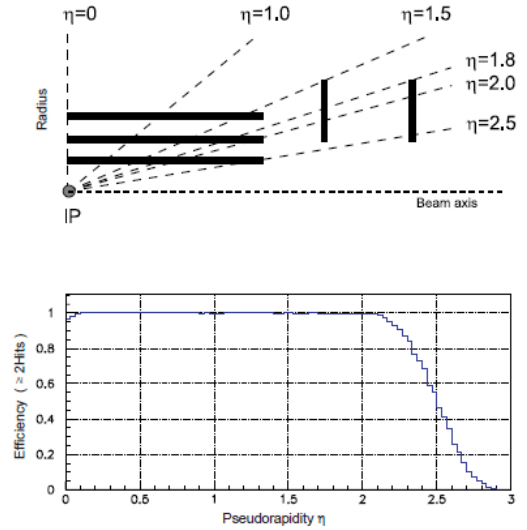


Figure 2.8: Tracker coverage.

The **pixel detector** is made of 3 barrel layers and endcaps of 2 disks for a total of 66 million

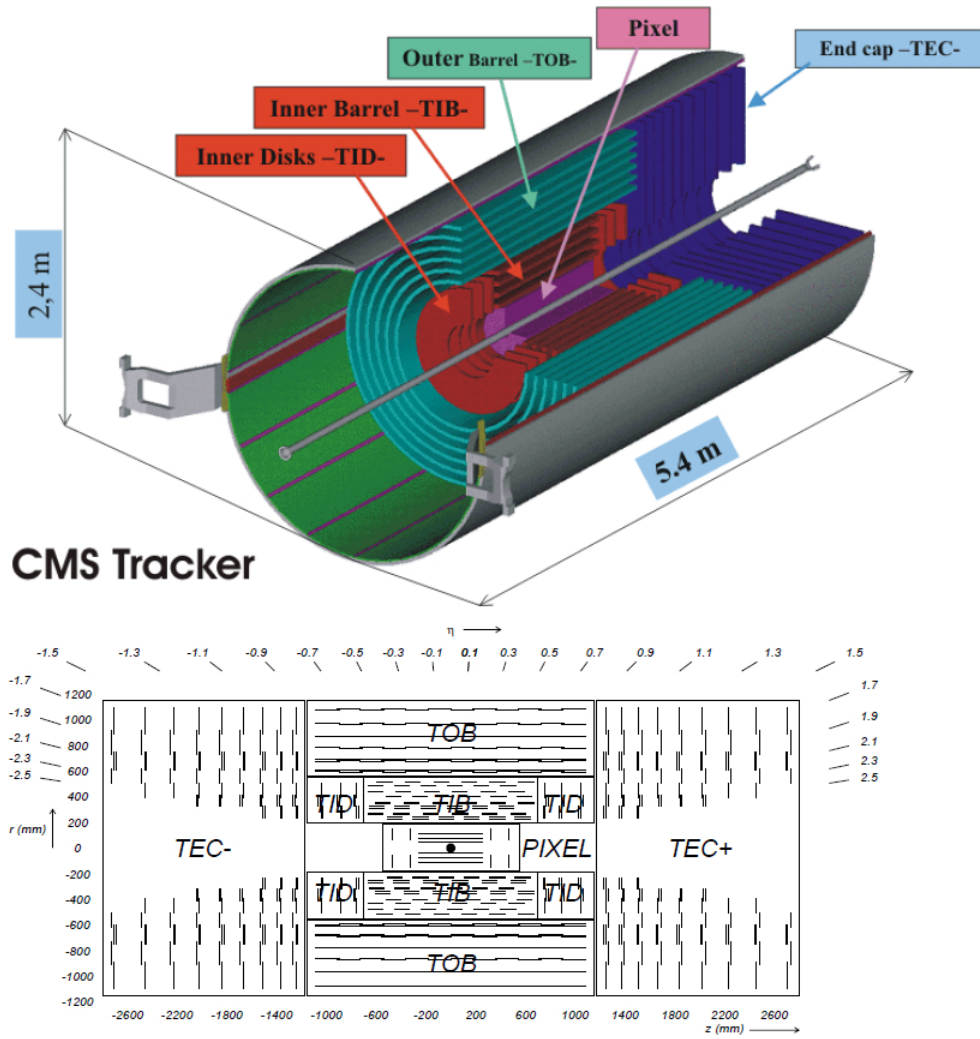


Figure 2.7: Scheme of the tracker. It shows the pixel detector, the TIB, TID, TOB and TEC and the reference coordinates r , z , η .

pixels (48 millions in barrel and 18 in endcaps) $285\ \mu\text{m}$ thick. The barrel layers are located at radii 4.4, 7.3 and 10.2 cm and provide the first three tracking points. The endcaps disks are at $z=\pm 34.5$ and ± 46.5 cm. The pixel resolution is $15\text{--}20\ \mu\text{m}$ for cells of $100\times 150\ \mu\text{m}^2$ and degrades with aging due to the strong irradiation so close to the interaction point. The design lifetime of the inner tracker is about 2 years.

The pixel detector structural components are:

- *BPix*, the barrel, 570 mm long. Figure 2.9 shows the cylindrical support structure that hosts the modules of pixel detector and supply tubes with the readout electronics on both sides. The half parts at positive and negative z are electrically separated and split in 16 independent sectors each. The basis of the support structure is made of cooling aluminium tubes 0.3 mm thick and carbon fiber blades 0.24 mm thick between adjacent tubes. At both ends the connectors for the module cables are mounted on flanges. The material of the pixel barrel amounts to 5% of a radiation length in the central region. The 50% of it due to the structural and cooling features while only one third to sensors and chips.

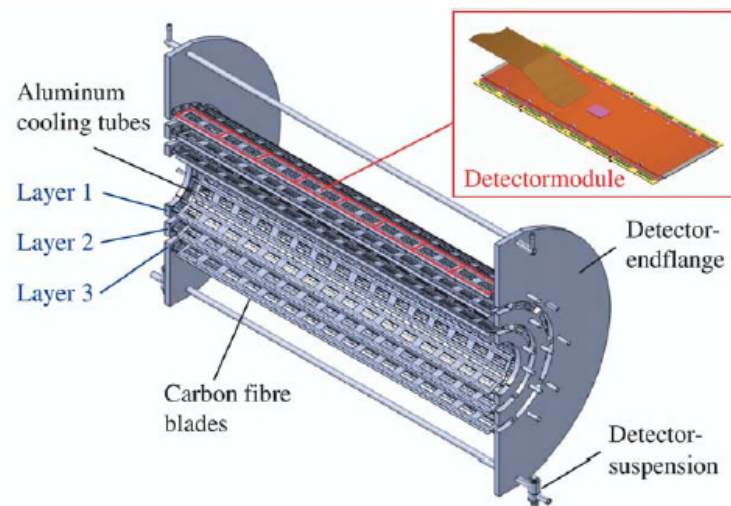


Figure 2.9: Support structure of the tracker.

- *FPix*, the endcaps. They are made of two completely separate sections at the two opposite sides with respect to the interaction point (IP), called half-cylinders, that support the actual pixel detector. They contain the mechanical and electrical infrastructures (support, position, cooling, power, control, readout) and they are made of berillium 0.5 mm thick to assure strength, stability and relatively low mass.

The **silicon strip detector** is made of 10 barrel layers and 3+9 disks for each endcap.

The main features are:

- *TIB*, inner barrel. Composed of 4 concentric cylinders at radii 255.0, 339.0, 418.5 and 498.0 cm,

it covers in z a region from -700 to $+700$ mm. Its sensors are $320\text{ }\mu\text{m}$ thick. The two inner layers are equipped with double sided modules and the pitch within the strips is $80\text{ }\mu\text{m}$. In the two outer layers the single sided modules have strips with a $120\text{ }\mu\text{m}$ pitch.

-*TID*, inner disk. It hosts 3 identical disks placed between $z=\pm 800$ mm and $z=\pm 900$ mm. Each one consists of 3 rings covering radii from 200 to 500 mm. The two inner rings house back-to-back modules, while the outer has single sided modules. The pitch of the $320\text{ }\mu\text{m}$ thick strips spans from 100 to $141\text{ }\mu\text{m}$.

-*TOB*, outer barrel. It is a single 2180 mm long wheel supporting 688 self-contained sub-assemblies called rods that provide support and cooling for 6 or 12 modules each. Its 4 disks of carbon fiber epoxy laminate are joined by three inner and three outer cylinders. The openings in the disks form six detection layers with average radii of 608, 692, 780, 868, 965 and 1080 mm. The strips are $500\text{ }\mu\text{m}$ thick with a pitch from 122 to $183\text{ }\mu\text{m}$.

-*TEC*, endcaps. They are called TEC+ and TEC- depending on the z direction with respect to the IP. They extend radially from 200 mm to 1135 mm and in z between ± 12400 mm and ± 2800 mm. Each one is made of 9 disks to support the detectors modules and 2 disks as front/back terminations. The strips are $500\text{ }\mu\text{m}$ thick and have a pitch 97- $184\text{ }\mu\text{m}$.

2.2.3 Electromagnetic Calorimeter ECAL

The electromagnetic calorimeter [62] is described in detail in section 2.3.

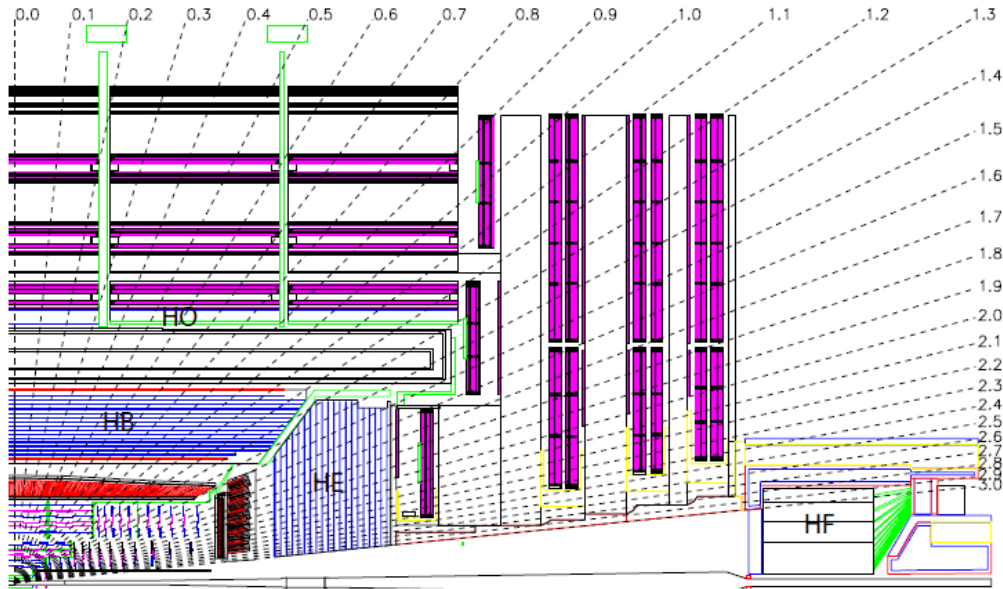


Figure 2.10: Longitudinal scheme of the CMS detector that shows the position of HB, HE, HO and HF.

2.2.4 Hadron Calorimeter HCAL

The Hadron Calorimeter [57, 63] is crucial for the measurement of hadronic jets and evaluation of the missing energy due to neutral particles like neutrinos, and they can be a signature of presence of Higgs bosons or supersymmetric particles. To assure the capture of all the particles ejected during the collision it must be hermetic. The space occupied by HCAL is limited by ECAL at $r=1.77$ m and by the magnetic coil at $r=2.95$ m.

The calorimeter is composed of four detectors, shown in figure 2.10, to assure coverage up to $|\eta| = 5$: HB, HE, HO and HF.

Barrel (HB) Built between ECAL and the solenoid coil, it is a sampling calorimeter, made of 36 identical azimuthal wedges parallel to the beam axis, that cover $|\eta| < 1.3$. The plates are assembled in order to avoid projective dead material.

A sampling calorimeter alternates layers of dense absorber, where the incident particles lose a large fraction of their energy, and of scintillator, the active material that produces fast light pulses at the passage of the particles. This trick reduces the dimensions of the detector in comparison to an homogeneous calorimeter.

The absorber is made of flat brass (70% Cu, 30% Zn) plates, with density 8.53 g/cm^3 , radiation length 1.49 cm and interaction length 16.42 cm . The inner and outer plates have the function of support, hence they are built with stainless steel. The total absorber thickness at 90° is $5.82 \lambda_I$ and increases with the pseudorapidity as $1/\sin \theta$, until it reaches $10.6 \lambda_I$ at $|\eta| = 1.3$. The granularity is $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$.

The absorber plates are interleaved by planes of scintillator. The last layer is thicker to contain the late developing showers that can cause energy leakages out of the calorimeter. Moreover, there is a layer 0 of scintillator, in front of the steel support plate, in order to sample the showers that develop in the inert material between ECAL barrel and HCAL barrel.

The light is collected by wavelength-shifting (WLS) fibers, made of a fluorescent material that absorbs a radiation at a given wavelength and emits at a lower frequency. The signal is read by hybrid photodiodes (HPD), characterized by low sensitivity to the magnetic field, large dynamic scale and gain of approximately 2000.

Endcaps (HE) Attached to the muon endcap yoke, they are the natural continuation of the HB and cover $1.3 < |\eta| < 3$.

Due to the high luminosity of LHC the HE are expected to be radiation tolerant and able to manage a large counting rate. They are designed to minimize the projective dead material, the cracks between EE and HE and the single-particle energy resolution. They must also support the weight of the EE and the preshower.

The 79 mm thick brass plates are interleaved by 3.7 mm thick scintillator layers. Also in HE there is a 9 mm layer 0 of scintillator to correct (partially) for the particles absorbed in the

support structure of EE and for the different response of EE to hadrons and electrons. The granularity is $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ in $|\eta| < 1.6$ and $\Delta\eta \times \Delta\phi = 0.17 \times 0.17$ in $|\eta| > 1.6$. As in HB, the light is collected by WLS fibers and the signal is read by HPDs.

Outer Hadron Calorimeter (HO) Also called tail catcher, it is placed in the first layer of each of the 5 rings of the iron yoke at $z=0, \pm 2.686$ and ± 5.342 m. Each HO level is constrained in 40 mm, 16 of them for the detector and the rest for the support structure. The scintillator is located just in front of the first layer of the barrel muon detector: 2 layers at $r=3.82$ and 4.07 m in the central ring (ring 0) and 1 layer at $r=4.07$ m in the others. The granularity is $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$.

This outer calorimeter is needed because the shower containment of EB and HB is not sufficient to have an adequate sampling depth at $|\eta| < 1.3$ that assures also a correct measurement of the missing energy. Hence, the HO is added outside the magnet, exploiting the solenoid coil as an addition absorber layer, to extend to a minimum of $11.8 \lambda_I$.

The tail catcher is essential to avoid leakages. In fact, without HO there is an excess of $E_{deposit}/E_{incident} < 1$, especially at $|\eta| = 0$ (ring 0), corresponding to the region of minimum material budget in HB. Moreover, the measurements with HO are more Gaussian since they recover the effect of leakage.

Forward Hadron Calorimeter (HF) The forward calorimeter is a quartz Cherenkov-based detector: it collects the light produced by charged particles above the threshold energy $\beta \geq 1/n$ (hence for the electrons 190 keV). It is a cylindrical steel structure with outer radius 130 cm and front face at 11.2 m from the IP, divided in 36 wedges.

Built in order to cover the pseudorapidity range $3 < |\eta| < 5$, this is the subdetector subject to the highest particle flux, hence a very radiation hard material is needed. Quartz fibres (fused-silica core and polymer hard-cladding) were chosen. Optical attenuation is induced by the accumulated dose, but this kind of material can survive more than ten years in such a harsh environment.

More than 1000 km of quartz fibres are inserted in the grooves of the steel absorber structure, parallel to the beam line. The granularity of the detector is $\Delta\eta \times \Delta\phi = 0.175 \times 0.175$.

Two sets of fibres of different length are used and read independently: half of them cover the full 165 cm of absorber while the other half starts at 22 cm from the front of the detector. This expedient allows to distinguish the showers due to electrons and photons from the hadronic showers, in fact the former lose most of their energy in the first 22 cm.

2.2.5 Muon detector

The muon chambers [57, 64] have the main aim to identify muons and measure their momentum, important for the search for the Higgs in the four leptons channel. Since the multiple scattering in the 16 interaction lengths of material in front of the muon detector degrades the

momentum resolution, information from the momentum measurement performed independently by the tracker are exploited.

There are three kinds of chambers with an overall acceptance up to $|\eta| = 2.4$ with no gap, hence muon identification is assured in the range $10^\circ < \theta < 170^\circ$. The total simulated reconstruction efficiency is 95 – 99%, except in the regions between two chambers.

In the barrel, where there is uniform magnetic field contained in the steel yoke and low muon rate, four sections of **drift tubes** (DT) are interleaved with the four flux return plates, concentric around the beamline. In the first three stations there are 8 chambers (4 to measure the coordinate in the bending plane and 4 for the z coordinate) while the fourth doesn't measure z. Each drift cell is offset of a half-cell width with respect to the neighbouring. In this way they cover up to $|\eta| < 1.2$ avoiding dead spots.

The DT have standard rectangular drift cells of 21 mm, that determine a maximum time of drift of 380 ns in a mixture of 89% Ar + 15% CO₂. This is an acceptable compromise between short occupancy and limited number of channels. A DT is composed of 3 (or 2) superlayers (SLs) each one made of four layers of cells. In the two external SLs the wires are parallel to the beam line in order to track in the bending $r - \phi$ plane, while in the internal SL (not present in the fourth station) they are orthogonal, to access the z position along the beam line. A single SL has time resolution of a few seconds that assures local efficient bunch crossing identification.

The goal is to achieve a spatial $r - \phi$ resolution of 100 μm (single wire resolution better than 250 μm) with a deviation from linearity of the space-time relation in the single cell less than 100-150 μm .

Figure 2.11 shows a cell composed by: five aluminium and mylar electrodes, one gold-plated stainless-steel anode wire, two field shaping strips and two aluminium and mylar cathode strips.

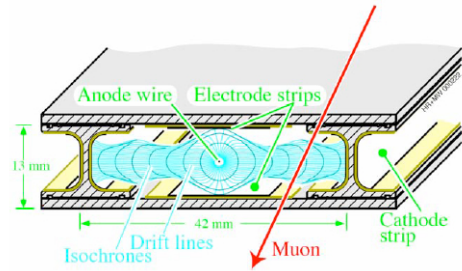


Figure 2.11: Drift cell scheme.

In the endcaps, region with non-uniform magnetic field and high muon and background rate, the detector is required to be characterized by fast response, fine segmentation, no need of precise gas, temperature or pressure control, radiation resistance and strip configuration easy to adapt to the endcap region. **Cathode Strip Chambers** (CSC) are chosen for this goal. The 540 trapezoidal multi-wire proportional chambers assure continuous ϕ coverage. A muon passing in $1.2 < |\eta| < 2.4$ crosses three to four CSC layer and in $0.9 < |\eta| < 1.2$ is detected by both DT and CSC.

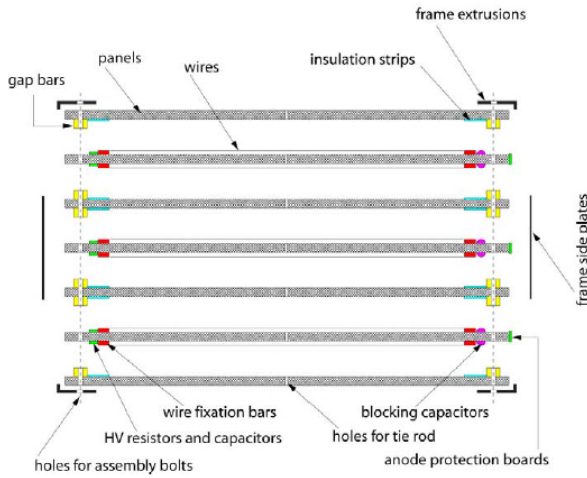


Figure 2.12: Structure of a cathode strip chamber.

view of the CSC structure.

The single plane resolution of the inner chamber is $80\ \mu\text{m}$ and assures a 6-plane chamber resolution about $33\ \mu\text{m}$. In the outer planes, depending on the strip pitch, the resolution is worse, but the strips in adjacent planes are staggered by half of the pitch width such that it corresponds to a strip with poor resolution in one plane another one with good resolution in the nearby plane in order to guarantee a 6-plane resolution about $80\ \mu\text{m}$.

The chambers are filled with 40% Ar, 50% CO_2 as a quencher and 10% CF_4 to avoid polymerization.

In addition, the region up to $|\eta| = 2.1$ is covered by **Resistive Plate Chambers (RPC)** for fast, independent and highly segmented trigger. Four layers of RPCs are located in the end-caps, while six in the barrel (two in each of the inner stations and one in the outer two). The RPCs are gaseous parallel-plate detectors operated in avalanche mode. They are made of two parallel 2 mm thick plates of Bakelite, with very high resistivity, one positively charged (anode) and the other negatively charged (cathode). A strip plane sits in between and it is detached from the electrodes by a 2 mm gap on each side. The gaps are filled with a mixture of 96.2% $\text{C}_2\text{H}_2\text{F}_4$ + 3.5% $i\text{C}_4\text{H}_{10}$ + 0.3% SF_6 and water vapor to maintain the bakelite resistivity. The layout of an RPC is pictured in figure 2.13. The hit strips measure very quickly the muon momentum for triggering. Moreover, the bunch crossing (BX) is unambiguously assigned to the observed tracks. An RPC is so fast that it is capable of tagging the event of an ionizing event in less than the 25 ns that separate two consecutive bunch crossings. The total induced signal is estimated as the sum of the signals induced in each of the gaps, hence high efficiency is reached with restrained voltage.

As in the barrel, the detector is made of 4 stations, interleaved with the iron yoke plates. Each chamber is composed by 6 anode wire planes interspersed with 7 cathode panels. The azimuthal wires define the radial coordinate of the track and those hosted inside the solenoid are tilted to assure an electronic drift parallel to the beam line. The strips milled on cathode panels, separated by 5 mm, are radial (they run lengthwise at constant $\Delta\phi$ width) and have a pitch from 8.4 to 16 mm. The ϕ coordinate of the muon along the wires is obtained by interpolating charges induced on strips by ions. In figure 2.12 is an exploded

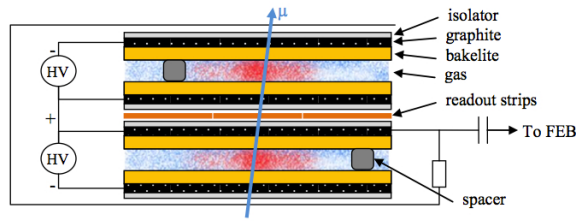


Figure 2.13: Layout of a double-gap RPC.

2.2.6 Trigger

At LHC there is a bunch crossing every 25 ns, i.e. with a frequency of 40 MHz, and during each of them several interactions take place and eject particles. Such a huge amount of data cannot be saved and stored, hence an important reduction of the rate is essential and challenging. The trigger system [57] has this job and it is composed by a programmable hardware trigger, Level-1 (L1) [65], that reduces the rate to 100 kHz and a software High-Level Trigger (HLT) [66] with an output of about 100 Hz. A Trigger Supervisor [67] checks the operation of the two triggers.

The L1 (in figure 2.14) is composed by a Calorimeter Trigger and a Muon Trigger whose information is then sent to the Global Trigger, the one which takes decisions. Each of the two

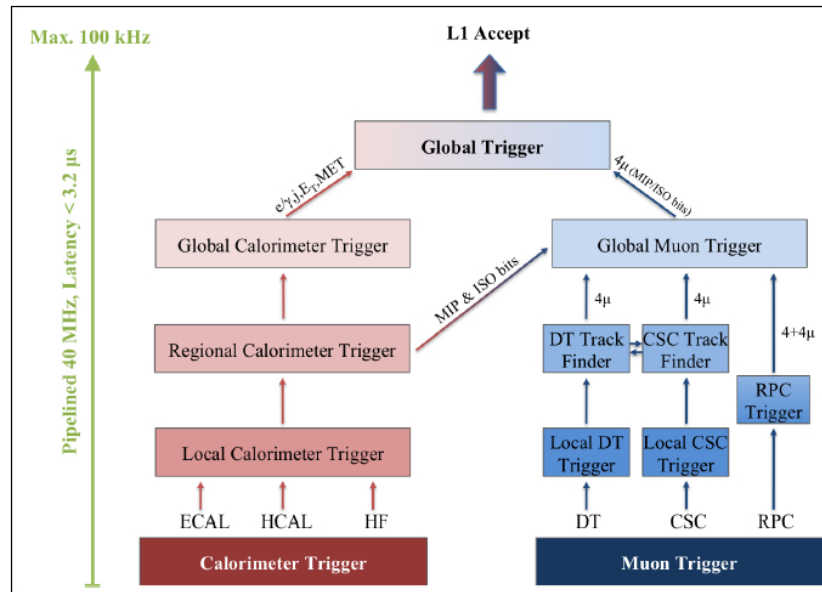


Figure 2.14: Scheme of the Level 1 trigger.

systems has a local, a regional and a global level. At local level (called Trigger Primitive Generator, TPS) the energy deposits in each trigger tower of ECAL, HCAL and HF and the hits in each DT, CSC and RPC detector are used. At the following step the information from the

TPSs is merged to find some candidates ordered for quality in specific regions. The top ranking candidates among them are selected by the Global Calorimeter Trigger and the Global Muon Trigger. The decision, based on E_T threshold cuts, about which of them must be rejected and which can be delivered to the HTL with their energy and isolation information is taken by L1 Global Trigger.

The HTL performs more complex operation: it uses data also from the tracker and makes calculations to reconstruct the events in some regions of interest. The selection takes place in 3 steps: Level 2 that uses only the data from the calorimeters, Level 2.5 that exploits the pixel detector and Level 3 that access the information from the whole detector. This trick is useful to save time, in fact the algorithms of the tracker are quite slow and if the event is rejected by Level 2 there is no need to go on.

All these expedients allow to work at high energy and luminosity (i.e. high statistics) and with excellent rejection of the undesired events that could overpower the interesting signals. Indeed, CMS was able to observe an event as rare as the Higgs boson over a large background during Run1 with an integrated luminosity of 24.6 fb^{-1} (4.9 fb^{-1} at $\sqrt{s} = 7 \text{ TeV}$ and 19.7 fb^{-1} at 8 TeV). For Run2, at increased luminosity and more harsh pileup conditions, the trigger system has been improved and the thresholds increased.

2.3 Electromagnetic Calorimeter ECAL

The electromagnetic calorimeter (ECAL) [57] is aimed to reconstruct the energy of electrons and photons and it is crucial for the study of the Higgs boson in the decay channels $\gamma\gamma$, $ZZ \rightarrow 4e$, $ZZ \rightarrow 2e2\mu$ and $WW \rightarrow 2e2\nu_e$. It detects the energy deposited through the showering of particles in matter exploiting the scintillation mechanism in crystals (i.e. the conversion of energy in light) and uses photodetectors to collect the light signal which is then converted to digital and stored for the analysis.

2.3.1 Showering and calorimetry

A technique used to detect particles and measure their energy is calorimetry [68]: the particle is absorbed in a dense medium where it creates a shower of secondary particles, with peculiar characteristics depending on the incident particle.

The electromagnetic showers are induced by an electron or a photon and they grow step by step due to subsequent pair production (conversion of a photon into a positron-electron system) and bremsstrahlung (emission of light by a particle which is deflected or decelerated in the medium). The showering stops below the critical energy $E_C = \frac{550 \text{ MeV}}{Z}$ at which pair production and bremsstrahlung become negligible and the remaining energy is lost through ionization and excitation. Indeed, at E_C the maximum of the shower in $t_{max} = \ln\left(\frac{E_0}{E_C}\right) + C_{e,\gamma}$ is reached, with E_0 incident energy, $C \simeq -0.5$ for photons and $C \simeq -1.0$ for electrons. Harder particles travel deeper in matter. The main contribution to the transverse development of the core of

the shower at high energy (beginning of the shower) comes from bremsstrahlung and pair production. The lateral tails of the shower are determined by low energy Compton scattering and photoelectric effect. Close to the critical energy there is also contribution from multiple scattering. The transverse containment is quantified by the Molière Radius $R_M(87\%) = \frac{21\text{MeV}}{E} X_0$. On the other hand, the hadronic showers are induced by a hadron that interacts strongly with the material of the detector. The two showers are distinguished from their shape (figure 2.15 [69]): the electromagnetic showers (left) are narrow while the hadronic ones (right) spread wider around the central axis and may show secondary showers.

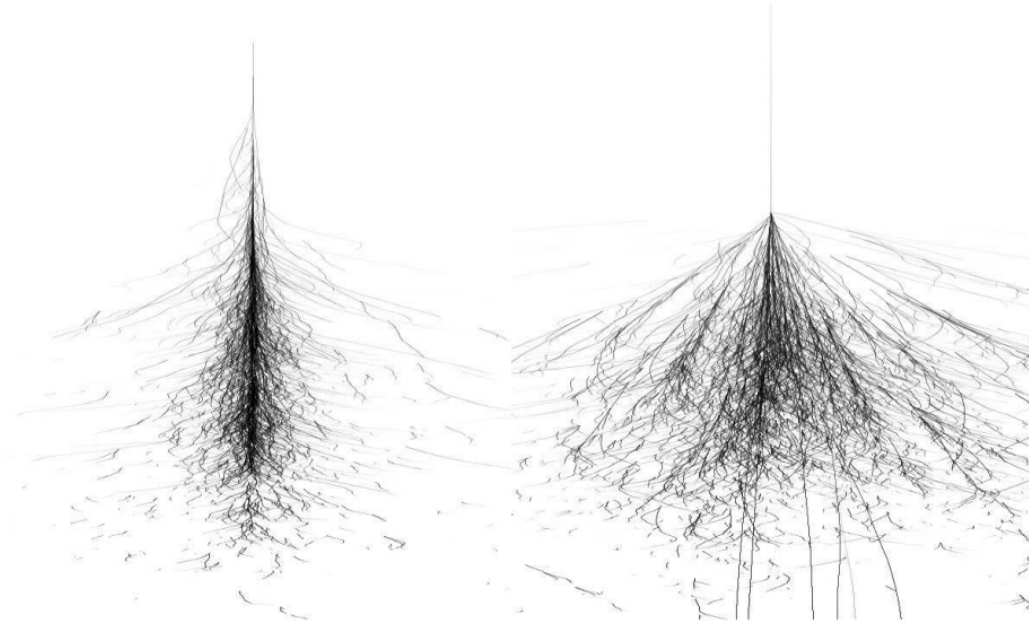


Figure 2.15: Electromagnetic (left) and hadronic (right) showers.

The mean path of an electron, i.e. the distance traveled before losing all but $(1/e)$ -th of its energy, is called radiation length X_0 , such that $E(x) = E_0 e^{-x/X_0}$. The 95% of the electromagnetic shower is longitudinally contained in $L(95\%) = t_{max} + 0.087Z + 9.6[X_0]$. The mean free path of a photon is $\lambda = \frac{9}{7} X_0$. The corresponding quantity for hadrons is the nuclear interaction length λ_I and it is larger than X_0 , in fact the hadronic showers begin deeper in the same medium.

This method of detection destroys the particle to convert its energy in a detectable signal (usually light), but it has the advantage that it can be applied also on neutral particles, indeed it is fundamental to detect neutral hadrons and photons. A precise measurement of their energy is also needed to estimate the missing energy. Calorimeters are largely employed in nuclear and subnuclear physics experiments and they are of two kinds: electromagnetic (ECAL) and hadronic (HCAL) with the proper design to reconstruct the corresponding showers.

In order to stop electrons or photons a reduced amount of material (measured in terms of X_0) is needed, hence an electromagnetic calorimeter with large Z can have compact design. The calorimeters are also classified in “homogeneous” (top) and “sampling” (bottom) calorimeters (figure 2.16). The former ones are entirely built with the active material that act as detecting medium (often a scintillator), they collect all the energy, i.e. have an excellent energy resolution. The latter ones intersperse the active medium with a very dense absorber (passive showering material) which has the function to decelerate quickly the particle but cannot detect it, hence part of the shower is not observed and the resolution is worse, but its dimension is much smaller. The hadronic showers that have a larger longitudinal development are detected with sampling HCAL.

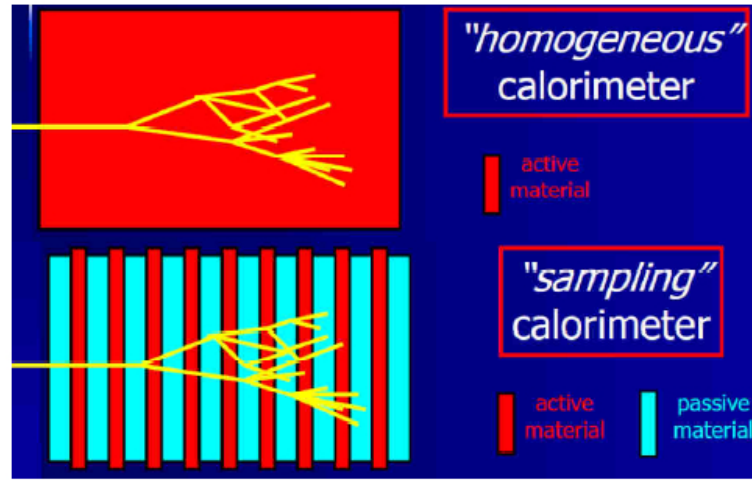


Figure 2.16: Homogeneous (top) and sampling (bottom) calorimeters.

The energy resolution of the calorimeter generally depends on three components:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \quad (2.2)$$

where a is the stochastic term, b the noise term and c the constant term.

The stochastic term consists in shower fluctuations (number of particles in the shower), sampling fluctuations (depending on the ratio $E_{active}/(E_{active} + E_{passive})$ in sampling calorimeters) and signal quantum fluctuations (number of detected photoelectrons). The noise term is due to electronic noise, pedestal (global offset) shifts and their uncertainty, radioactivity, pileup and shower leakages upstream ECAL. The constant term, dominating at high energy, is generated by gain (global scale) shifts and their uncertainty, lack of hermeticity (presence of dead material, longitudinal leakages), detector response non uniformity, mechanic tolerance, imperfections in calorimeter construction and geometry and technology effects (recombination, light attenuation, gas saturation).

2.3.2 Material and layout

In CMS an homogeneous calorimeter of scintillating crystals was chosen as electromagnetic calorimeter because its excellent resolution enhances the discovery potential in $H \rightarrow \gamma\gamma$. The choice of material fell on lead tungstate $PbWO_4$ because of:

- the high density that allows to build a very compact object;
- the short radiation length such that the shower begins early and it is longitudinally contained in the detector;
- the small Molière radius R_M that assures transverse containment;
- the fast scintillation that guarantees low occupancy;
- the already acquired knowledge in its construction.

The main features of $PbWO_4$ are summarized in table 2.2.

density [g/cm ³]	8.28
radiation length X_0 [cm]	0.89
interaction length [cm]	22.4
Molière radius [cm]	2.19
light decay time [ns]	5 (39%), 15 (60%), 100 (100%)
refractive index	2.30
peak emission [nm]	440 nm

Table 2.2: Parameters of the lead tungstate.

The electromagnetic calorimeter [62] is composed of about 76000 truncated-pyramid shaped crystals whose pointing is tilted of 3 degrees in η and ϕ with respect to the interaction point to eliminate gaps between the crystals. The side of the front face of the crystals has the dimension of about R_M and their length corresponds to about $26X_0$ in the central region. The calorimeter covers the range of $|\eta| < 3$ with a few cracks and gaps. Precise separation between electrons and photons and energy measurement is obtained up to $|\eta| = 2.5$, limit of the geometrical acceptance of the tracker. The remaining ECAL region improves the energy-flow measurement in the forward direction. The layout of ECAL is shown in figure 2.17. It is divided in barrel and endcaps. A preshower is added in the endcaps region.

The **barrel EB** covers $|\eta| < 1.48$ and has an inner radius of 1.29 m. Its crystals are grouped in arrays 2x5 (in $\eta - \phi$), the subchannels. In turn, they are assembled to form modules and supermodules. Half of EB contains 18 supermodules, each of them covers 20° in ϕ and 85° in η and is made of 4 modules, the first one of 500 crystals and the others of 400. The **endcaps EE** cover $1.48 < |\eta| < 3.0$ and in each endcap the 7320 crystals are split in 36 supercrystals. 18 supercrystals form a dee. Here the crystals are a bit shorter than in the barrel because in the

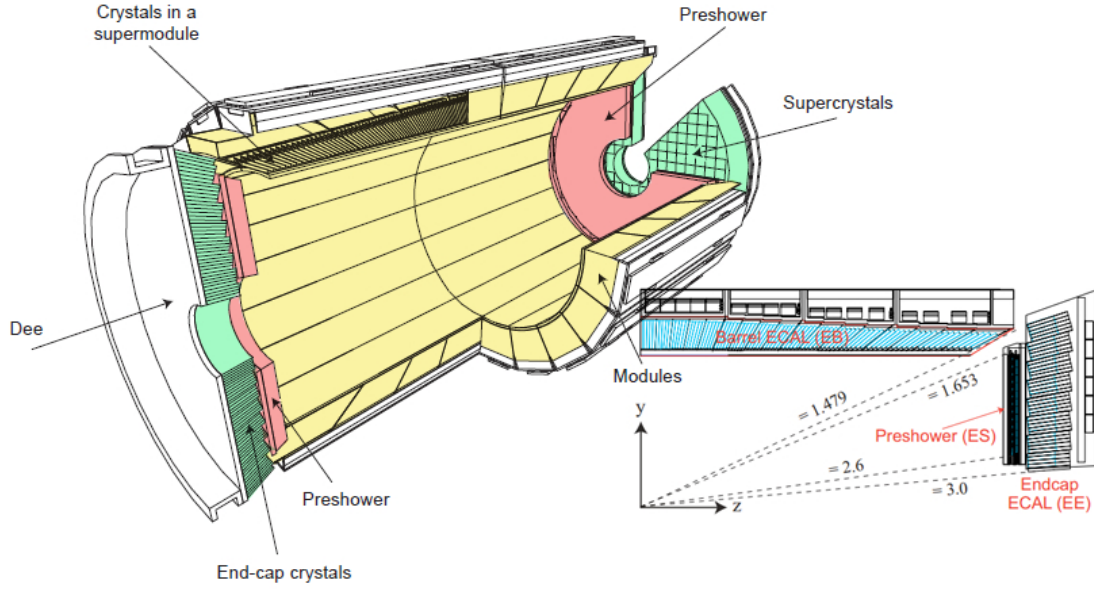


Figure 2.17: Layout of the electromagnetic calorimeter.

region $1.64 < |\eta| < 2.6$ the **preshower ES** accounts for $3X_0$. It is made of 2 lead converters and corresponding detector planes of silicon strips with pitch 2 mm. Its goal is to separate photons from pions that may be misidentified as photons and increase the background for example of the $H \rightarrow \gamma\gamma$ measurement (15% contribution). About 6 – 8% of the energy is lost in the preshower, less for harder incident particles that shower deeper in the calorimeter. The parameters of EB, EE and ES are listed in table 2.3.

The light produced by the scintillation is collected with silicon avalanche photodiodes (APDs) in EB and with vacuum phototriodes (VPTs) in EE. Their peak of absorption in the green-blue region is well matched with the maximum of emission of $PbWO_4$. The APDs cover 50 mm^2 of the crystal surface, have a gain of 50 and quantum efficiency of approximately 80% at the emission wavelength of lead tungstate. The VPTs cover 180 mm^2 with gain about 10 in the 4 T magnetic field and efficiency 15%. The VPTs have been chosen for the EE because despite their lower signal to background ratio they are more resistant to radiation and work better in a strong quasi-axial magnetic field.

The front-end electronics of the EB, EE, and ES samples the analogue signals from the detectors (APDs, VPTs and silicon sensors) at 40 MHz. It reads the signal with a 12-bit analogue-to-digital converter (ADC). Ten samples are acquired and stored for each bunch crossing (at each trigger signal): the first 4(3) in EB(EE) are the pedestal while the signal has the maximum in the 6th (sample 5). In figure 2.18 the amplitudes are peaked normalized to 1. Only three samples are stored in the ES.

	Barrel	Endcaps
Pseudorapidity coverage	$ \eta < 1.48$	$1.48 < \eta < 3.0$
r_{inner}, r_{outer} [mm]	1238, 1750	316, 1711
z_{inner}, z_{outer} [mm]	$0, \pm 3045$	$\pm 3170, \pm 3900$
Granularity $\Delta\eta \times \Delta\phi$	0.0175×0.0175	0.0175×0.0175 to 0.05×0.05
Crystal dimension [mm ³]	$21.8 \times 21.8 \times 230$	$24.7 \times 24.7 \times 220$
No. of crystals	61200	14640
Depth in X_0	25.8	24.7
Modularity	36 supermodules	4 dees
1 supermodule/dee	1700 crystals	5382
1 supercluster unity	-	36 crystals
Operating temperature	18°	18°
Preshower		
Pseudorapidity coverage	$1.65 < \eta < 2.61$	
Silicium detector	4512	
Strip pitch	1.9 mm	
Pitch length	61 mm	
Operating temperature	-5°	

Table 2.3: Parameters of EB, EE and ES.

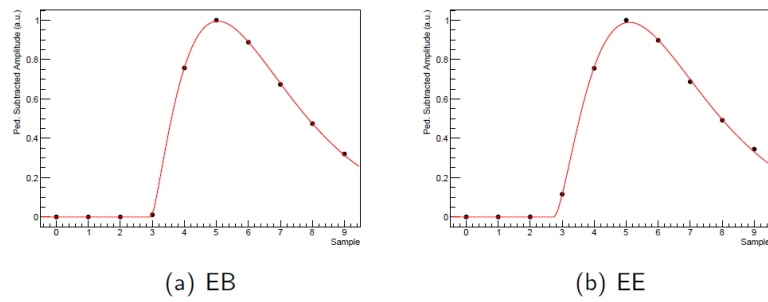


Figure 2.18: Pulse shape reconstruction in ECAL using the 10 samples collected for each signal pulse.

At the design luminosity of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ about 10^9 inelastic pp interactions take place per second, causing a very strong irradiation of the detector, hence provoking damages. This hard radiation induces leakage current on the APDs that increase the noise and can also degrade the VPTs response. The strongest effect of radiation is the creation of colour centers in the crystals that drop the transparency [70] (up to 6% in EB and 30% in EE), hence the light yield. When the irradiation stops the transparency restore through spontaneous annealing. A light monitoring system [71] that injects a laser at $\lambda = 440 \text{ nm}$ into each crystal is build to check and correct the energy measurements for changes of transparency.

From a mechanical point of view it was important to minimize the material in front of ECAL, where electrons can undergo bremsstrahlung energy loss and photons may convert, to optimize the interface and the alignment with the tracker in presence of a strong magnetic field, to ensure hermeticity (minimize the gaps between the crystals and between EB and EE), to minimize the gap between ECAL and HCAL (otherwise the energy of jets and the missing energy would be mis-estimated) and to maintain the temperature stable. Variations of temperature, that degrade the light yield of crystals and the APD gain ($-2\%/^{\circ}\text{C}$), and of bias APD voltage that drop the gain ($-2\%/V$) are kept under control, down to a negligible level.

2.3.3 Calibration

The energy reconstructed with ECAL must be corrected for single channel response variation and containment and an absolute energy scale is needed [72, 73]. The energy corrected for detector effects is $E_{e/\gamma} \propto F_{e,\gamma}(\eta) \sum_{clusters} C_i \cdot S_i(t) \cdot A_i$. Indeed, the terms $S_i(t)$, C_i , $F_{e,\gamma}$ and G are measured during calibration.

As mentioned in the previous subsection, the harsh radiation induces formation of color centers in the crystal. The consequent variation of transparency is accounted in the constant $S_i(t)$ which is checked continuously (every 40 minutes) by a light monitoring system. In fact, the radiation does not affect the scintillation mechanism itself and the intrinsic energy resolution, but it changes the properties of the detector and its performance. The change of response to scintillating light from the initial value S_0 to the one at time t , $S(t)$, is estimated in comparison to the response to laser light R :

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

The exponent α is measured in beam test, varies between 1 and 1.5 depending on the crystal producer and has a spread of 10% for residual transparency differences. During in-site tuning α is optimized with $Z \rightarrow e^+e^-$ events to achieve the best energy resolution.

The contributions of the calibration for time dependent instabilities to the constant term of

resolution resides in the precision of the individual laser monitoring measurements ($< 0.1\%$), long term instability of the single channel ($< 0.2\%$), spread of α (0.3%) and uncertainty on the choice of an effective $\alpha = 1.16$ for the endcaps to compensate the VPTs response drop due to radiation (1.5%).

The intercalibration constants C_i that equalize the response of all the single channels are measured with different methods, among them the preferential ones are [74]:

- ϕ -*symmetry*: the correction is found from the ratio between the total transverse energy in a crystal over the average of the sum of the total transverse energies of all the crystals at same η , based on the assumption that the total energy deposit should be the same for all crystals at a fixed pseudorapidity. The measurements are performed with minimum bias events with an energy trigger of 150 MeV in EB and 650 MeV in EE, a lower E_T threshold at six times the RMS noise of the channel and an upper E_T threshold at 1 GeV above the lower to cut the rare events with very large transverse energy. The precision of the method is $1.5\% - 3\%$, due to dishomogeneities of the material that drop the above assumption of symmetry;
- π^0 and η^0 *calibration*: the energy deposits of unconverted photons from the decay of pions or eta resonances are measured and used to construct the invariant mass of the diphoton, which is fitted with a Gaussian (for the signal) and a polynomial of 4th order (background). A constant is added to each channel to correct the fitted mass and match the expectations. Only photons that are collected far from the boundaries of the fiducial region are used, to avoid leakages. The precision of the method is $0.5\% - 1\%$ in EB and $2\% - 4\%$ in EE;
- calibration with *isolated electrons from W or Z decays*: the calibration constants are obtained from the comparison of the energy reconstructed in ECAL and the momentum measured in the tracker, by solving iteratively the system $\sum_i C_i E_i = p_{track}$. The precision of the calibration is $0.8\% - 4\%$ in EB and up to 6% in EE due to material inhomogeneities and it is limited by the statistics. Electrons without bremsstrahlung loss have a narrower E/p peak and grant the best quality calibration..

The constants obtained with the different methods are then combined (in figure 2.19), weighted by the relative precision.

The contribution of the intercalibration to the constant term of the energy resolution is 0.3% in $|\eta| < 1.0$, 0.5% in $1.0 < |\eta| < 1.48$, 1.0% in $1.6 < |\eta| < 2.3$ and 1.5% up to $|\eta| = 2.5$. The main sources of channel-to-channel response variations are the crystal light yield spread in EB (15%) and the gain spread of the photodetectors in EE (25%). The effect of intercalibration on the invariant mass is shown in figure 2.20.

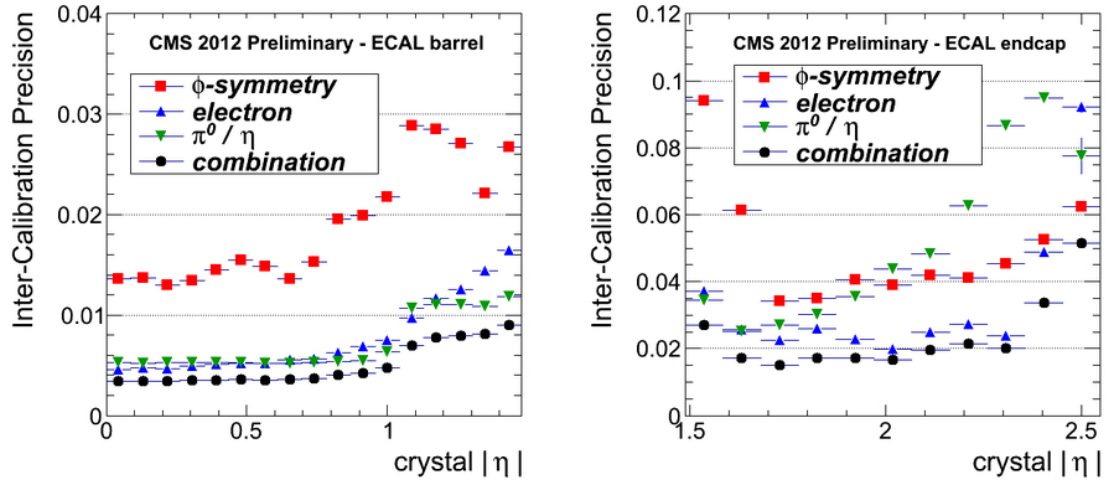


Figure 2.19: Precision of the different methods of intercalibration at 8 TeV and their combination (black).

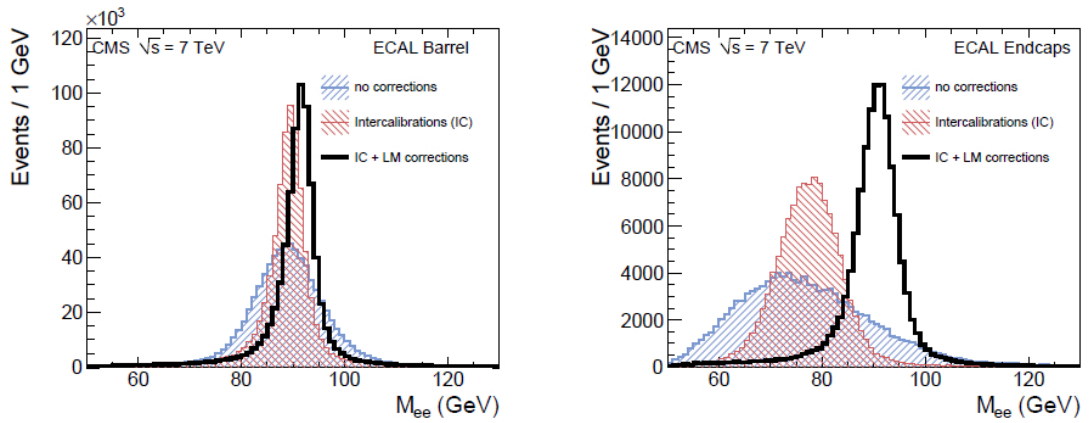


Figure 2.20: The peak of the Z boson in the dielectron invariant mass is shown, for EB on the left and EE on the right, in blue without corrections, in red with intercalibration corrections and in black with both intercalibration and laser monitoring corrections.

The corrections $F_{e,\gamma}$ account for the containment, both of the shower inside the calorimeter and of the supercluster energy, for particles that start the shower in the material upstream ECAL. These constants are optimized with a multivariate analysis based on η and ϕ coordinates of the supercluster, shower shape variables (like those studied in section 4.2), number of primary vertices (to account for multiple interactions) and local cluster coordinates to measure the distance of the cluster from boundaries, cracks and gaps of ECAL, where part of the energy can be lost. The corrections are dominated by material effects.

The absolute energy scale G (ADC-to-GeV conversion) was initially introduced during a test beam, where the sum of the energies in a 5×5 crystals matrix was equalized to the one of the electrons incident beam. After the installation in the detector the in-site calibration is performed on electrons from Z decay, selected with $E_T > 25$ GeV (2% background contamination left). The invariant mass of the dielectron system is fitted in the range $60 < m_{ee} < 120$ GeV using a Breit-Wigner (that describes the resonant peak) convoluted with a Crystal Ball (CB) function that accounts for the gaussian term of the resolution σ_{CB} . A maximum likelihood is performed with the mass and width of Z fixed at $m_Z = 91.188$ GeV and $\Gamma_Z = 2.495$ GeV, and the peak shift of the CB function Δm and the resolution σ_{CB} to be determined. The same fit is performed on the Monte Carlo distribution of m_{ee} and the energy scale G is computed as $1 - (\Delta m_{MC} - \Delta m_{data})/m_Z$.

2.3.4 Resolution

The energy resolution [72, 73] described in equation 2.2 and corrected with the calibrations explained in the previous section is measured experimentally on dielectron systems from the decay of Z bosons. At the test beam it was measured to be:

$$\frac{\sigma_E}{E} = \frac{2.81\%}{\sqrt{E(\text{GeV})}} \oplus \frac{12\%}{E(\text{GeV})} \oplus 0.3\%.$$

The energy resolution is estimated in different η bins, in the range $89 < m_{ee} < 100$ GeV around m_Z and in two regions that separate converted and unconverted electrons. A maximum likelihood fit is performed on the convolution (*Voigt*) of a Breit-Wigner with a Gaussian response function of width $\sigma_{m_{ee}}$:

$$\mathcal{L} = \prod_i \text{Voigt}(m_{ee}^i, \sigma_{ee}^i; m_Z, \Gamma_Z),$$

where the mass and width of Z are fixed parameters, while the mass and uncertainty $\sigma_{m_{ee}}/m_{ee} = (\sigma_{E_1}/E_1 \oplus \sigma_{E_2}/E_2)/2$ of the dielectron are the free parameters to be determined. In figure 2.21 the energy resolution of electrons from Z decay is shown, for unconverted electrons on the left and converted on the right, in red for simulations, in black with unmodified calibrations of 2012 and in light blue with improved corrections using the 2015 dataset [75].

It is observed that the energy resolution depends strongly on the pseudorapidity because it is affected by the presence of material upstream of ECAL. In fact it degrades above $|\eta| = 2$ and in

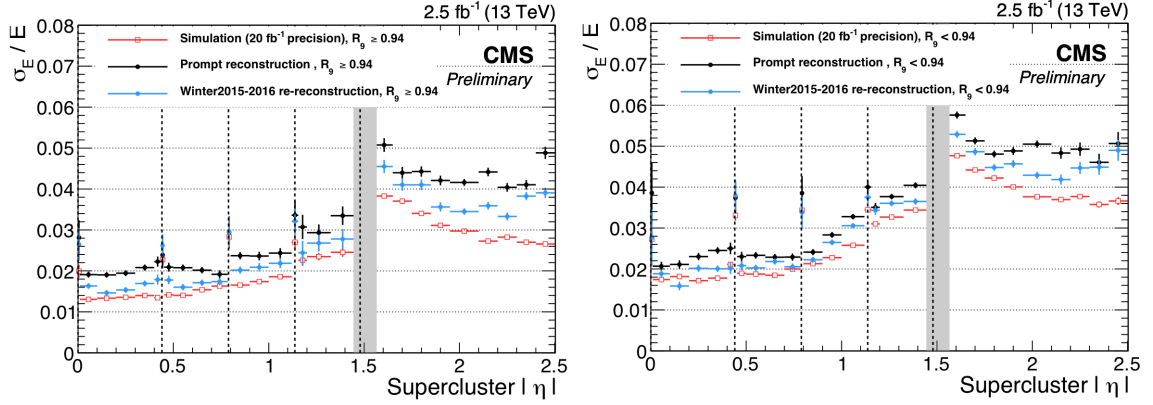


Figure 2.21: Energy resolution of electrons from Z decay is shown, for unconverted electrons on the left and converted on the right, in red for simulations, in black with unmodified calibrations of 2012 and in light blue with improved corrections using the 2015 dataset.

the vicinity of module borders.

The mis-modeling of the material of the tracker in front of ECAL causes also discrepancies between data and Monte Carlo. In addition, at larger pseudorapidities the single-channel response variations are not fully modeled in simulations. To account for this deviation a gaussian smearing term is added to the simulation. It is the quadratic difference between the electron resolution in data and MC simulation and it is summed in quadrature to the electron and photon energy.

The resolution of photons $\sigma_{E_\gamma}/E_\gamma$ is extracted from simulated $H \rightarrow \gamma\gamma$ events with the same selections applied in the Higgs data analysis. It is estimated by fitting the distribution of the ratio between the reconstructed photon energy and the true one. The observed width of the Higgs is dominated by the resolution of the diphoton mass, that in turn is affected by the resolution of the reconstructed energy of the two photons $\sigma_{E_{\gamma_1}}/E_{\gamma_1}$ and $\sigma_{E_{\gamma_2}}/E_{\gamma_2}$ and the uncertainty on the measurement of the angle between the two photons $\sigma_{\theta_{\gamma\gamma}}/\tan(\theta_{\gamma\gamma}/2)$. The latter becomes negligible if the primary vertex is correctly assigned (section 3.6). The diphoton invariant mass resolution

$$\frac{\sigma_{m_{\gamma\gamma}}}{m_{\gamma\gamma}} = \frac{1}{2} \left(\frac{\sigma_{E_{\gamma_1}}}{E_{\gamma_1}} \oplus \frac{\sigma_{E_{\gamma_2}}}{E_{\gamma_2}} \oplus \frac{\sigma_{\theta_{\gamma\gamma}}}{\tan \frac{\theta_{\gamma\gamma}}{2}} \right)$$

is continuously improved and reached the excellent value of about 1% in the best photon conditions.

The σ_{E_γ} is corrected for the smearing which is added as a constant term. A discrepancy remains between data and simulation with dependency on η . The main sources of discrepancy are:

- description of the tracker material in simulation. An incomplete or incorrect modeling of the material in front of ECAL can explain the degrading of the resolution in data with respect to Monte Carlo. A study of the material is described in section 4.1;

- clustering improvements and energy corrections. They must be optimized to mitigate for material effects, especially in the vicinity of borders, cracks and gaps;
- imperfect knowledge of the channel-by-channel α parameter whose spread degrades the resolution. An uncertainty of 10% in α produces a response change of the 3% in EB and up to the 15% in EE, i.e. it has an effect of the 0.3% – 1.5% on the energy resolution;
- imperfect estimate of the intercalibration systematic uncertainty. The uncertainties of the three combined methods are assumed completely independent, but they may be overestimated because geometrical and other experimental effects are not accounted;
- the lack of implementation of the change of transparency in the simulation.

With the smearing the single photon energy resolution is found 1.1% – 2.6% in EB and 2.2% – 5% in EE.

Chapter 3

Electrons and photons reconstruction for Higgs boson studies

One of the main goals for the construction of the CMS detector was to observe the Higgs boson. The benchmark channel $H \rightarrow \gamma\gamma$ is the cleanest one, and despite its low branching ratio ($\sim 0.2\%$) it is the favored one for the search of the Higgs boson, because it allows to measure its mass with great precision.

As explained in section 1.3.4, there are several mechanisms of production for the Higgs boson. The main one at LHC is the gluon fusion (ggH) which is an “untagged” channel because it does not have any additional signature (“tag”).

The “tagged” mechanisms, that take place with lower probability at LHC, are the vector bosons fusion (VBF), the associated production with a vector boson (VH), also called Higgsstrahlung, and the associated production with a top-antitop quark pair ($t\bar{t}H$). The tags of these mechanisms are the presence of two jets separated by large rapidity (back-to-back jets) for the VBF channel, leptons or missing energy E_T^{miss} for the VH channel. The $t\bar{t}H$ channel is accompanied by a b-quark pair plus two W bosons and might have in the final state also charged leptons from top decay or additional jets from top decay.

Two kinds of analysis are applied when studying the SM Higgs boson: a more simple cut-based analysis and a more sophisticated Multivariate Analysis (MVA). A better event selection (15% higher sensitivity) is achieved with the latter.

A system of detectors built around the interaction point, like the CMS experiment, is needed for the selection of the interesting events: the decay products of the Higgs boson and possible tags of its mechanism of production. To distinguish the Higgs event over the background it is important to identify the different particles that come out of the interaction. This is done exploiting information from four different subdetectors in CMS. In particular, the electrons [76] and photons [77] are reconstructed using the energy deposits in ECAL and tracking information. With a precise reconstruction of the energy deposits, the photons from the Higgs boson decay in the diphoton channel are selected. The presence of hits in the tracker allow to separate them from the electrons. Electrons may be signature of the decay of the Higgs boson in a

pair of Z bosons. Otherwise, they might come from a Drell-Yan Z boson and are exploited for calibration of the method of analysis.

In section 3.2 the energy measurement and cluster reconstruction are described. In sections from 3.3 to 3.10, as an example, the selection of the diphoton candidates from $H \rightarrow \gamma\gamma$ is explained.

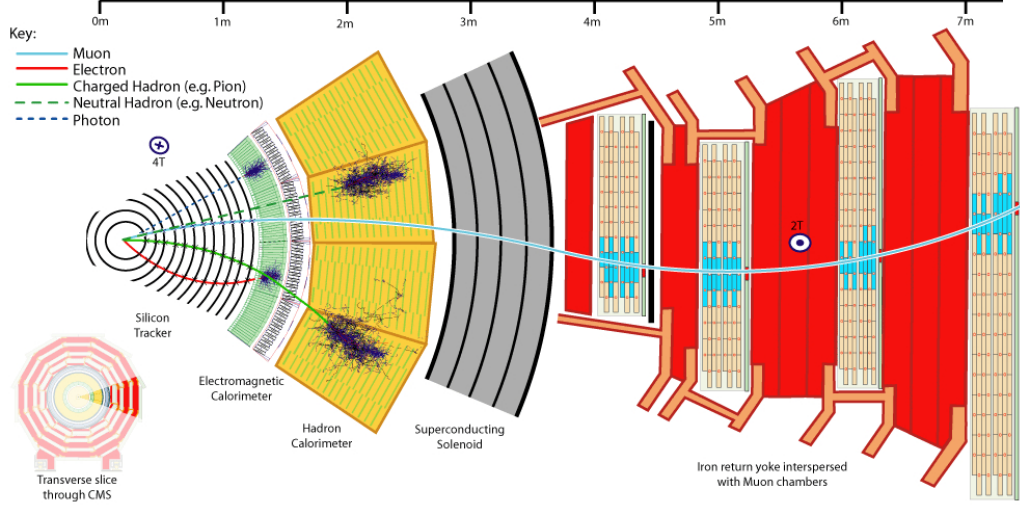


Figure 3.1: CMS section. It shows how each particle interacts with the different subdetectors.

3.1 Particle identification

The particle identification (particle ID) [78] takes place through the global event reconstruction (or particle-flow event reconstruction) and it is optimized combining information from all the subdetectors. Photons are identified by means of ECAL energy clusters which are not associated to a charged track. Electrons are recognized from their primary charged particle track linked to an ECAL energy cluster and the deposit due to the photons lost by bremsstrahlung in the material upstream of ECAL. The muons show a track both in the tracker and in the muon system. The tracks which are not associated to an electron or a muon are reconstructed as charged hadrons, while the neutral hadrons signature is an HCAL energy cluster without a corresponding charged track.

The technique to reconstruct the energy is different depending on the particle. The photon energy is taken from the ECAL. The energy of the electron is accessed from the momentum at the main vertex and the energy deposits from the electron and its bremsstrahlung photons in the ECAL. The energy of the muon is estimated from the momentum of the track. The energy of the neutral hadrons is measured from the calibrated energy in ECAL and HCAL; when the hadrons are charged the track momentum is summed to the energy measured in the calorimeters. The transverse missing energy E_{miss}^T (e.g. of neutrinos) is the negative vector sum of the

transverse momenta of all the reconstructed particles. Figure 3.1 shows how particles interact with the different CMS subdetectors.

3.2 Energy measurement and clustering

As explained in section 2.3, the most important detector for the identification of electrons and photons is the calorimeter, where their energy is measured. In addition for the electrons this information is matched with the one coming from the tracker. The reconstruction of photons is crucial to access the mass of the Higgs through the diphoton channel. Moreover, to tune and calibrate the analysis a well known resonance is needed, hence the electrons from Z decays, reconstructed as if they were photons, are used to measure the trigger, reconstruction and identification efficiencies and the energy scale and resolution.

The energy of the particle is not deposited only in one crystal, it is spread during showering over many crystals, grouped into clusters. Most of the energy of the electromagnetic shower is released in a 5×5 crystals matrix.

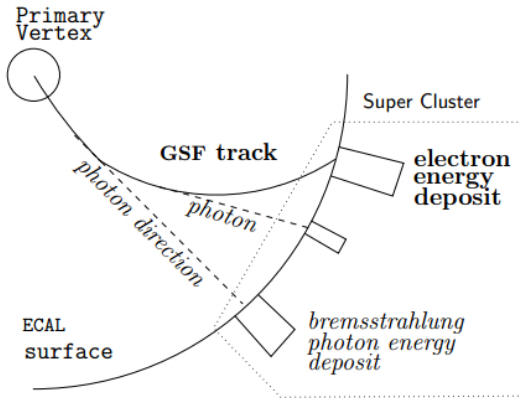


Figure 3.2: Reconstruction of an electron with its bremsstrahlung photons.

In the case of electrons or converted photons some energy may also be lost by bremsstrahlung: about the 33%(68%) of the electron energy is lost at $|\eta| = 0(1.4)$, region of minimum (maximum) material. Since a charged particle is bent in the $r - \phi$ plane by the strong magnetic field, its track is deflected, while its bremsstrahlung photons move straightforward. Hence, there is a spread of the energy in ϕ as shown in figure 3.2, due to the displacement of the electron in the magnetic field with respect to the initial direction of motion. In this case, to collect also the bremsstrahlung photons, dynamical clustering is needed and the ϕ dimension of the cluster is adjusted to collect the secondary deposits of energy.

The photons convert into electrons in the tracker with a probability $P_{conv} = 1 - e^{-\frac{7x}{9X_0}}$. The ϕ dimension of the shower can be used to distinguish converted and unconverted photons. Since the photons do not bend in the magnetic field, the unconverted photons have a better energy resolution. The converted photons show a track, so they might make easier to assign the primary vertex.

3.2.1 Clustering

The algorithm used for the reconstruction of the clusters of energy deposits in ECAL [79] consists in two steps:

1. identification of the seeds, i.e. the local energy maxima which lays above a given threshold;
2. growth of the topological cluster starting from the seed and sum of the energy deposits in the neighboring crystals (that have at least one side in common with a clustered one) where the energy deposit is above a threshold and merging of clusters into super clusters (SC).

At first, the photon energy is estimated as the bare sum of the calibrated energies of the reconstructed hits plus the corresponding energy deposit in the preshower for the events in the EE regions. Then several corrections must be taken into account:

- crystal level correction equalizing the channel-to-channel response variations;
- containment in the clustered crystals and corrections for the energy lost from the converted photons;
- energy scale of the photons measured from the comparison between data and simulation in $Z \rightarrow e^+e^-$ mass distribution;
- gaussian smearing applied to the simulations, extracted from the data-simulation comparison in $Z \rightarrow e^+e^-$ events.

Indeed, as mentioned in section 2.3.3, the energy $E_{e/\gamma}$ is obtained from the amplitude A_i of the signal measured by each channel, corrected for conversion from ADC signal to GeV (G absolute scale), intercalibration (to equalize all the channels) C_i , single channel time response variation $S_i(t)$ (e.g. due to irradiation), containment and geometry effects $F(\eta)$.

$$E_{e/\gamma} \propto F(\eta) \sum_{clusters} C_i \cdot S_i(t) \cdot A_i \quad (3.1)$$

The energy deposit in the preshower is also added, when present.

The clusters are built in different ways in endcaps and barrel. In EB the hybrid algorithm is used to create a window that is small in η and large in ϕ . It starts from the seed, the crystal with the highest peak above the minimum threshold $E_{T,seed}^{min}$. Then, arrays 5x1 above E_{array}^{min} around the seed are added to the cluster in both ϕ directions in N_{steps} steps. Finally, the clusters with energy higher than $E_{seed-array}^{min}$ are grouped in superclusters.

In EE the multi 5x5 algorithm starts from seeds with energy $> E_{T,EEseed}^{min}$ and builds crystals 5x5. Those with energy larger than $E_{T,cluster}^{min}$ are merged in SC, in a range $\eta \pm \eta^{range}$ and $\phi \pm \phi^{range}$.

	Barrel	Endcaps
$E_{T,seed}^{min}$	1 GeV	$E_{T,E_{seed}}^{min}$ 0.18 GeV
$E_{seed-array}^{min}$	0.33 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: Clustering thresholds.

The thresholds are summarized in table 3.1.

In order to reconstruct the particle showers individually, the particle flow (PF) algorithm [80] is used, adding to the seed the neighboring crystals with energy being at least two standard deviations above the electronic noise: 80 MeV in the barrel and up to 300 MeV in the endcaps depending on $|\eta|$.

The energy of the supercluster is computed as the sum of the energies of the clusters. Its position is the mean of the positions of the clusters weighted by their energies. The energy weighted SC position is estimated as $x = \frac{\sum x_i w_i}{\sum w_i}$, where the weights vary logarithmically with the contained energy $\frac{E_i}{\sum E_j}$.

Figure 3.3 shows that the dynamical merging reconstructs the energy better than a simple 5x5 clustering.

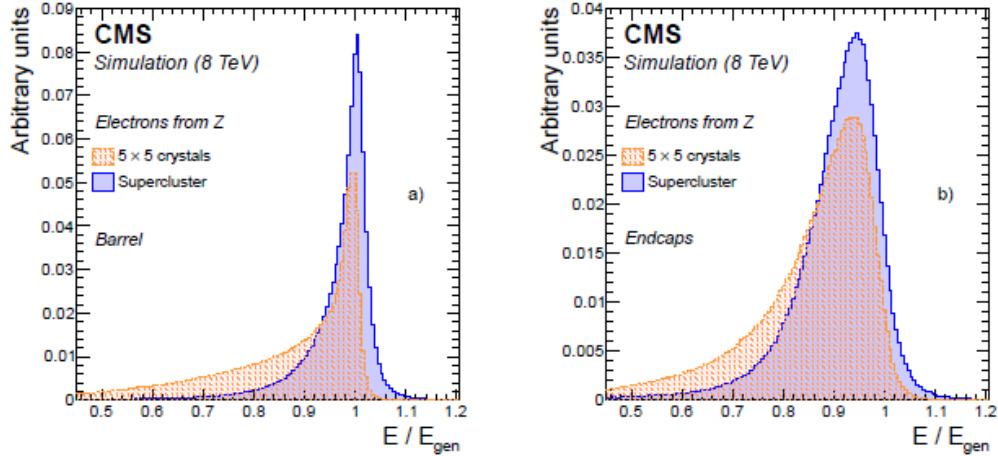


Figure 3.3: Comparison of the reconstructed energy over the generated energy with dynamical and fixed 5x5 clustering.

3.2.2 Track assignment

The electrons and the converted photons have a track that has to be matched to a corresponding energy cluster in ECAL. The tracking can be ECAL-driven or tracker-driven. The tracker-based

algorithm uses the Kalman Filter (KF), equivalent to a global least square minimization, to reconstruct the track from the hits in the innermost layers up to the position of the SC. Since the electrons are subject to bremsstrahlung, the KF is used together with a Gaussian Sum Filter (GSF) that fits the energy loss in each layer as a mixture of Gaussians with different probabilities [81], in order to follow the change of the track curvature.

The ECAL- based tracking, more efficient for $p_T > 10$ GeV, starts from the ECAL information about the clusters, then from the energy weighted position of the SC it extrapolates the track by propagating back in the magnetic field down to the innermost tracker layer. If a compatible hit is found, the vertex is estimated, and from it a new GSF track is propagated (see figure 3.4).

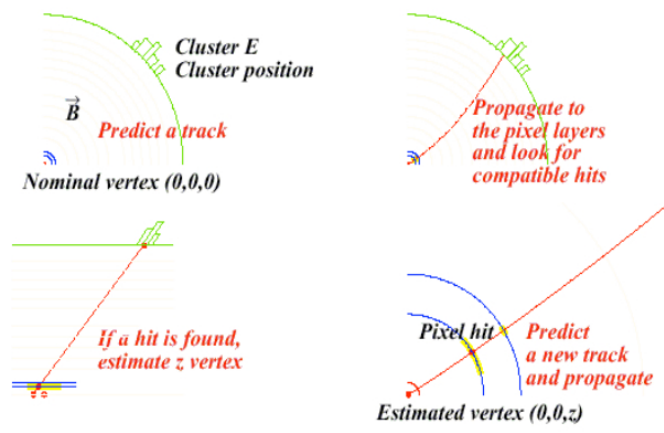


Figure 3.4: ECAL-driven tracking.

The matching between GSF tracks and ECAL superclusters requires

$$|\Delta\eta| = |\eta_{SC} - \eta_{in}^{extrap}| < 0.02$$

$$|\Delta\phi| = |\phi_{SC} - \phi_{in}^{extrap}| < 0.15$$

The two algorithms are complementary and their overall combined efficiency for $Z \rightarrow e^+e^-$ events is approximately 93%.

The converted photons tracks are recognized because of the missing hits in the pixel detector (while for electrons at most one missing hit in the inner layer is accepted) and the conversion vertex can be extrapolated.

3.3 Photons preselection for Higgs analysis

In this and in the following sections, as an example, the main steps of the diphoton candidates selection from the decay of the Higgs boson are illustrated.

The two photons entering the analysis must satisfy a preselection and pass a p_T^γ cut which depends on the invariant mass of the diphoton [78, 79].

The photon preselection consists in an electron veto, cuts on the transverse momentum and loose calorimetric identification. Namely, both the candidates must pass the electron veto: their energy deposit should not be associated to a track or if an associated track is present it must have missing hits in the inner layers, signature of a photon that converted in the tracker. When the electrons from the Z decay are reconstructed as photons in the calibration of the analysis the veto is inverted. The leading photon is required to have $p_T^{\gamma_1} > 30$ GeV and the subleading $p_T^{\gamma_2} > 20$ GeV.

In addition, loose cuts are applied on isolation and shower shape variables:

- the energy-weighted standard deviation of η in the super cluster, $(\sigma_{i\eta i\eta})^2 = |\sum (\eta_i - \bar{\eta})^2 w_i| / \sum w_i$, where η_i is the energy of the i -th crystal, $\bar{\eta} = \frac{\sum w_i \eta_i}{\sum w_i}$ and the weight w_i depends logarithmically on the contained energy $w_i = \max \left[0, 4.7 + \log \left(\frac{E_i}{E_{5x5}} \right) \right]$. It quantifies the lateral extension of the shower along η direction;
- the ratio R_9 of the energy in a 3x3 matrix, centered on the seed, over the SC energy E_{SC} . It quantifies the extent of the shower in ϕ and it is used to separate converted and unconverted photons;
- the ratio $\frac{H}{E}$ of the energy deposit in HCAL (just behind the SC) and ECAL. H is collected in a cone $R=0.15$ around the seed of the SC;
- the photon isolation I_γ : sum of the transverse momenta of the particles identified as photons in a cone of radius $R = 0.3$ around the direction of the photon candidate;
- the track isolation I_{Trk} : sum of the transverse momenta of the tracks with $p_T > 1$ GeV reconstructed in a cone of radius $0.04 < R < 0.3$ around the direction of the photon candidate;
- the charged hadron isolation I_{HCAL} : sum of the transverse energies collected in HCAL inside a cone of radius $R = 0.3$ around the direction of the photon candidate.

The preselection is summarized in table 3.2.

R_9	$\frac{H}{E}$		$\sigma_{i\eta i\eta}$		I_{HCAL}	I_{Trk}
	barrel	endcaps	barrel	endcaps		
≤ 0.9	< 0.075	< 0.075	< 0.014	< 0.034	< 4 GeV	< 4 GeV
> 0.9	< 0.082	< 0.075	< 0.014	< 0.034	< 50 GeV	< 50 GeV

Table 3.2: Photons preselection.

3.4 Photon identification for Higgs analysis

The photon identification is performed on the candidates that passed the preselection [77]. A simple identification method is to apply the selections in table 3.3 on the isolation variables (photons I_γ , charged hadrons I_π and neutral hadrons I_n , with $I_{\pi,n}$ defined like I_γ) and ratio of energy deposit in HCAL. The different working points (loose, medium and tight) have efficiencies for photon identification of about 93%, 84% and 76%, respectively.

		Loose	Medium	Tight
I_γ	barrel	1.3 GeV + 0.005 p_T^γ	0.7 GeV + 0.005 p_T^γ	0.7 GeV + 0.005 p_T^γ
	endcaps	-	1 GeV + 0.005 p_T^γ	1 GeV + 0.005 p_T^γ
I_n	barrel	3.5 GeV + 0.04 p_T^γ	1.0 GeV + 0.04 p_T^γ	0.4 GeV + 0.04 p_T^γ
	endcaps	2.9 GeV + 0.04 p_T^γ	1.5 GeV + 0.04 p_T^γ	1.5 GeV + 0.04 p_T^γ
I_π	barrel	2.6 GeV	1.5 GeV	0.7 GeV
	endcaps	2.3 GeV	1.2 GeV	0.5 GeV
$\sigma_{i\eta i\eta}$	barrel	0.012	0.011	0.011
	endcaps	0.034	0.033	0.031
H/E		0.05		
Electron veto		conversion-safe		

Table 3.3: Photon identification cuts at three different working points.

A more sophisticated analysis can be performed using a Boosted Decision Tree (BDT), a Multi-variate Analysis (MVA) technique [82] trained on simulations.

The inputs to the MVA are [78]:

- topology based variables: $(\sigma_{i\eta i\eta})^2$, the off-diagonal elements of the covariance matrix $cov(i\eta, i\phi)$, the S_4 ratio of the energy deposit in a 2x2 matrix over that in the SC, H/E and the energy weighted second moments σ_η , σ_ϕ and $\sigma_{RR} = \sigma_x \oplus \sigma_y \oplus \sigma_{preshower}$;
- photon, track and charged hadrons isolation variables;
- energy density per unit area ρ to account for the pileup;
- photon kinematic observables, like η and E .

In output the MVA associates to each candidate a score between -1 and +1 and it is used to distinguish the signal (high score) from the background (low score). In the signal selection both the photons are required to have BDT output > -0.2 . This loose selection keeps the 99% of the signal and rejects about the 35% of the background. The distribution of the photon MVA output for data and MC simulations in EB is shown in figure 3.5 [83].

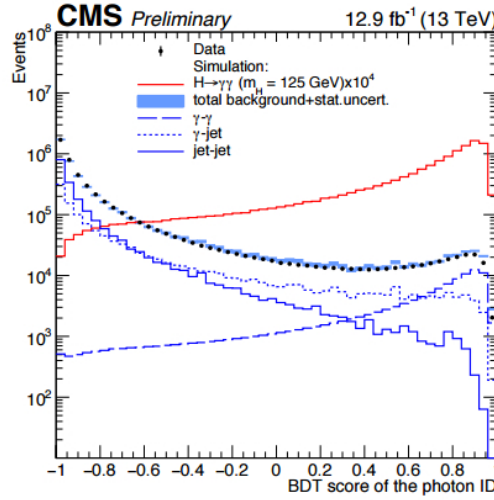


Figure 3.5: BDT score of the lower-scoring photon of diphoton pairs with an invariant mass in the range $100 < m_{\gamma\gamma} < 180$ GeV, for events passing the preselection in the 13 TeV dataset (black points), and for simulated background events (blue histogram). The sum of all background distributions, generated at leading order, is scaled up to data. The red histogram corresponds to simulated Higgs boson signal events.

3.5 Pileup

The pileup events are events other than the signal one from the main interaction. They can have origin in multiple bunch crossings or simultaneous interactions, hence two kinds of pileup are distinguished.

The **in-time pileup** is an effect of the high luminosity of LHC that induces additional proton-proton collisions that take place during the same bunch crossing as the collision of interest. At 8 TeV during 2012 the LHC instantaneous luminosity reached $7 \cdot 10^{33} \text{ cm}^{-2} \text{ sec}^{-1}$ and the number of superimposed primary vertices per bunch crossing were measured to be on average about 20. During 2016, the instantaneous luminosity is expected to increase up to $2 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ and the number of pile-up events may be up to 70 per bunch-crossing. Indeed, for the Run2, with the central collision energy of 13 TeV, and the time separation of 25ns between bunches, it was expected a luminosity such that on average the number of pileup per bunch crossing is around 40 [84]. In addition, there are soft in-time interactions of the partons which are not involved in the hard scattering. Most of these events are easily discarded with a cut on the charged particles momentum.

The **out-of-time (OOT) pileup** are events overlaid from different bunch crossings. Due to the integration time of the electronics the signal pulses last more than the 25 ns that separates two subsequent bunches in the train. High pileup will overwhelm the low energy pulses making difficult to determine the correct bunch crossing assignment from the pulse shape alone.

Pileup is not measured by itself, but it is estimated from the number of reconstructed pri-

mary vertices in the event (the quality of the estimate depends on the vertex reconstruction efficiency). In picture 3.6 an example of the distribution of the number of vertices is plotted for $Z \rightarrow w^+e^-$ events at 13 TeV [85]: black dots for the data, blue histogram for the pileup-reweighted simulations.

The OOT pileup can be mitigated using time information. The same approach is much more difficult to apply on in-time pileup since an incredible resolution would be required to distinguish the timing of each pp interaction inside the same bunch crossing. To maintain a low occupancy high granularity and thin pixel detector are envisaged.

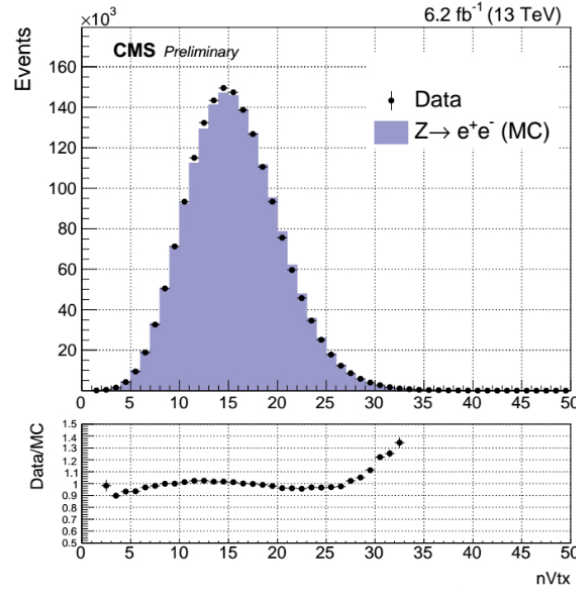


Figure 3.6: Distribution of the number of reconstructed vertices.

3.6 Vertex identification for Higgs analysis

Since several interactions may take place at each bunch crossing, the determination of the primary vertex associated with an Higgs boson decaying in two photons is critical. The vertex assignment, together with the measured photon energy resolution, affects the mass resolution [78, 79].

The selection of the vertex is made through a BDT multivariate analysis based on

- the sum of the transverse momenta squared of each charged track associated to the given vertex over all the particle flow PF candidates, $\sum_i |\vec{p}_T^i|$;
- the balance between the track momentum and the momentum of the diphoton system in the transverse plane, $-\sum_i \left(\vec{p}_T^i \cdot \frac{\vec{p}_T^{\gamma\gamma}}{|\vec{p}_T^{\gamma\gamma}|} \right)$;

- the asymmetry between the momentum of the charged track and the diphoton system, $(|\sum_i \vec{p}_T^i| - p_T^{\gamma\gamma}) / (|\sum_i \vec{p}_T^i| + p_T^{\gamma\gamma})$.

The output of this BDT is used in the classification (in section 3.8). If one of the photons is associated to a conversion track the pull $g_{conv} = |z_{vtx} - z_e|/\sigma$ is introduced. It corresponds to the difference between the longitudinal position of the reconstructed vertex z_{vtx} and the vertex location estimated using conversion tracks z_e , weighted on the z_e uncertainty.

Another BDT is employed to estimate the vertex finding efficiency, defined as the per-event probability that the assigned vertex is within 1 cm from the true vertex. This multivariate analysis technique is trained on simulated $Z \rightarrow \mu^+ \mu^- \gamma$ events and it receives in input

- the number of vertices in each event;
- the BDT output for the three most likely vertices of each event;
- the distance between the chosen vertex and the second and third best choices;
- $p_T^{\gamma\gamma}$;
- the number of converted photons.

The efficiency to identify a vertex closer than 1 cm from the true one, integrated over the full p_T spectrum of the Higgs, is 85.4% at 8 TeV. In these conditions the mass resolution is dominated by the photons energy resolution.

As shown in figure 3.7 the distributions of the correct and the misaligned vertices are clearly distinct. The correct vertices are shown in indigo for the simulation and in white circles for the data. The misaligned vertices are in white (with red contour) for the simulation and in black circles for the data.

3.7 Diphoton selection for Higgs analysis

After the steps described above, the diphoton candidates are identified, as two photons that pass the p_T^{γ} trigger, scaled for the invariant mass of the diphoton system to avoid distortions at the low end of the $m_{\gamma\gamma}$ spectrum: $p_T^{\gamma 1} > m_{\gamma\gamma}/3$ and $p_T^{\gamma 2} > m_{\gamma\gamma}/4$ with an efficiency above 95%. Their invariant mass is also requested to be in a range between 100 and 180 GeV and if multiple diphoton candidates are found for the same event the one with largest $p_T^{\gamma 1} + p_T^{\gamma 2}$ enters the analysis.

Then an MVA is used on these diphoton systems to separate those produced by the Higgs from the background [78, 79].

The inputs of the BDT are:

- $p_T^{\gamma 1, \gamma 2} / m_{\gamma\gamma}$. The relative momenta are used to have an analysis independent on $m_{\gamma\gamma}$;
- η coordinate of the SC of both the photons;

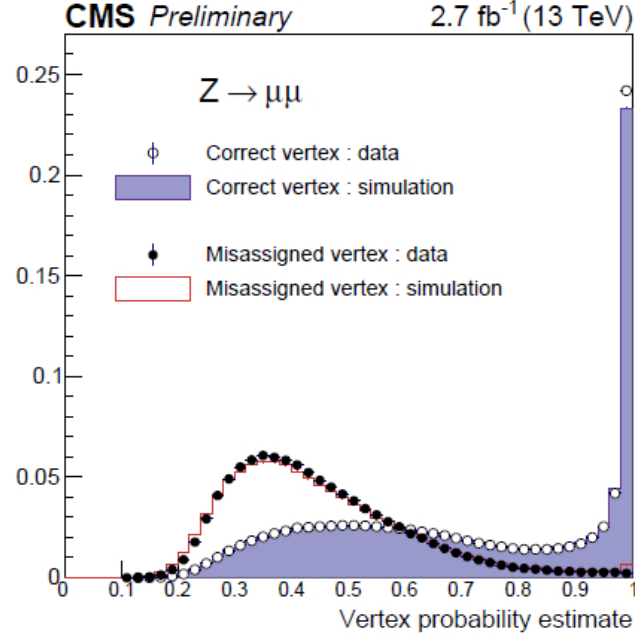


Figure 3.7: Normalized distributions of the per-event probability to assign the vertex within 1 cm from the true one in $Z \rightarrow \mu^+ \mu^- \gamma$ events.

- $\cos(\phi_1 - \phi_2)$ (of the angle between the two photons);
- relative diphoton mass resolution for correct vertex assignment, $\sigma_m^{right}/m_{\gamma\gamma}$;
- relative diphoton mass resolution for incorrect vertex assignment, $\sigma_m^{wrong}/m_{\gamma\gamma}$;
- per-event probability that the right vertex was used in the reconstruction;
- BDT score from the identification of both the photons.

The output of the diphoton BDT is a score between 0 and 1. Hence the diphoton candidates are split in the following categories (for Run1) to enhance the sensitivity of the analysis:

- $[0, 0.104]$ discarded
- $[0.104, 0.428]$ untagged3
- $[0.428, 0.753]$ untagged2
- $[0.753, 0.983]$ untagged1
- $[0.983, 1.0]$ untagged0

The class with lowest diphoton BDT output is discarded because it contains almost only background. The one with highest score selects unconverted photons, hence, good signal that assures the best resolution on the estimation of the Higgs mass. In figure 3.8 the distribution of

the diphoton MVA output is shown. The contributions from different backgrounds are also shown [83].

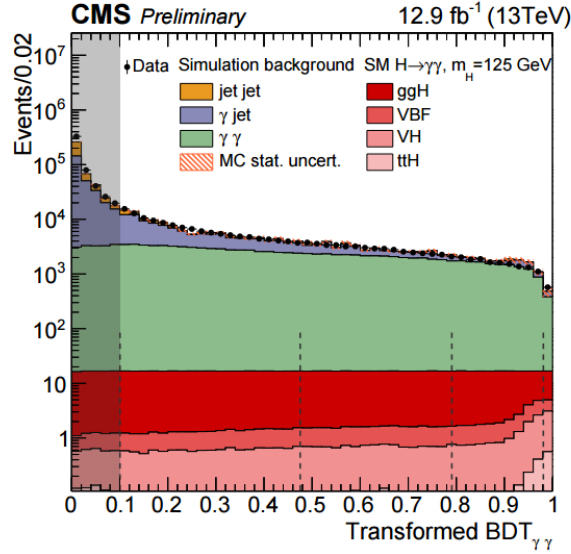


Figure 3.8: BDT diphoton classifier score in data (black points) and simulation (stacked histograms) for events in the region $100 < m_{\gamma\gamma} < 180$ GeV. The contributions from different background processes are shown.

3.8 Tags and classification for Higgs analysis

The events are categorized [78, 79] beginning from those that present some signatures of a tagged production mechanism in the final state. First the events produced through the more rare mechanisms, $t\bar{t}H$ and VH , are selected. Then if dijets events that did not pass the previous selection are tagged the events are classified as VBF. The remaining inclusive events (99%) are split using the diphoton multivariate classifier as explained in section 3.7. The selection criteria for each class are summarized in table 3.4.

3.9 Signal and background models for Higgs analysis

The distributions of the data acquired by CMS and selected as promising diphoton candidates must be compared with the theoretical expectations for a SM Higgs boson. Indeed, a signal model [79] is developed, based on simulations with next-to-leading order (NLO) matrix-element generators. The Higgs boson mass shape is modeled using MC events, corrected for tuning, scales, efficiencies and resolution after comparison to the data in $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-\gamma$.

The $m_{\gamma\gamma}$ distribution of the simulated events is fitted individually in each category and for

Label	N° of classes		Main requirements
	7 TeV	8 TeV	
$t\bar{t}H$ lepton tag	-	1	$p_T^{\gamma_1} > m_{\gamma\gamma}/2$ 1 b-tagged jet + 1 electron or muon
VH tight ℓ tag	1	1	$p_T^{\gamma_1} > 3m_{\gamma\gamma}/8$ [e or μ , $p_T > 20$ GeV, and $E_T^{miss} > 45$ GeV] or [2e or 2 μ , $p_T^\ell > 10$ GeV; $70 < m_{\ell\ell} < 110$ GeV]
VH loose ℓ tag	1	1	$p_T^{\gamma_1} > 3m_{\gamma\gamma}/8$ e or μ , $p_T > 20$ GeV
VBF dijet tag 0-2	2	3	$p_T^{\gamma_1} > m_{\gamma\gamma}/2$ 2 jets; classified using combined diphoton-dijet BDT
VH E_T^{miss} tag	1	1	$p_T^{\gamma_1} > 3m_{\gamma\gamma}/8$ $E_T^{miss} > 70$ GeV
$t\bar{t}H$ multijet tag	-	1	$p_T^{\gamma_1} > m_{\gamma\gamma}/2$ 1 b-tagged jet + 4 more jets
VH dijet tag	1	1	$p_T^{\gamma_1} > m_{\gamma\gamma}/2$ jet pair, $p_T^j > 40$ GeV and $60 < m_{jj} < 120$ GeV
untagged 0-4	4	5	The remaining events, classified using diphoton BDT

Table 3.4: Properties of fermions.

each production process with an analytic function built as a sum of Gaussians (at most five). The distributions for events with wrongly and correctly assigned vertices are managed separately. In fact, as previously mentioned, the vertex finding affects and changes the invariant mass distribution. The obtained parameters are interpolated between each mass point. Finally, the functions of each category and mode are summed together, weighted by the relative cross section of the process.

The background contribution (referred as the continuum) is made of prompt photons (irreducible background) and of a reducible background due to events where at least one jet has been reconstructed as a photon (γ -fake, fake-fake), as depicted [86] in figure 3.9.

The modeling of the background [78] is entirely data driven. A few families of reasonable functions are chosen (e.g. polynomials, sum of exponentials). Then, in each class the data are fit in the range $110 < m_{\gamma\gamma} < 180$ GeV with functions of the same family at increasing degrees of freedom. When the improvement in the F-test (where the test statistics has Fisher-Snedecor distribution in the null hypothesis [87]) moving from N to N+1 degree is negligible, the function with N degrees of freedom (e.g. polynomial of Nth order, sum of N exponentials) is selected as the “truth” function for that family. At that point, the “truth” is used to create a toy MC that will be “fit” with the same functional family. In each class the potential bias on the estimate of the signal is computed for each “truth-fit” couple. The one with the largest

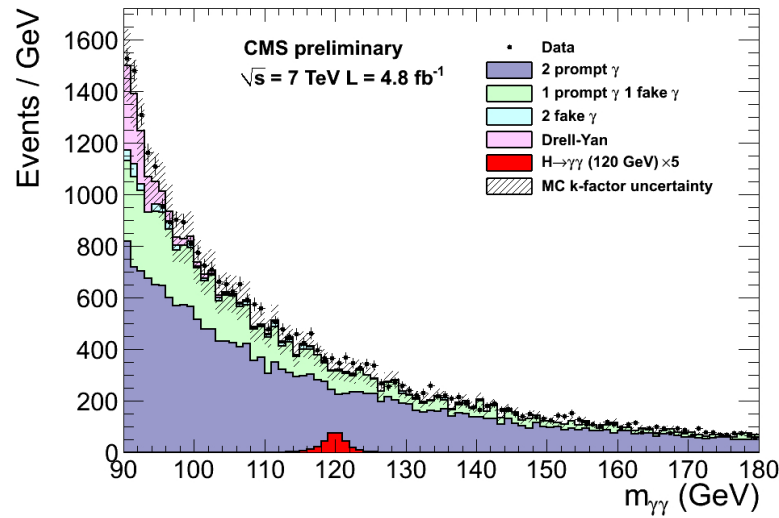


Figure 3.9: Diphoton mass distribution for data (data points), and Monte Carlo simulation of SM processes which constitute the background to the search (histograms). In red a simulated signal for a Higgs boson with a mass of 120 GeV is shown.

potential bias on the fitted signal strength describes the final model, under the condition that the bias is less than five times the statistical uncertainty on the background.

3.10 Systematic uncertainties for Higgs analysis

All the measurements and the theoretical predictions introduce systematic uncertainties on the estimate of the mass of the Higgs boson. The main sources of systematics are [78, 83, 88]:

- theory [25]:
 - PDF uncertainties (typically less than 2% depending on the category);
 - factorization and renormalization scale (about 7%);
 - α_S uncertainty (at most 3.7%);
 - underlying event and parton shower uncertainty (depending on the choice and tuning of the generator);
 - QCD scale uncertainty (a few percents depending on the process, typically around 5-10%);
 - uncertainty on the branching ratio of the channel $H \rightarrow \gamma\gamma$ (estimated 2.08%);
- signal model:
 - integrated luminosity, estimated from the data (2.2% at 7 TeV, 2.6% at 8 TeV and 6.2% at 13 TeV);

- trigger efficiency, measured using $Z \rightarrow e^+e^-$ events (slightly less than 0.1%);
 - photon preselection, uncertainty on $\epsilon_{data}/\epsilon_{MC}$ (from 0.1% to 2.3% depending on the category);
 - photon energy scale due to the difference between electrons and photons: the electrons from Z boson decays are reconstructed like photons and used to validate the calibration of the analysis to be applied to the photons, but their interaction in the material in front of ECAL is different and not perfectly modeled (effect from 0.03% in the central barrel to 0.3% in the outer endcap). In addition, there is a variation of the fraction of scintillating light that reaches the photodetector depending on the longitudinal depth of the emission in the crystal (0.015%);
 - energy scale nonlinearity, determined from the difference between data and Monte Carlo in the extrapolation of the energy scale from $Z \rightarrow e^+e^-$ to $H \rightarrow \gamma\gamma$ (0.1-0.2%);
 - non-uniformity of light collection: uncertainty on the response of the ECAL crystals, estimated to be 0.07%;
 - estimate of the energy scale, smearing and resolution with $Z \rightarrow e^+e^-$, corrections due to imperfect agreement between data and MC (below 0.5%);
 - photon identification BDT score uncertainty on the ratio $\epsilon_{data}/\epsilon_{MC}$, due to imperfect simulation of the shower shape (4%), and estimate of the per photon energy resolution estimate due to imperfect simulation of the shower shape (up to 10%);
 - photon preselection efficiency, measured in the data over MC preselection efficiency scale factors from $Z \rightarrow e^+e^-$ (0.05-0.1%);
 - vertex finding efficiency, uncertainty from the measurement of the data/MC scale factor in $Z \rightarrow \mu^+\mu^-\gamma$ (about 1.5%);
 - diphoton efficiency, diphoton mass scale, and diphoton mass resolution propagated from the single photon quantities;
- classification with tags:
 - dijet tagging efficiency, probability that additional jets are produced in ggH, probability in ttH that a gluon converts in a $b\bar{b}$ couple, rejection of jets from pile up (10-30%);
 - jet energy scale and smearing corrections (migration of 4-15% from the scale and 2% from the resolution);
 - b-tagging efficiency: b-tagging efficiency scale factors in data and MC (2% in leptonic category and 5% in hadronic-tagged class);
 - E_T^{miss} selection efficiency (4-15%);
 - leptons identification uncertainty, in data/simulation identification efficiency scale (0.2-0.5%);
 - background modeling and uncertainty on the choice of the function that describes better the background.

Chapter 4

Detector studies

4.1 Study of the material budget in front of ECAL

The material in front of ECAL can strongly affect the energy resolution of the calorimeter, hence the resolution on the Higgs mass, as mentioned in section 2.3.4. For this reason it is important to know in detail the structure of the tracker and to implement it in the simulations, in order to describe the actual data.

The material of CMS upstream of the calorimeter is studied by means of electrons from the decay of Z bosons, created in the pp collisions at LHC. Data from Run2 have been used in the following analysis.

Standard selections have been applied on the electron candidates to assure the rejection of most of the background events:

- two opposite charged electron candidates, reconstructed using the ECAL driven algorithm (section 3);
- electron ID medium selection, based on kinematics, shower shape and isolation [76];
- transverse momentum $p_T > 25$ GeV for each electron candidate;
- mass of the Z boson, reconstructed as the invariant mass of two electron candidates, in the range $70 \text{ GeV} < M_Z < 110 \text{ GeV}$;

The same selections have been applied to the simulated events with the goal of investigating the agreement between the distributions of the data and the Monte Carlo (MC) events.

The variable f_{brem} has been studied, which represents the relative energy lost by an electron in the tracker, equal to the difference between the initial momentum (at the point of closest approach to the vertex) and the final momentum (obtained from a fit of the momentum measurement in the last 3 track points), divided by the initial momentum:

$$f_{brem} = \frac{p_i - p_f}{p_i}.$$

The average remaining energy is defined as

$$\langle E(x) \rangle = E_0(1 - \langle fbrem \rangle) = E_0 e^{-\frac{x}{X_0}},$$

and from this quantity it is possible to estimate the material budget in terms of radiation lengths using the following equation:

$$\frac{x}{X_0} = -\log(1 - \langle fbrem \rangle). \quad (4.1)$$

The number of interactions that take place in each bunch crossing is not the same in data and simulations and the distributions of the number of vertices do not overlap perfectly: the pileup is accounted for in the MC, but the distribution of the number of pileup events in simulation and events is slightly different and the simulations must be reweighted in order to properly describe the data. In practice, a weight is assigned to each simulated event to compensate for this difference, such that the distributions of the number of interactions in data and MC become identical.

During Run1 some differences in the material budget were observed between data and MC and the simulations were corrected. This work is meant to verify how good the agreement is after this improvement. At first approximation, the difference between x/X_0 measured in data and simulations is entirely attributed to a mismodeling of the material.

In figure 4.1 $fbrem$ as a function of the pseudorapidity η (top plots) and of the azimuthal angle ϕ (bottom plots) both for the data (left) and for the simulation (right) is plotted for electron candidates.

In figure 4.2 is the ratio between data and simulation for $fbrem$ as a function of η (top) and $fbrem$ versus ϕ (bottom). The plots are in quite good agreement, but there are regions which are not well described in the MC. In fact we can notice some discrepancies:

- in η at the gap between the EB and the EE (in $1.4442 < |\eta| < 1.566$);
- in ϕ around 1.8, probably due to mismodeling of the supports of the tracker.

The support structure of the tracker, briefly described in section 2.2.2, is shown in figure 2.9. Using equation 4.1 the material budget is computed and it is mapped in η versus ϕ in figure 4.3, that indeed shows roughly the structure of the tracker. This result is in agreement with the previous one in figure 4.4a [89], made using simulated neutrinos on Run 1.

In figure 4.3 (right) the ratio between the mapping with data and with MC is also shown. An excess of material in the data with respect to simulation is evident only in the limit regions of the tracker at high η .

In figure 4.5a the profiles of the material budget in η for the data (top) and the MC (bottom) are shown. They are compared in figure 4.5b. The material goes from a minimum of $0.3X_0$ in the center of the barrel to a maximum of $1.2X_0$ around the gap between EE and EB. The main

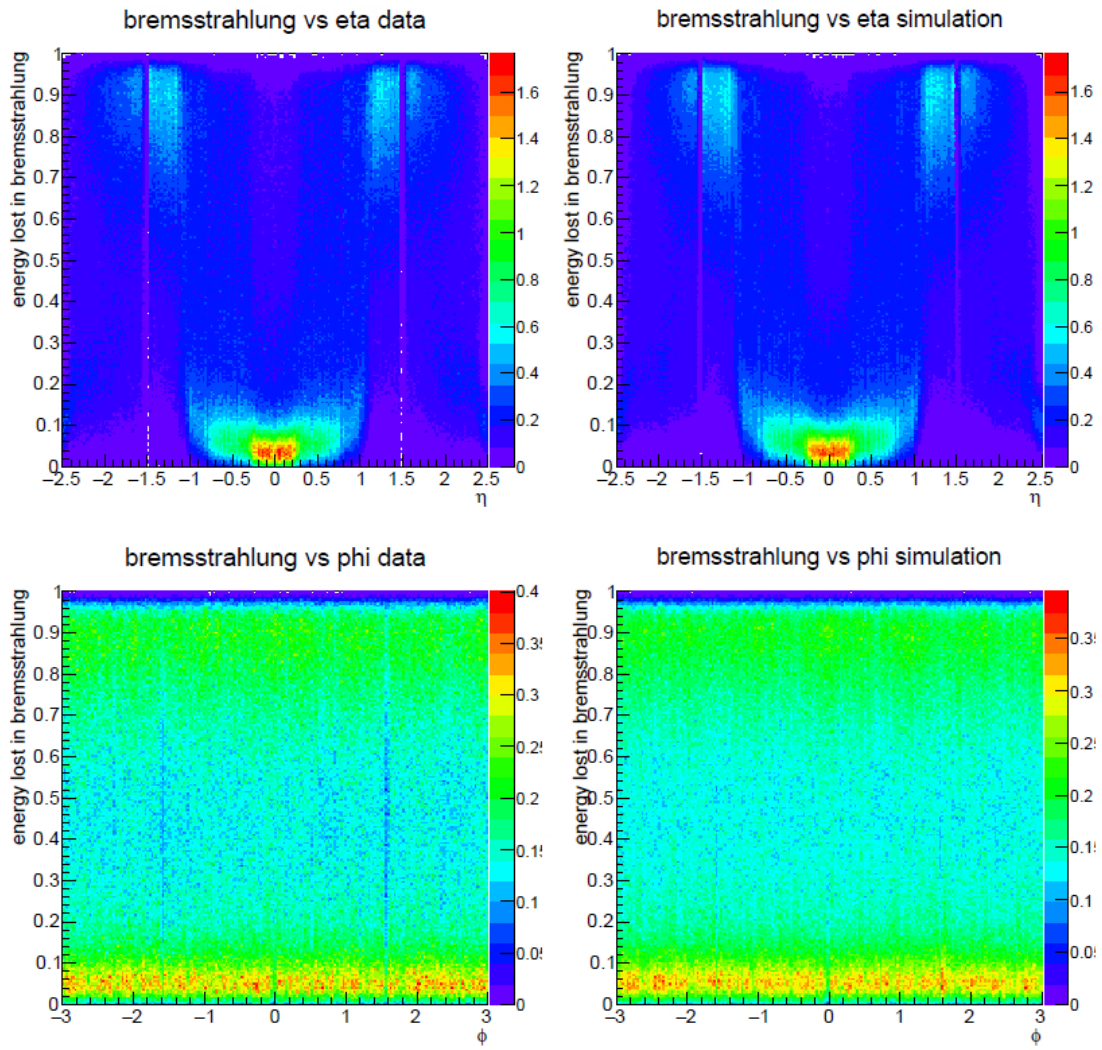


Figure 4.1: Plot of the energy lost by the electron by bremsstrahlung as a function of the coordinate η (top plots) and of the coordinate ϕ (bottom plots) both for the data (left) and for the simulation (right).

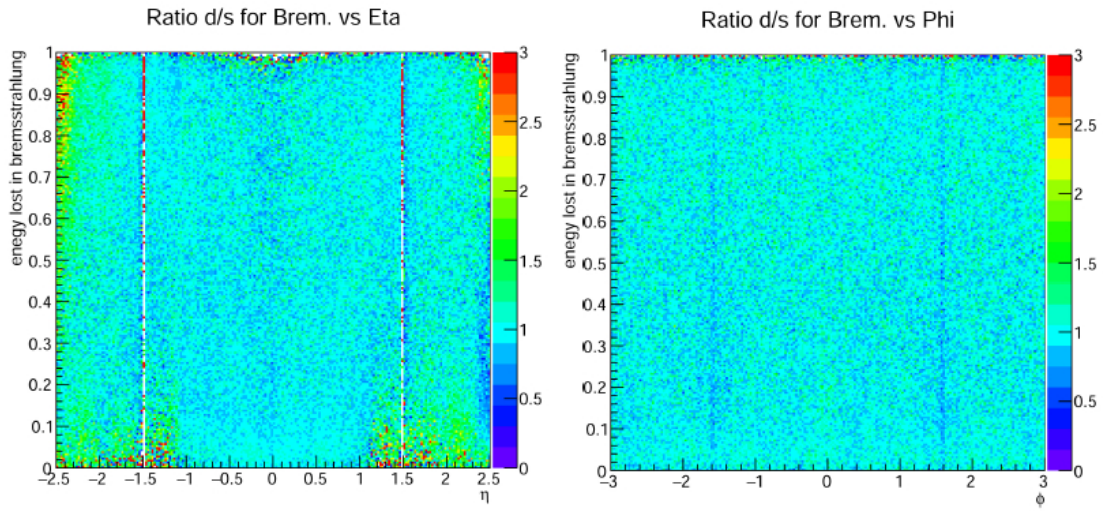


Figure 4.2: Ratio between data and simulation for f_{brem} as function of η (top) and of ϕ (bottom).

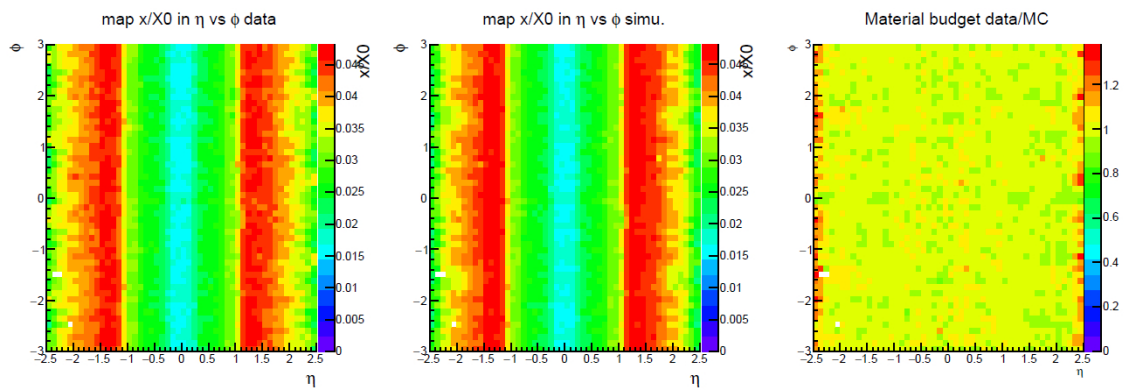
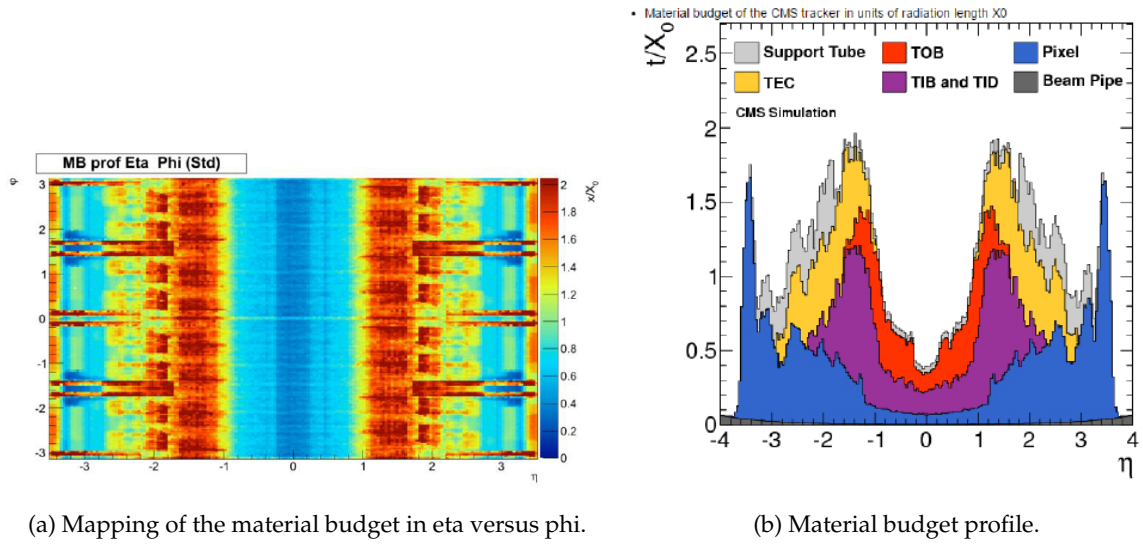


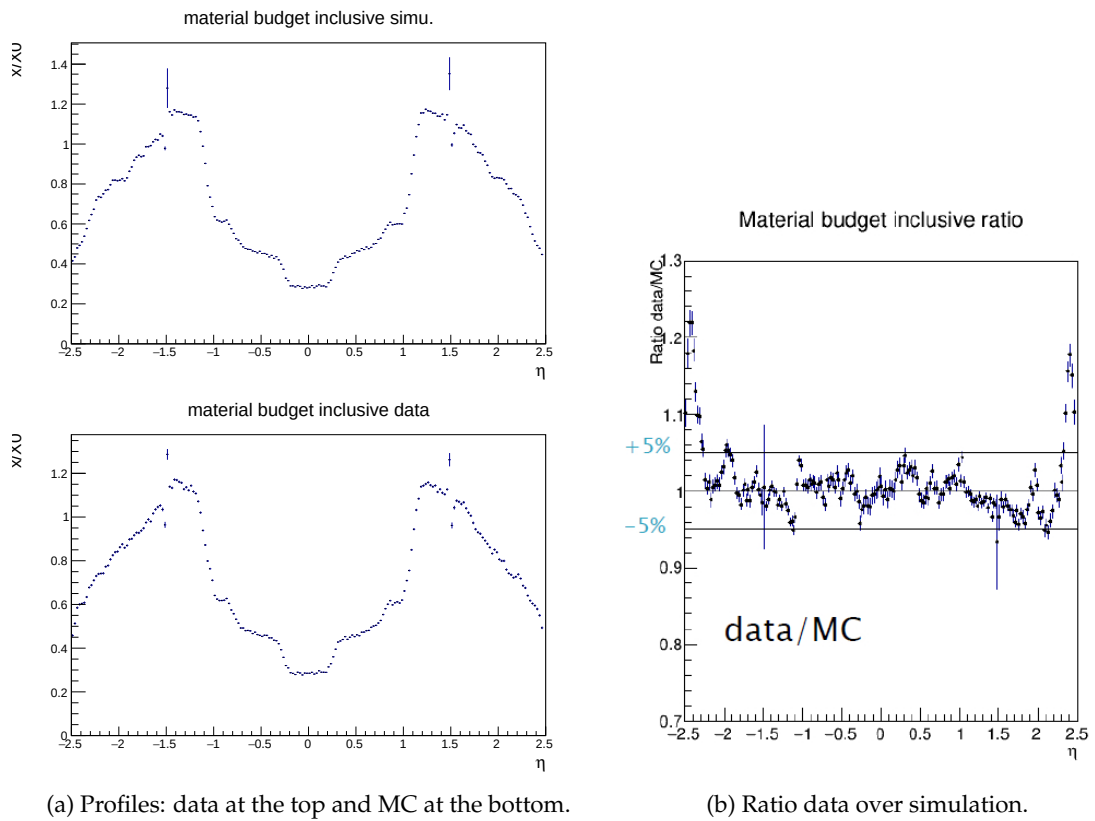
Figure 4.3: Mapping of the material budget in η versus ϕ for data (left) and simulation (center) and their ratio (right).



(a) Mapping of the material budget in eta versus phi.

(b) Material budget profile.

Figure 4.4: Studies of the material upstream ECAL with simulated neutrinos on Run 1.



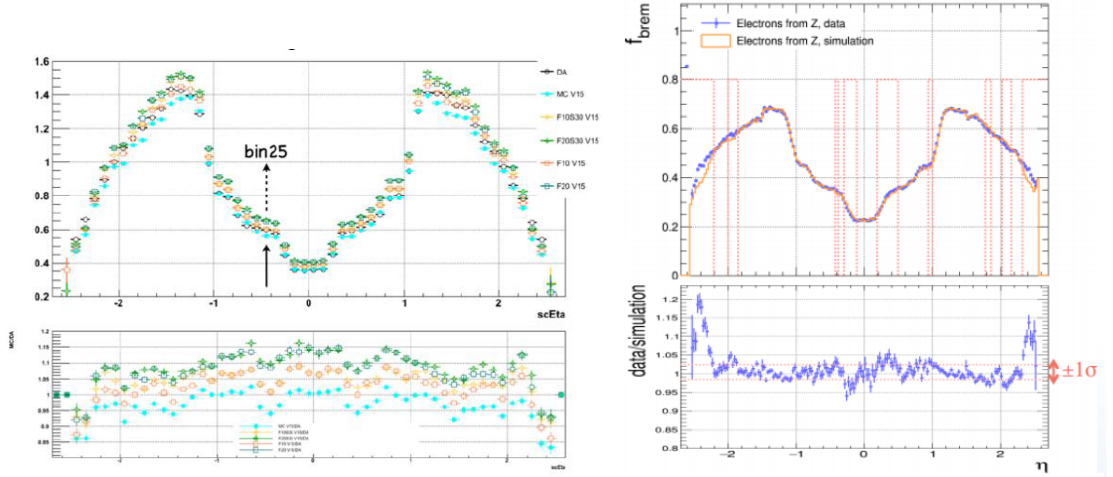
(a) Profiles: data at the top and MC at the bottom.

(b) Ratio data over simulation.

Figure 4.5: Profiles of the material budget in η for data and MC and their ratio.

contribution to the material is given by the support structure and the TECs.

There is very good agreement in $-2 < \eta < 2$ and a discrepancy of about the 20 % at high η , where the tracking efficiency degrades (see figure 2.8). The previous results on Run 1 simulations with neutrinos [89], in figure 4.4b, show the same shape, in agreement also with other studies [90] of the same period in figure 4.6a, which are made on Run1 exploiting the variable f_{brem} . The data are displayed in black. Minimum material budget of $0.4X_0$ and maximum of $1.5X_0$ are found. There is a discrepancy ranging from $0.1 X_0$ at $\eta = 0$ to $0.3X_0$



(a) Profiles of the material budget on Run 1 and their ratio.(b) Profiles of the material budget on Run 2 and their ratio.

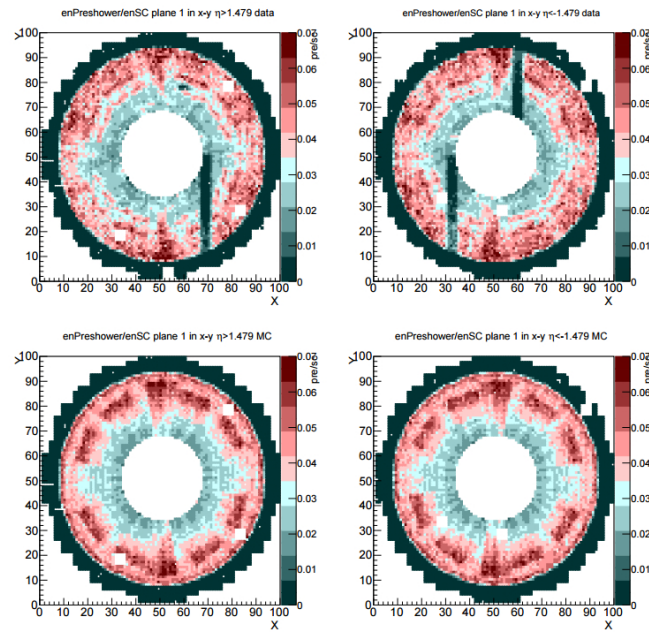
Figure 4.6: Profiles of the material budget in η for data and MC, and their ratio for Run 1 (left) and for Run 2 (right).

at $|\eta| = 1.5$ between the value of the minimum and maxima of the profile of x/X_0 on Run2 dataset and those obtained on Run1. Another recent study [91] on the data of Run2 in figure 4.6b shows peaks at 0.7 in f_{brem} that correspond exactly to the one obtained by this analysis at $1.2 X_0$. This may be due to a change of the tracking algorithm from Run1 to Run2, that changed the sensitivity to the material in front of ECAL.

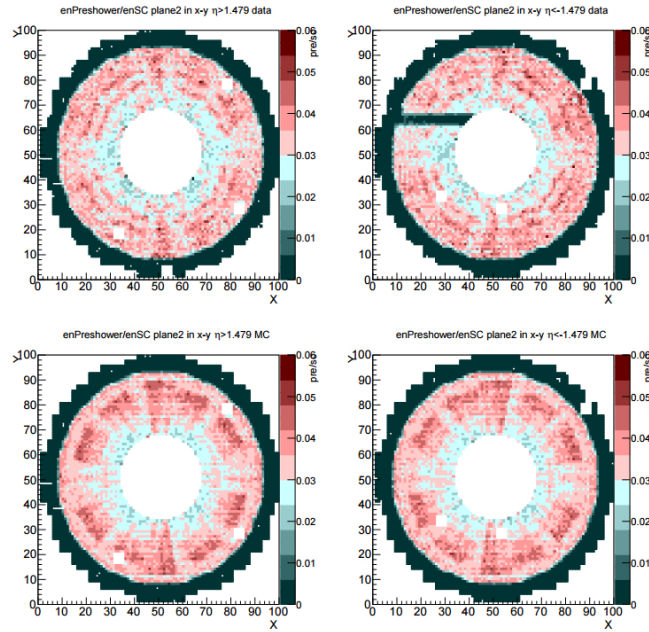
In order to better understand the structure in front of the endcaps, the ratio between the energy deposit in the preshower and the energy in the super cluster for both the endcaps, at $\eta > 1.497$ (EE_+) and $\eta < -1.497$ (EE_-), using data and MC events, has been mapped in X versus Y. It gives information about the material between the preshower layers.

In figure 4.7a the mapping for the deposit in the first plane of the preshower is shown, in figure 4.7b for the second plane and in figure 4.8 just in front of the super cluster.

In figure 4.9 the ratio between the mapping in data and simulation is shown. This ratio shows rough agreement between simulations and data, but in the data some structures can be identified that are not implemented in the MC: these correspond to the boxes that



(a) Mapping of the ration between the energy deposit in the first plan of the preshower and the energy of the supercluster for EE+ on the left and EE- on the right, with data at the top and MC at the bottom.



(b) Mapping of the ration between the energy deposit in the second plan of the preshower and the energy of the supercluster for EE+ on the left and EE- on the right, with data at the top and MC at the bottom.

Figure 4.7: Energy deposit in the first and second layers of the preshower, relative to the energy of the corresponding SC.

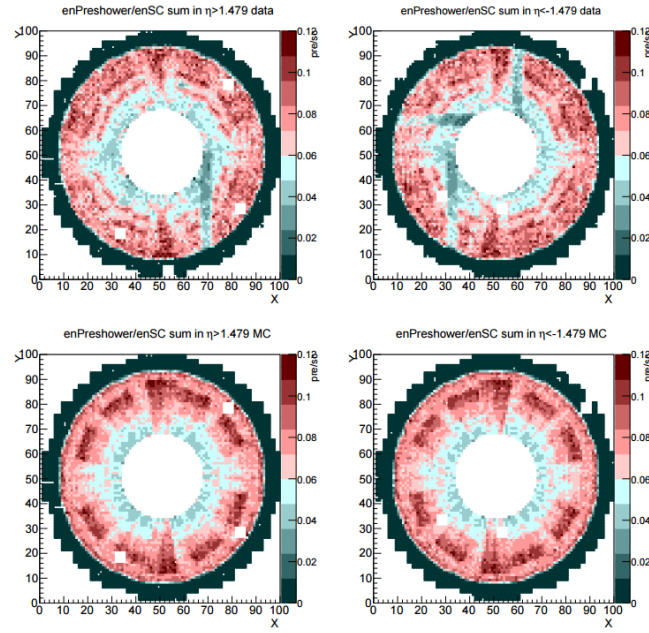


Figure 4.8: Mapping of the ratio between the energy deposit in the the preshower (just in front of the SC) and the energy of the supercluster for EE+ on the left and EE- on the right, with data at the top and MC at the bottom.

contain the detector cables (the dark rectangular regions).

In conclusion, the discrepancies between the simulated material budget and the real one are negligible in the fiducial region up to $|\eta| = 2$, while at higher pseudorapidities there is more material in the detector than in the model, up to the 20%. The region above $|\eta| = 2$ is used in the analysis of the Higgs taking into account that here the coverage of the tracker is less efficient and some discrepancies may arise between MC and data. Thanks to this study of the material budget and to the measured agreement between data and MC the effect of the material on the systematics on the energy scale of the mass of the Higgs boson in $H \rightarrow \gamma\gamma$ have been observed to amount to at most the 0.17% [83]. The incidence on the systematics have been halved with respect to Run1 when the effect of the material on the energy scale reached the 0.3% [78].

It is then important to study shower shape variables which may be affected by this small material mismodeling. In addition, a study in absence of the magnetic field can be exploited to check if eventual discrepancies between data and MC are due to material effects [92].

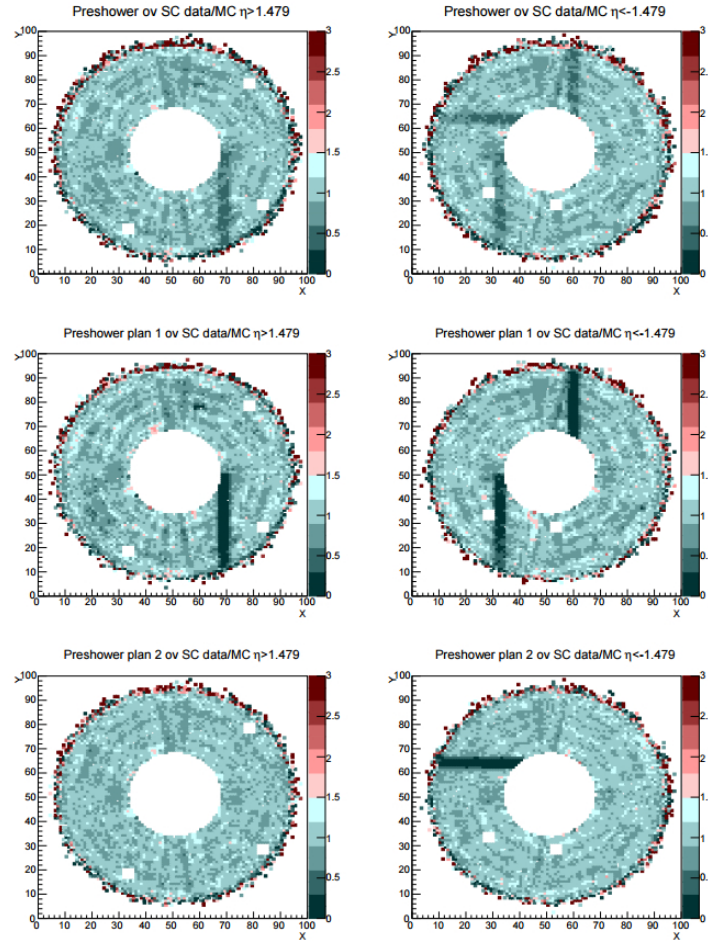


Figure 4.9: Ratio data over MC of the mapping of the preshower energy (plane 1 at the top, plane 2 in the center, sum at the bottom) over the super cluster energy for EE+ (left) and EE- (right).

4.2 Shower shape studies

Some of the most relevant variables related to the electromagnetic shower shapes in the calorimeter are compared in this section. Variables sensitive to material effects are chosen:

$R_9 = E_{3 \times 3} / E_{SC}$, the ratio between the energy deposited in a 3×3 crystal matrix around the crystal seed of the SC and the energy deposited in the corresponding supercluster;

$S_4 = E_{2 \times 2} / E_{5 \times 5}$, ratio between the energy deposited in a 2×2 crystal matrix around the seed and in a 5×5 crystal matrix defined in the same way;

$S_5 = E_{1 \times 5} / E_{5 \times 5}$, defined as S_4 for a 1×5 crystal matrix;

the η -dimension of the supercluster $\sigma_{i\eta i\eta}$,

$\Delta\eta$ between the track and the supercluster.

The distributions are normalized to the number of events and a standard selection is performed on the events that pass the electron veto and have unitary charge. The electrons of this study are asked to be products of the Z boson decay, i.e. with invariant mass $70 < m_{ee} < 110$ GeV. The transverse momentum thresholds are chosen to be 32 GeV for the leading electron and 25 GeV for the subleading one. The algorithm of reconstruction is the ECAL driven one. The gap between EB and EE is excluded. The data have been collected during Run2 while the simulations are performed with the events generator MadGraph5_aMC@NLO [93].

In figures 4.10 and 4.11 some shower shape variables are compared at 3.8 T (in blue) and 0 T (in red). In absence of the magnetic field the electrons are not bent and decelerated, the material effects due to bremsstrahlung loss are avoided, hence reduced spread of the cluster, i.e. better energy resolution, is expected. Other variables, like f_{brem} or $\Delta\eta$, that bring information from both ECAL and the tracker, are not considered in the comparison because their distribution is strongly affected by the misalignment caused by the change of magnetic field (the tracker and ECAL are carefully aligned only at 3.8 T).

The distribution of the ratios of energies in different crystal matrices, $E_{1 \times 5} / E_{5 \times 5}$ (figures 4.10a and 4.10b), $E_{2 \times 2} / E_{5 \times 5}$ (figures 4.10c and 4.10d) and $E_{3 \times 3} / E_{5 \times 5}$ (figures 4.10e and 4.10f), clearly exhibit the expected reduction of energy spread at 0 T. The field effect is stronger in EE, where more material is crossed by the electrons and the distributions are less peaked at high values. The bending due to the magnetic field takes place in the $\phi - r$ plane, so the variables with dependence only in η are expected to remain almost unchanged. Indeed, the energy-weighted standard deviation $\sigma_{i\eta i\eta}$ of the dimension of the supercluster that quantifies the spread of the SC in η , in figure 4.11b, has the same shape and width at 0 T and 3.8 T. The small change of the distribution in EB, in figure 4.11a, comes from the weak focusing effect of the magnetic field in η .

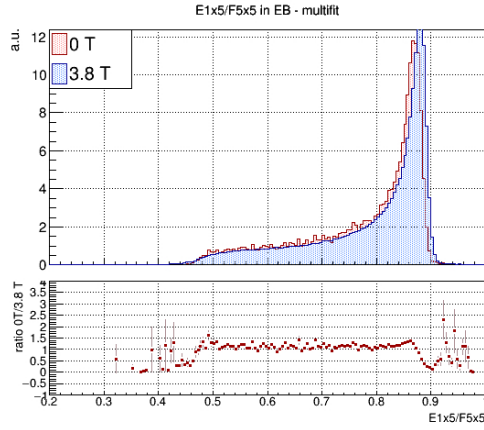
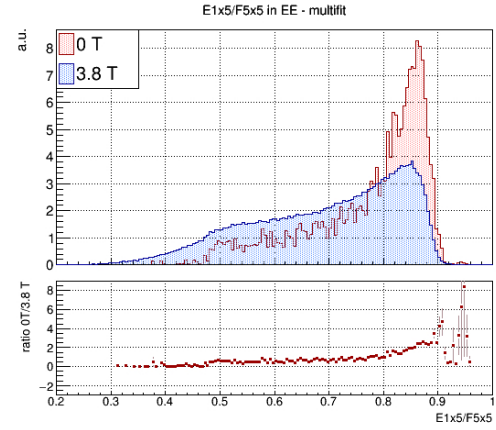
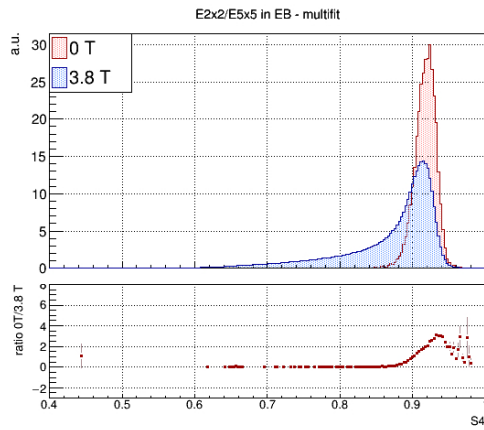
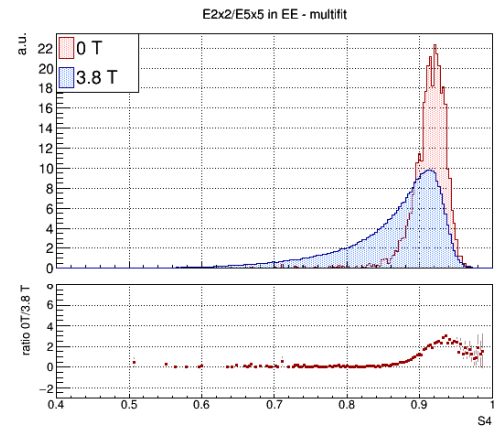
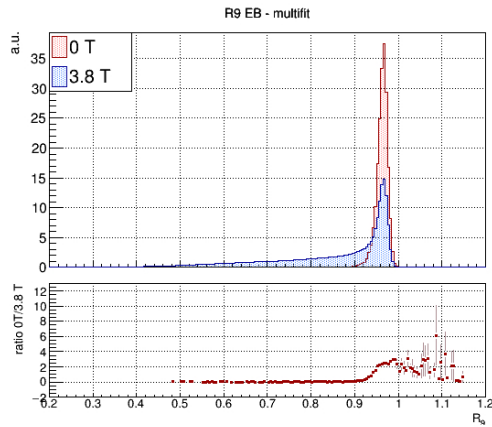
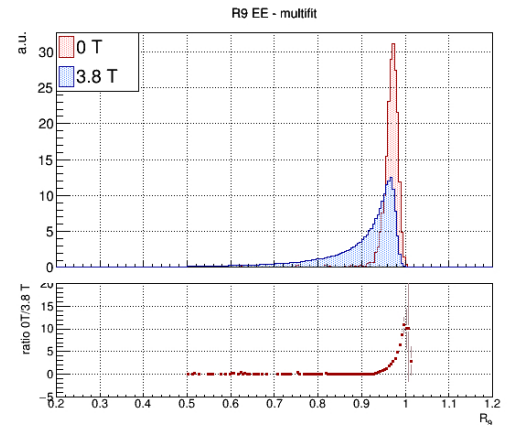
(a) Ratio $E1x5/E5x5$ in barrel.(b) Ratio $E1x5/E5x5$ in endcaps.(c) Ratio $E2x2/E5x5$ in barrel.(d) Ratio $E2x2/E5x5$ in endcaps.(e) Ratio $R_9=E3x3/E5x5$ in barrel.(f) Ratio $R_9=E3x3/E5x5$ in endcaps.

Figure 4.10: Comparison of energy ratios in different crystal matrices at 0 T (red) and 3.8 T (blue).

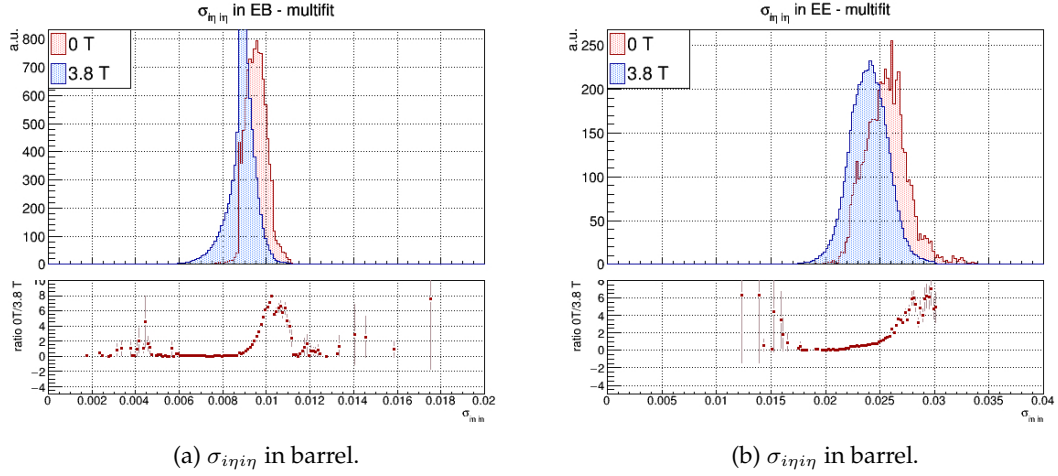


Figure 4.11: Comparison of the spread of the cluster in η at 0 T (red) and 3.8 T (blue).

In Run1 and Run2 two different methods are used to reconstruct the signal in the electromagnetic calorimeter. As mentioned in section 2.3.2 and stressed in figure 2.18, the measurement of the pulse shape consists in 10 samples. Among them, 4 are pure pedestal in EB, and 3 in EE. The peak of the signal lays on the 6th sample, called $BX0$, which corresponds to the bunch crossing time of the event that triggers the readout. Every crystal signal has a similar shape but a different amplitude, depending on the energy deposit.

Overall the pulse shape can be described by an alpha-beta functional form:

$$A(t) = Ped + A \left(1 + \frac{\Delta t}{\alpha \beta} \right)^\alpha e^{-\frac{\Delta t}{\beta}},$$

where the parameters α and β are obtained from test beam data or through a fitting of an average pulse in simulations.

On Run1 the standard “weights” method is applied [73, 76]. The amplitude is computed with a simple linear weighting of the single samples, as the weighted sum of the discrete samples $A = \sum w_i S_i$, where the weights w_i have been chosen in order to optimize the energy resolution. The 4th and the 10th samples are discarded, since the former is the rising edge of the signal and the latter is subject to non negligible crystal-to-crystal variations. The first three samples have negative weights to subtract the pure pedestal, while the other five are used to reconstruct the shape of the signal. In fact, this method is often referred to as “3+5”.

Since out-of-time signal (see section 3.5) with same shape can be observed in bunch crossings before or after $BX0$ (e.g in figure 4.12 [94]), for Run2 a new method has been developed to account for the overlap of signals from different bunch crossings. In fact the presence of OOT pileup breaks the assumption of fixed shape of the “weights” method and degrades the resolution. Several components of pileup can be detected at high η where the pileup occupancy is larger. Two methods were suggested that use the time information: matrix pulse subtraction and multfit [94].

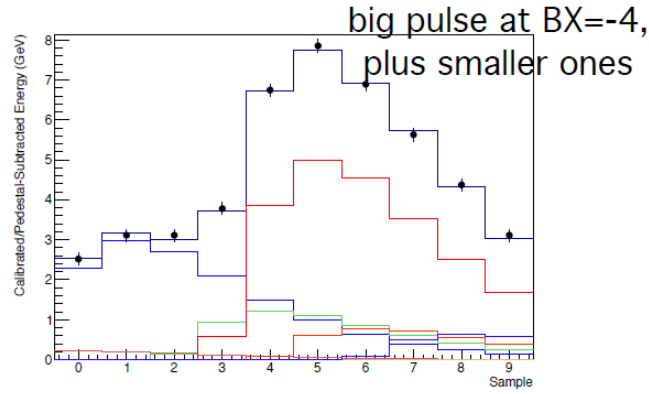


Figure 4.12: The total pulse (in dark blue) shows contributions from the in-time pulse (in red), an out-of-time pulse at $BX=-4$ (in blue) and other smaller pulses [94].

The first method is based on the tagging and subtraction of the pileup in 5×5 crystals and a subsequent weighting of the samples. It works with fixed pedestals and can be performed in $BX \in [-\infty, 4(3)]$ for EB(EE), but the improvement with respect to the “3+5” method is relevant only in the barrel. As shown in figure 4.13 [94], the matrix pulse subtraction method (in green) does not improve the energy resolution in EE (right plot) in comparison to the “3+5” method (in black).

The second method has been chosen for Run2. It is based on a multiple fit [95] of the single

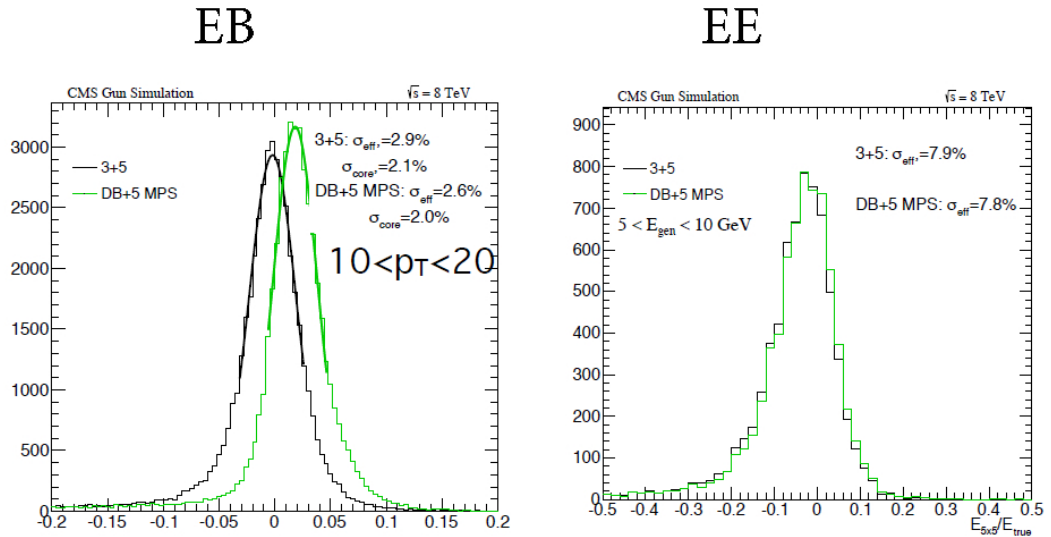


Figure 4.13: Comparison between the matrix subtraction method (in green) and the standard “3+5 method” (in black). Improvements are observed in EB (left) but not in EE (right) [94].

reconstructed signal with 10 components (one for each bunch crossing), and pulse shapes measured in data with low pileup events.

Since basically the light deposited in each ECAL crystal is collected by a photodetector, digitally converted and recorded in 10 samples, 25 ns spaced, and each sample has an amplitude determined not only by the in-time signal, but also by the overlap with out-of-time signals in the previous or following BXs, then a 10-components-multifit can be performed to check in each sample if there are contributions from both interesting signal and one or more OOT pileup interactions. The goal is to understand from the measured signal which fraction of the amplitude is due to pileup and from which BX it comes from, given a fixed pulse shape for each measured signal.

The pileup pulses are tagged in 5x5 crystals with a series of iterative template fits to the crystal amplitude. Those with amplitude larger than 5σ of the noise determine which BXs are “active”. The framework for the multifit of the reconstructed signal is a χ^2 minimization and the inputs are the templates for the pulses, the covariance matrix of the noise and the pedestal (which was subtracted in the templates and samples) [94, 95]:

$$\chi^2 = \sum_i^{N_{samples}} \frac{(\sum_j^{N_{pulses}} A_j p_{ji} - S_i)^2}{\sigma_{S_i}^2} \quad (4.2)$$

Here A_j is the pulse amplitude of the j -th pulse to fit, S_i the measured value of the i -th sample, p_{ij} the pulse template value of the j -th pulse in the i -th sample and σ_{S_i} the noise. Equation 4.2 can be put in matrix form:

$$\chi^2 = \left(\sum_j^{N_{pulses}} A_j \mathbf{p}_j - \mathbf{S} \right)^T \mathbf{\Sigma}_S^{-1} \left(\sum_j^{N_{pulses}} A_j \mathbf{p}_j - \mathbf{S} \right). \quad (4.3)$$

The covariance matrix $\mathbf{\Sigma}_S = \sigma^2 \rho$ is estimated from the correlation matrix ρ measured in the detector and the total noise σ^2 (RMS of the pedestal in the data base).

In case of uncorrelated pulses, the uncertainty on the template can be added in quadrature:

$$\chi^2 = \sum_i^{N_{samples}} \frac{(\sum_j^{N_{pulses}} A_j p_{ji} - S_i)^2}{\sigma_{S_i}^2 + \sum_j^{N_{pulses}} A_j^2 \sigma_{p_{ji}}^2} \quad (4.4)$$

or in the more generic matrix form:

$$\chi^2 = \left(\sum_j^{N_{pulses}} A_j \mathbf{p}_j - \mathbf{S} \right)^T \left(\mathbf{\Sigma}_S + \sum_j^{N_{pulses}} A_j^2 \mathbf{\Sigma}_{p_j} \right)^{-1} \left(\sum_j^{N_{pulses}} A_j \mathbf{p}_j - \mathbf{S} \right), \quad (4.5)$$

where $\mathbf{\Sigma}_{p_j}$ are the covariance matrices associated with each of the pulse templates.

The events reconstructed with the multifit (in black) of 10 simultaneous pulses in $BX \in [-5, +4]$ are compared in figure 4.14 [95, 96] with those obtained through:

- fit for in-time pulse and additional pulses tagged (using 10-pulse fit) at 3σ above the fit uncertainty (in pink);

- fit for in-time pulse and additional pulses tagged (using 10-pulse fit) at 5σ above the fit uncertainty (in green);
- the default “3+5” method (in red).

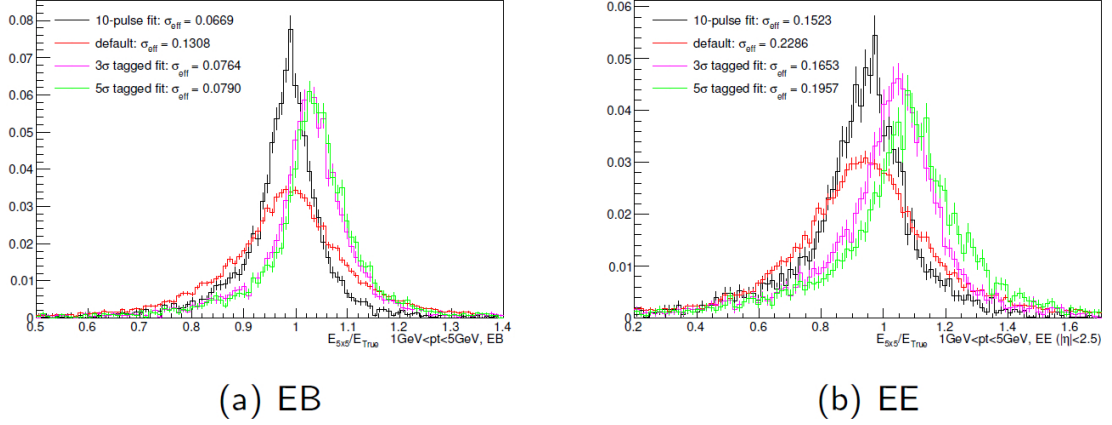


Figure 4.14: Comparison of the signal amplitude reconstructed with different fit methods [95].

In figure 4.15 the distribution of the variable $ootAmpli/Ampli$ is shown at 0 T (top) and 3.8 T (bottom) in barrel (left) and endcaps (right). $ootAmpli/Ampli$ represents the fraction of amplitude of the signal (i.e. energy) in the sample that comes from out-of-time (OOT) pileup. In the in-time sample (the 6th) $ootAmpli/Ampli$ is 0 by construction. As expected, more OOT pileup is observed in the endcaps and without magnetic field (the electrons are not deviated and gather in the same region). Since the probability to have OOT pileup is the same in each sample, these distributions are expected to be almost the same in any sample. A small deviance from what expected is observed. To study this disagreement in more details the means of $ootAmpli/Ampli$ are plotted in different bins of the distribution of the number of vertices, n_{vtx} , in figure 4.16. No structure appears from this plot due to the number of vertices, in fact the only visible fluctuations are in the ranges of n_{vtx} with low statistics. The means of $ootAmpli/Ampli$ are also plot in different η ranges in figure 4.17, but also in this case it is not possible to individuate a dependence on the pseudorapidity of the discrepancy in the different samples. In the end, it was not found a clear explanation of the dishomogeneity of the deposits from out-of-time pileup in the different samples: the origin of the bias in $ootAmpli/Ampli$ was not easily identified but the n_{vtx} and the η variables were excluded as sources. If the bias is constant it is absorbed in the intercalibration constants. Another possibility, that may be checked in the future, is that it depends on the energy or the BX and in this case it is more difficult to handle.

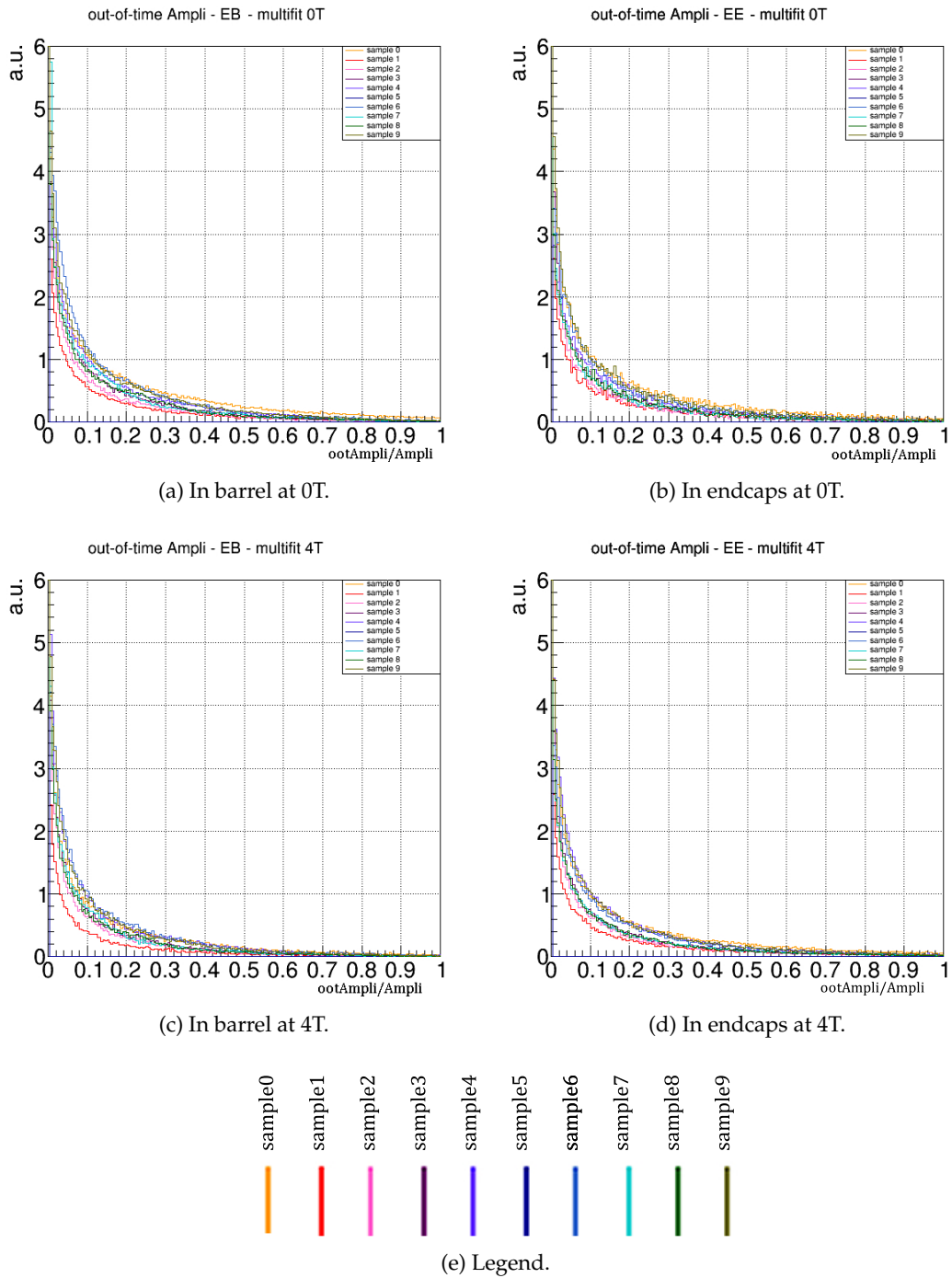


Figure 4.15: Ratio of amplitude of the measured signal which comes from out-of-time pileup. The legend is shown at the bottom.

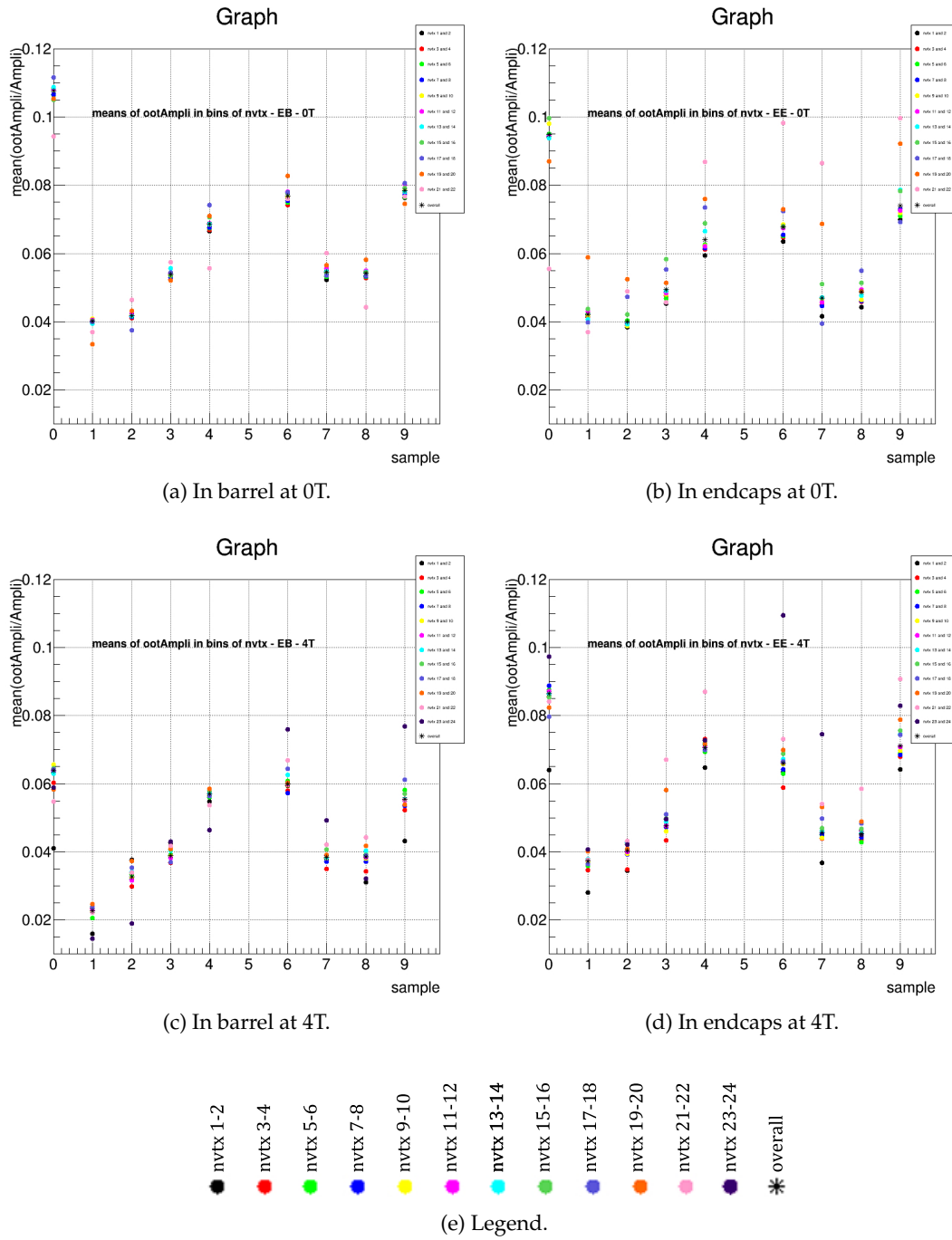


Figure 4.16: Mean of the ratio of amplitude of the measured signal which comes from out-of-time pileup in different ranges of nvtx . The legend is shown at the bottom.

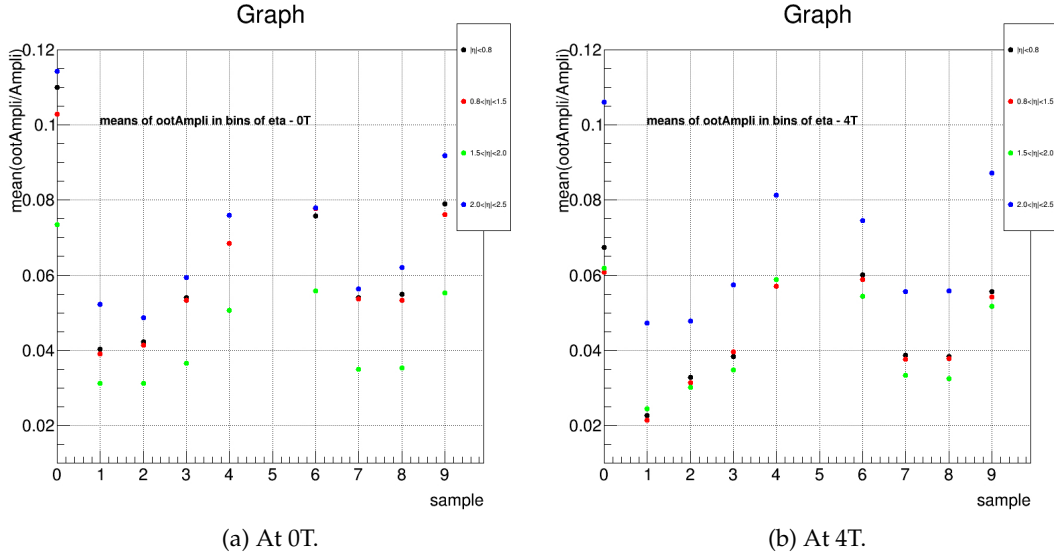


Figure 4.17: Mean of the ratio of amplitude of the measured signal which comes from out-of-time pileup in different ranges of η : $|\eta| < 0.8$ in black, $0.8 < |\eta| < 1.5$ in red, $1.5 < |\eta| < 2.0$ in green and $2.0 < |\eta| < 2.5$ in blue.

In figure 4.18 the distributions of S_4 and $\sigma_{i\eta i\eta}$ are compared for the standard “weights” method (purple) and the “multifit” method (green). In these distributions the main discrepancies between the two reconstruction methods are shown, in fact the clusters reconstructed with the “multifit” method are expected to be slightly narrower. Indeed, it is observed that the η dimension of the cluster is a bit smaller in the case of the “multifit” and in the same conditions the ratio of energy collected in a 2x2 crystals matrix is larger than with the “weights” method. The average pileup (in-time and out-of-time) is about 11 (with RMS of 3.4), deduced from the distribution of the number of reconstructed vertices.

The clusters, in presence of the 3.8 T magnetic field, are reconstructed with both the “multifit” and “3+5” methods and compared to the corresponding Monte Carlo. In figure 4.19 the comparison is shown for the “3+5” method on shower shape variables that depend on the construction of the SC. In figure 4.20 it is shown for clustering independent variables. A mismodeling of just a few percents (see section 4.1) in the simulation of the material upstream of ECAL, that globally does not screw the analysis, can still have visible effects on variables which are very sensitive to material effects like R_9 and $\sigma_{i\eta i\eta}$, where the simulated material appears respectively overestimated (figure 4.19a) for R_9 in EB (probably because approaching the peak close to 1 the variable is less sensitive) and underestimated (figure 4.19d) for $\sigma_{i\eta i\eta}$ in EE. The largest mismodeling, about 20%, observed above $|\eta| = 2.5$, does not have disrupting effects on the shower shapes in endcaps (figures 4.19b and 4.20b). Important discrepancies are also observed at a local level, in E1x5, that cannot be explained only with a material mismod-

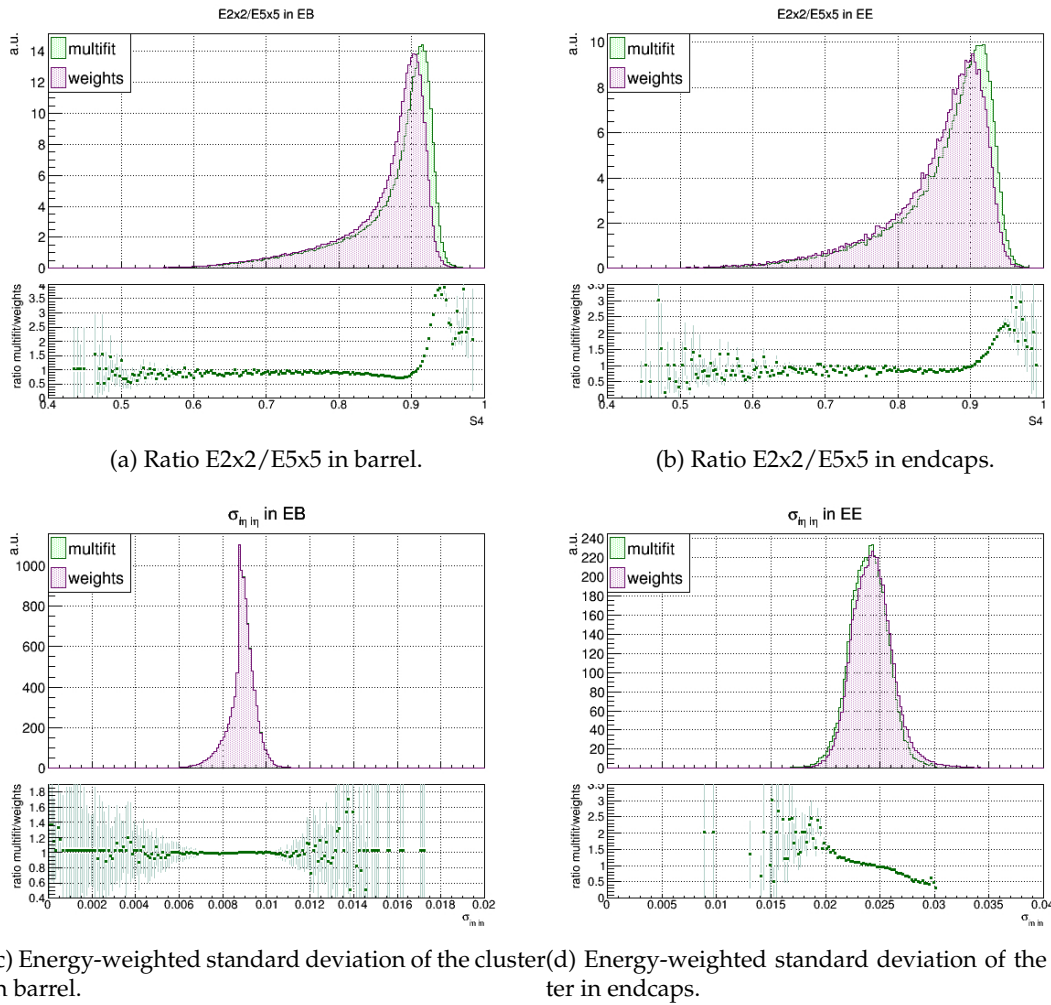


Figure 4.18: Comparison of shower shape variable with multfit (green) and weights (purple) methods.

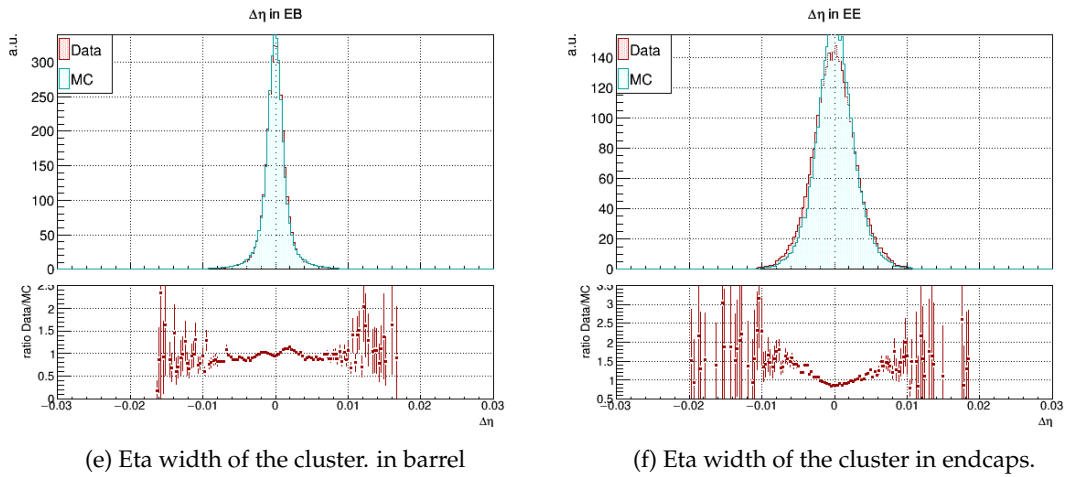
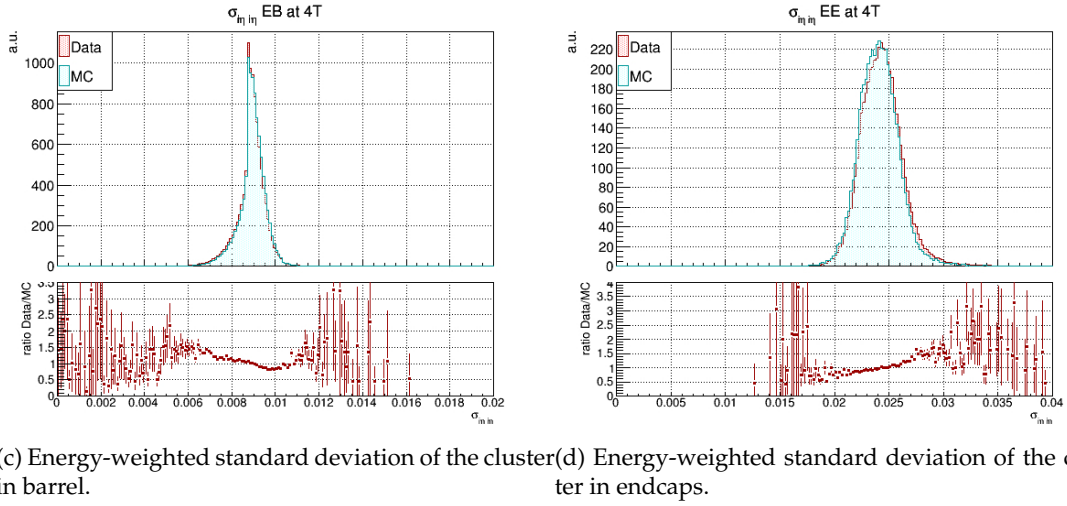
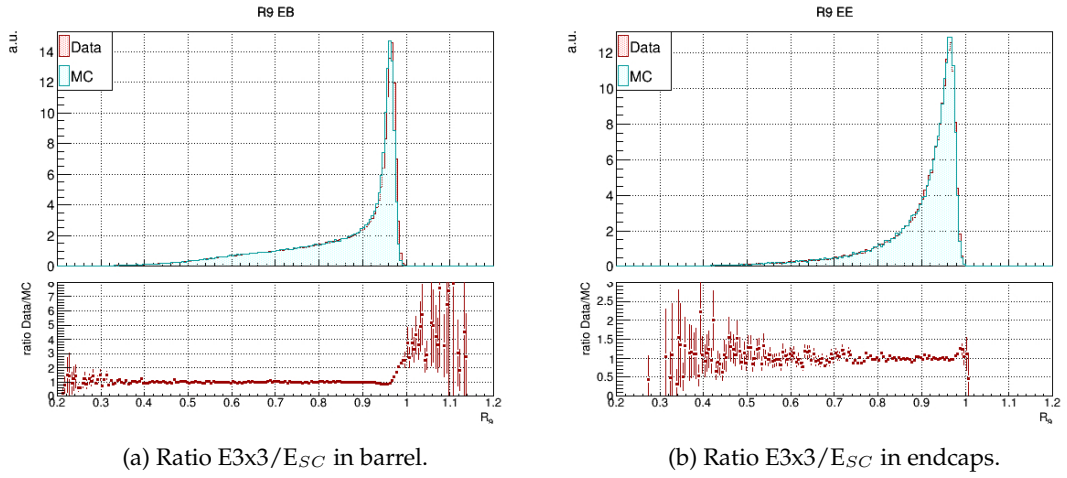


Figure 4.19: Comparison of cluster-dependent shower shape variable in data (red) and simulation (light blue), reconstructed with the “weights” method.

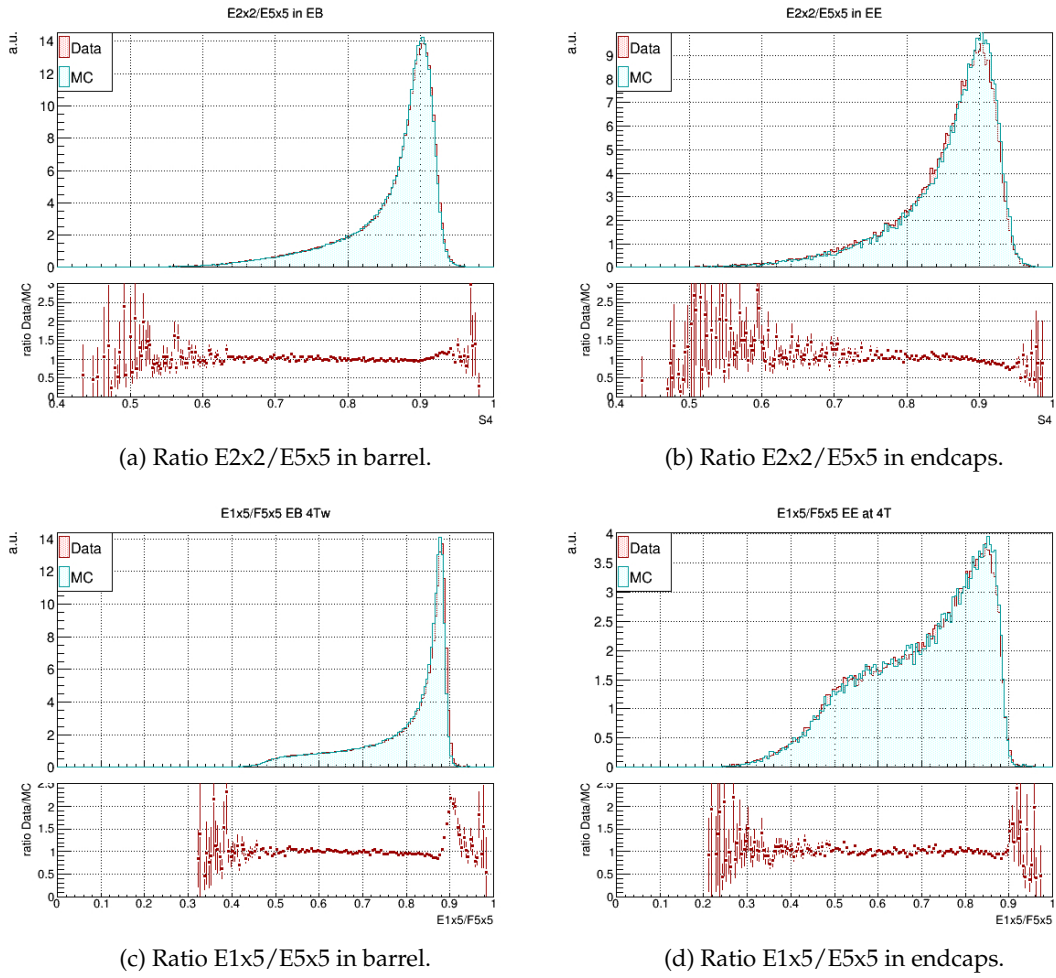


Figure 4.20: Comparison of cluster-independent shower shape variable in data (red) and simulation (light blue), reconstructed with the “weights” method.

eling, hence there must be other effects. Some hypothesis for the sources of discrepancy are the choice of the signal reconstruction method and clustering effects. The latter is excluded as a main cause of discrepancy because the disagreement is observed both in the clustering dependent and independent variables.

To investigate the first hypothesis the same shower shape variables are plotted in figures 4.21 for clustering dependent and 4.22 for clustering independent variables using signals reconstructed with the new “multifit” method. The discrepancies are still observed in equal measure. Only in one variable, $\sigma_{i\eta i\eta}$ in EE (figure 4.21d), an improvement is visible in the case of “multifit”. Hence, it is deduced that the disagreement between data and MC in the shower shape distributions does not depend heavily on the choice of the signal reconstruction method.

In conclusion, discrepancies are observed between data and simulations. They are present both in clustering dependent and clustering independent shower shape variables, hence the disagreement is not due to clustering effects. In addition, they are observed with the same behavior in variables where both the “weights” and “multifit” methods are applied (except for $\sigma_{i\eta i\eta}$ in EE), hence the disagreement is not avoided with the new signal reconstruction method. Moreover, due to the study in the previous section, it is confidently excluded that the discrepancies come entirely from a mismodeling of the material in front of ECAL, which accounts only for the 5% below $|\eta| = 2$ and the 20% above. A confirmation that the material is only a partial cause of disagreement would come from the study of the MC in absence of magnetic field, condition where it was shown that the material effects are less significant. Unfortunately these simulations were not available yet at the time of the writing of this work.

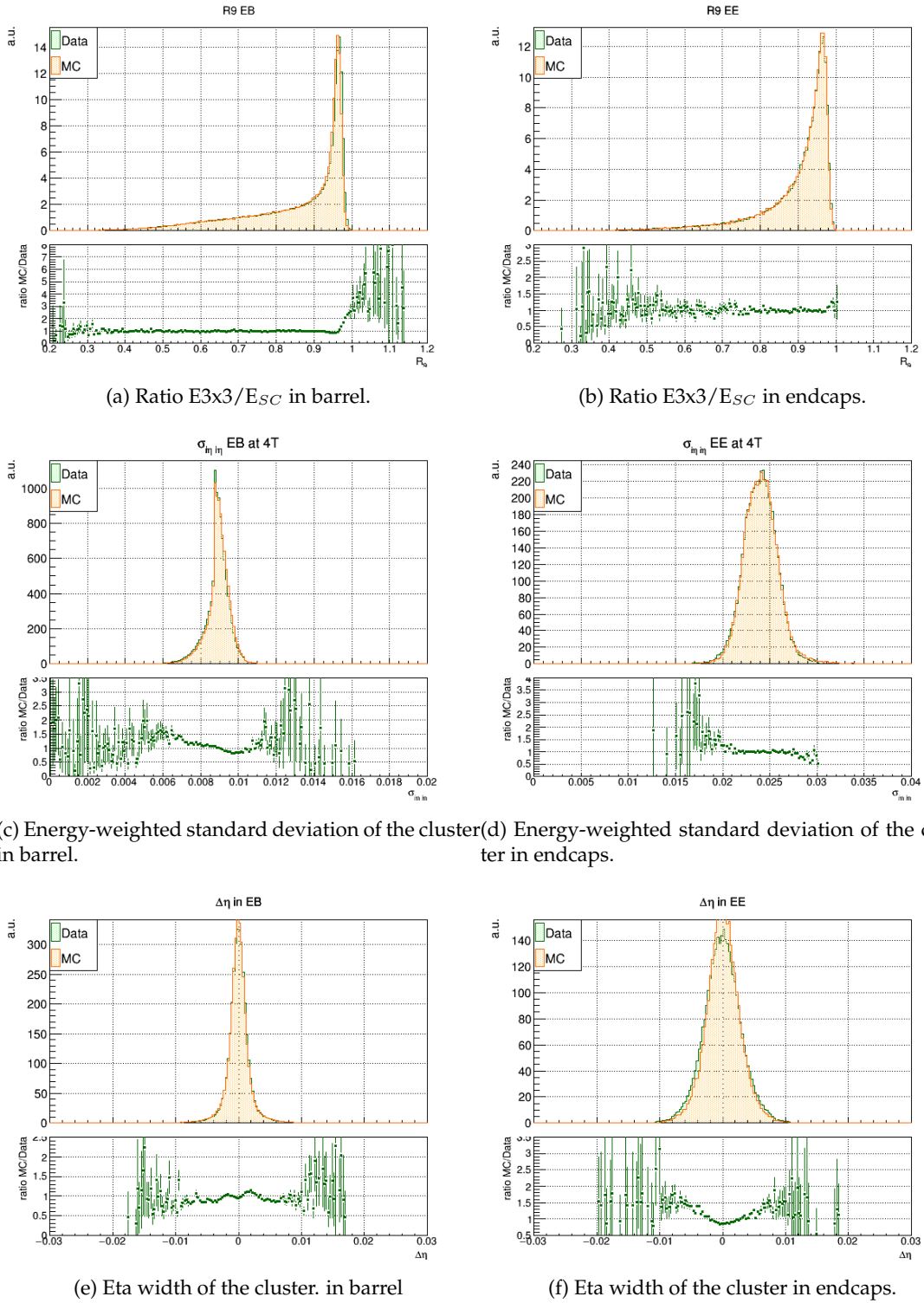


Figure 4.21: Comparison of cluster-dependent shower shape variable in data (green) and simulation (orange), reconstructed with the “multifit” method.

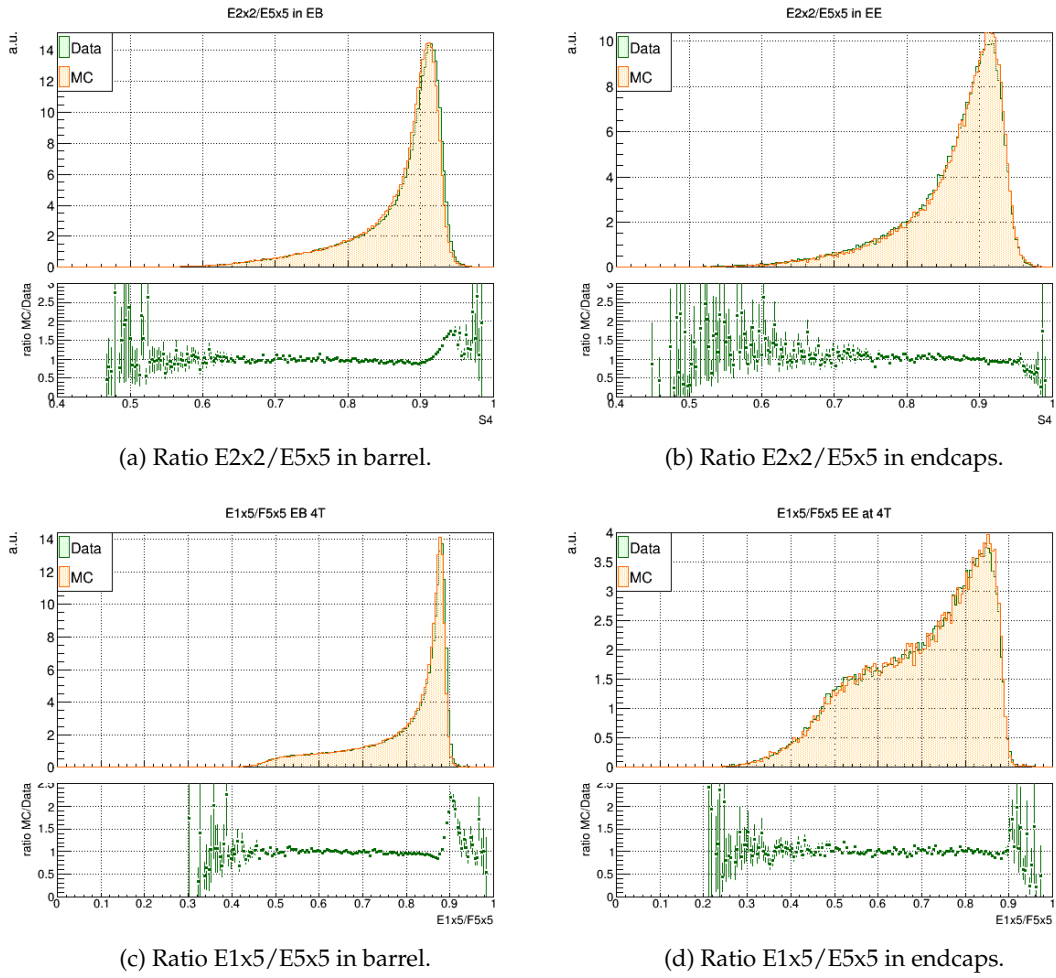


Figure 4.22: Comparison of cluster-independent shower shape variable in data (green) and simulation (orange), reconstructed with the “multifit” method.

Chapter 5

Analysis of the Higgs boson

5.1 Monte Carlo studies

For the study of the Higgs boson production it is crucial to have appropriate generators of events to compare with the real data. These generators must be able to perform computations of cross sections for the different production and decay modes and generate events which are then passed through simulations in order to take into account the detector response.

Studies of the events generators used in CMS at 13 TeV can be made by comparing the transverse momentum (p_T) distributions of the Higgs boson from gluon fusion obtained with different generators, in fact this variable is very sensitive to the choice of the configuration of the generators. In particular, the accurate theoretical resummed computations at Next-to-Next-to-Leading Order and Next-to-Next-to-Leading Logarithm, NNLO+NNLL, performed with HRes [97, 98] are compared with the distributions from the fixed order POWHEG [99] and aMC@NLO [100] generators routinely used at CMS, that merge NLO computations and parton shower Monte Carlo.

The generator HRes is able to simulate the full kinematics of the Higgs boson and of its decay products. It combines calculations of the cross section of production of the Higgs boson up to the perturbative order NNLO and the resummation of the logarithmically enhanced terms NNLL. The cross section computations for the decay channels $H \rightarrow \gamma\gamma$, $H \rightarrow WW \rightarrow \ell\nu\ell\nu$, $H \rightarrow ZZ \rightarrow 2\mu 2e$ and $H \rightarrow ZZ \rightarrow 4e$ are included in the program.

HRes results depend on the choice of the following flags which set the scales of the computation:

- $a0$ for the large top-mass limit [101, 102], $a1$ for the finite top mass [103] only, and $a2$ when the masses of both the top and bottom quarks (involved in the loops at the production of the Higgs) are accounted exactly;
- $p1$ for calculations at the NNLO+NLO perturbative order, and $p2$ for calculations at the NNLL+NNLO perturbative order. The resummation of the multiple soft gluon emissions at small p_T is introduced because around the kinematical boundary at LO in $p_T = m_H/2$

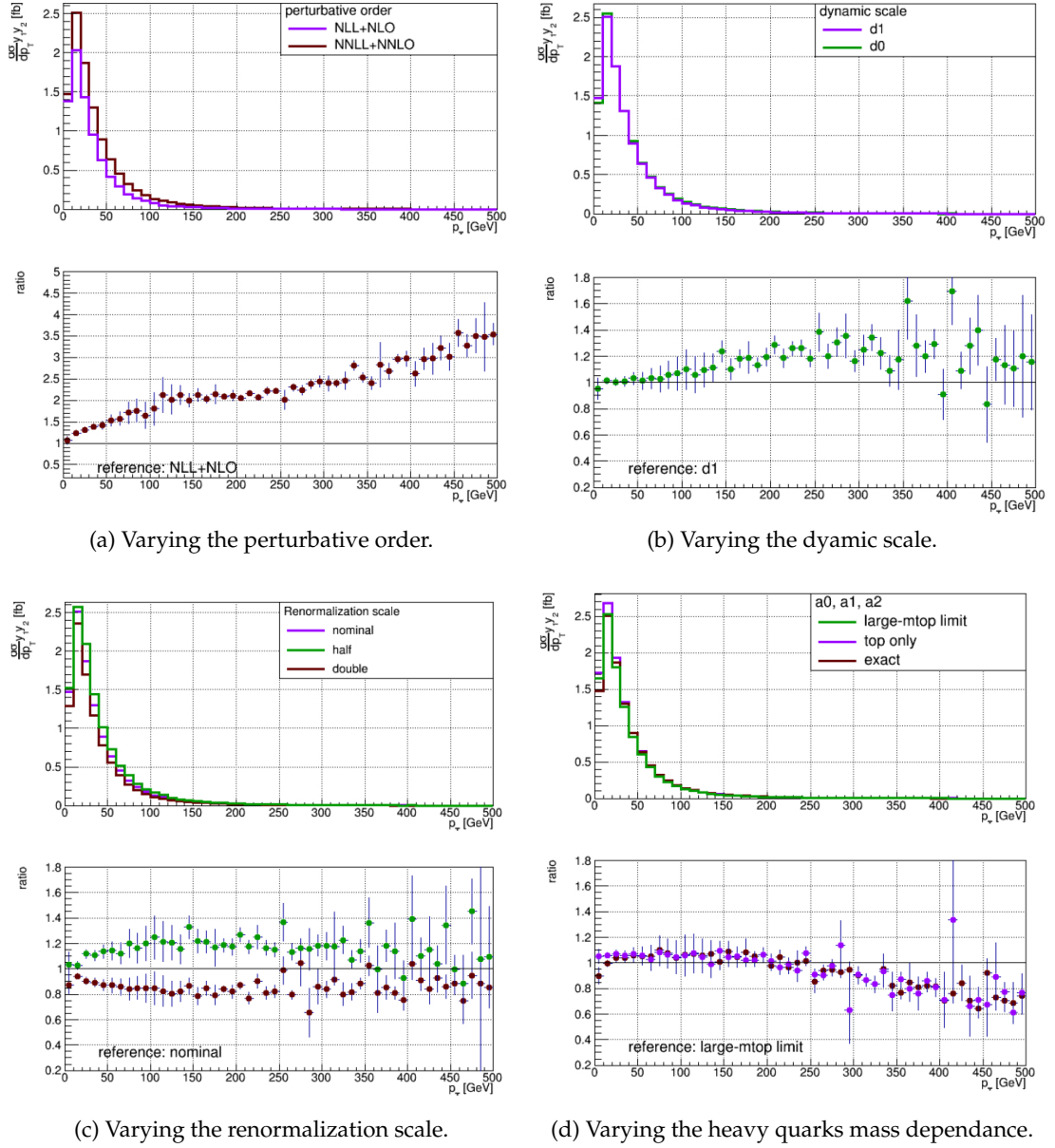


Figure 5.1: Transverse momentum distributions simulated with the HRes generator. The comparison of the spectra while changing one of the flags and keeping the others unchanged is shown.

due to the narrow width approximation (acceptable for an approximate Higgs width of about 4 MeV as predicted by the SM) some logarithmic singularities of the form $\alpha_S^n \ln^{2n} m_H^2/q_T^2$ are observed at higher perturbative orders NLO and NNLO, and they can be suppressed using the resummed NLL+NLO and NNLL+NNLO [101, 104];

- *d0* to set a fixed scale for the central value of the renormalization and the factorization at m_H , and *d1* to set the dynamic scale at m_H for the resummed calculation and at $\sqrt{m_H^2 + p_T^2}$ (transverse energy) for the fixed order calculation;
- *sR1* (nominal), *sR2* (doubled) and *sR05* (halved) renormalization scale. This cut-off is introduced in the definition of the strong coupling to absorb the effects of very short-time physics that lead to UV divergences (virtual loops);
- *sF1* (nominal), *sF2* (doubled) and *sF05* (halved) factorization scale. This arbitrary scale is added to separate short-distance finite and long-distance infinite physics. Indeed, it is used to isolate and absorb the IR divergences (gluon emission).

The nominal configuration is obtained with the parameters choice: *a0p1d1sR1sF1*.

Figure 5.1 shows how the p_T distribution of the Higgs boson produced in gluon gluon fusion changes, while varying one of HRes parameters with respect to the standard configuration and keeping the others at the nominal point.

At NNLL+NNLO perturbative order (figure 5.1a) an enhancement of the distribution is observed. This is not surprising since going from LO to NLO and then NNLO the cross section increases (at LHC expected about +25% from NLO to NNLO and a contribution of slightly less than +10% from the resummation of NNLO and NNLL at 13 TeV in ggH [105]).

At the nominal working point the large-top mass limit is assumed (effective Lagrangian approach with one loop less). This limit is used to compute the high order correction, but it works well only at small momenta ($p_T < 2m_{top}$), hence for the cases with top mass only and exact top and bottom masses we observe, as expected, the same decrease at high momenta (figure 5.1d). In fact, at large p_T the top contribution dominates and reduces the cross section with respect to the large top mass limit. Focusing on the small p_T , it is evident that the bottom contribution becomes significant and it changes the shape of the spectrum.

The distribution increases and decreases of about the same amount while varying the factorization or renormalization scale (figure 5.1b) from the nominal value to a doubled and an halved one, without changing the shape of the distribution.

The employment of the dynamic scale shows the p_T dependence due to the introduction of the transverse energy as factorization and normalization scale, that increases with the transverse momentum. The dynamic scale is believed to be more appropriate at high p_T [106].

The same comparison was performed on the distributions of rapidity, pseudorapidity and azimuthal angle. The results show less significant differences, mainly because these variables are less sensitive to the radiative corrections, accounted in different ways in HRes depending

on the choice of the flags and parameters.

The two main generators used in CMS to simulate events are POWHEG and aMC@NLO [107, 108]. POWHEG exponentiates the full real emission matrix element to generate the hard-

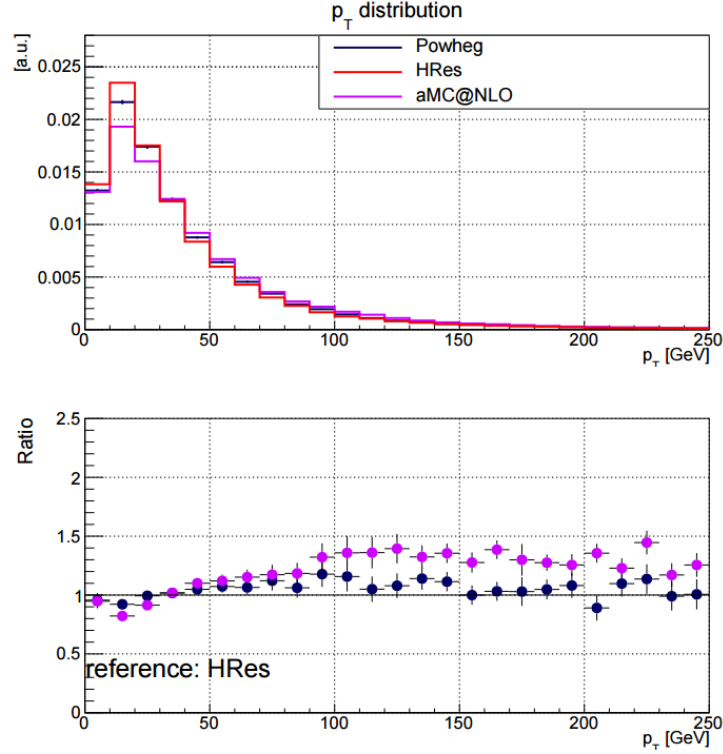


Figure 5.2: Comparison between the p_T distribution obtained by HRes generator and with the two event generators routinely used by CMS: POWHEG and aMC@NLO.

est radiation implementing NLO calculations and it is interfaced to a shower Monte Carlo. It simulates events with 0 jets at NLO, with 1 jet at LO and with 2 jets at the parton shower (PS) level. In POWHEG [99] infrared-safe observables have NLO accuracy and collinear emissions are summed at the leading-logarithmic level. Its peculiarity, from which it takes the name of Positive Weight Hardest Emission Generator, is that it produces positive-weighted events. POWHEG predictions are weighted to match by construction those by HRes, which accounts for NNLO and NNLL QCD corrections [39].

aMC@NLO is an automated matrix element generator capable of generating LO and NLO processes and interfaced with a parton-shower program. It is one of the first programs where NLO Higgs boson production via gluon fusion is merged with a shower Monte Carlo program. It generates events with 0, 1 or 2 jets at NLO, with 3 jets at LO and with 4 jets at PS level. Its goal is to compute tree-level and one-loop amplitudes for arbitrary processes from a theoretical model and a set of process-independent building blocks in a general way that can be applied to several processes.

In figure 5.2 these two generators are compared with the theoretical prediction from HRes. The p_T distributions are normalized to the number of events and a simple selection assures to select two photons that come from the Higgs decay. POWHEG fits very well the theoretical distribution of HRes, while aMC@NLO shows a small discrepancy, especially at high p_T , but it has the advantage of a more accurate description of jets in the final state, particularly important when differential measurements and Higgs boson coupling measurements are performed. The discrepancy between aMC@NLO and POWHEG at large p_T is due to NNLO effects in the POWHEG formula for the differential cross section, that push POWHEG closer to the NNLO results [109].

A standard preselection is then applied to the generated events in the channel $gg \rightarrow H \rightarrow \gamma\gamma$ to check how it influences the shape of the transverse momentum distribution. The variables involved in the selection are R_9 , the η coordinate of the SC, the p_T of the photons, the energy-weighted spread of the super-cluster along η , the electron-photon isolation, H/E , the sum of the p_T of the tracks in a cone of $R = 0.3$ and the score of the photon identification MVA. The preselected events are then split in classes by means of the estimator of the diphoton MVA that finds the best candidate diphoton from the Higgs boson, based on kinematics, vertex identification, quality of the energy measurement, etc., as explained in section 3.7. The result of the categorization is shown in figure 5.3: in orange the class untagged3 in the

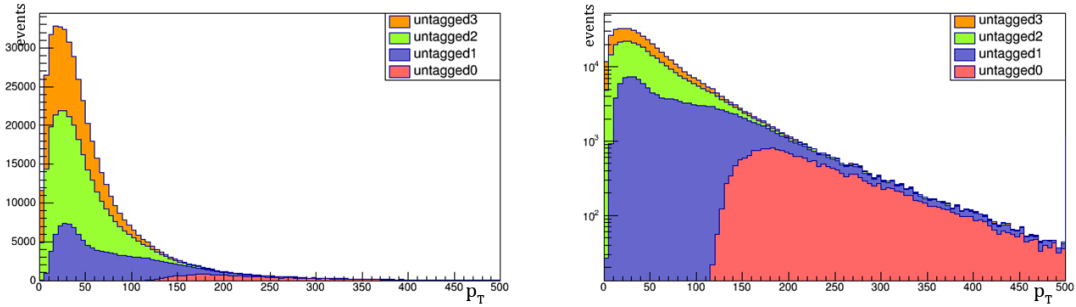


Figure 5.3: p_T spectrum of $gg \rightarrow H \rightarrow \gamma\gamma$ events simulated with POWHEG in different diphoton BDT classes. In magenta the class that selects mainly unconverted photons.

diphoton MVA range $[-0.338, 0.273]$, in green untagged2 in $[0.273, 0.617]$, in indigo untagged1 in $[0.617, 0.870]$ and in magenta untagged0 in $[0.870, 1.0]$. The latter is the cleanest sample of Higgs events. The events with diphoton MVA score less than -0.338 are rejected as background events.

As explained in section 1.3.4 there are different production modes for the Higgs boson and they can all be simulated with the two generators. aMC@NLO, unlike POWHEG, creates also events with negative weights which might not be well treated when used as inputs to train a diphoton MVA. Figure 5.4 shows that the effect of the negative weights on the transverse momentum distribution is important for ggH (figure 5.4a) and VH (figure 5.4b), while it is negligible for ttH (figure 5.4c) and VBF (figure 5.4d). Hence, due to the discrepancy in presence of events with negative weights, the distribution of the transverse momentum of the Higgs boson

is not reliable when using aMC@NLO to simulate the ggH and VH production mechanisms. In any case, this disagreement does not affect the complete analysis where the selections are based on the distributions of the photons and the diphoton systems.

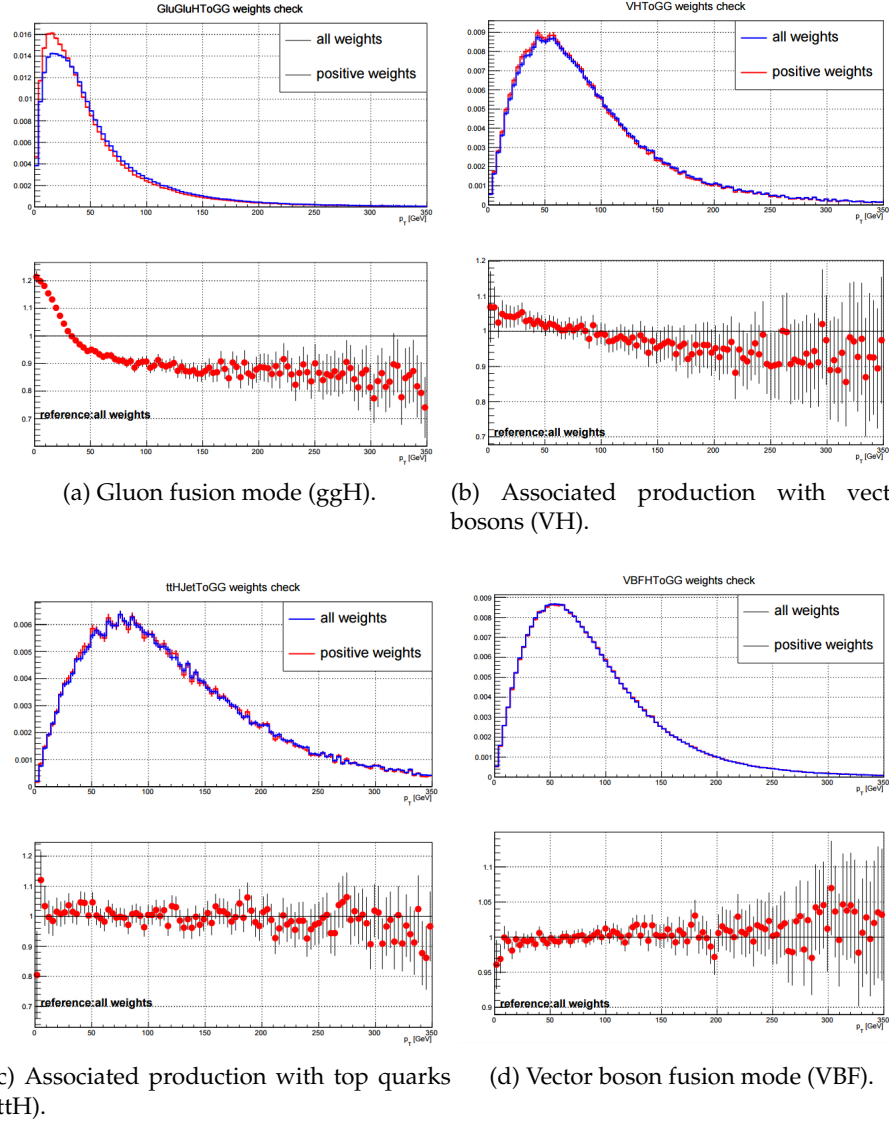


Figure 5.4: Comparison of the p_T distribution of the Higgs events simulated by aMC@NLO in the different production channels with (blue) and without (red) negative weights.

5.2 Higgs boson associated production with a vector boson and decay in two photons

After the studies of the material of the detector (section 4.1), the examination of the agreement between simulations and data (section 4.2) and the checks on the MC generators performances (section 5.1), this section will describe the optimization of the selection criteria used to classify the events which are good candidates as signature of Higgs boson decay and reject the background events. Part of the events classification for the search of the Higgs boson described in section 3.8 is used here on Monte Carlo simulations prepared for Run2 at 13 TeV. The VH selection (with its four subcategories, described in section 3.8) is applied on simulated Higgs events in the ggH, VBF and VH modes with H, decaying into a diphoton system, where ggH and VBF are considered as background for this class. The whole classification includes, in this order, the selections: “ttH ℓ ”, “VH tight ℓ ”, “VH loose ℓ ”, “VBF dijet 0-2”, “VH met”, “ttH multijet”, “VH dijet” and “untagged 0-4”. Thus, in this particular study, the events produced in the ttH channel are not included in the categorization because they are very efficiently selected by the ttH selection, preceding the VH selection, hence, on first approximation, they can be excluded from this analysis. The VH channel is one of the most tricky Higgs boson production channels due to the very small cross section (see figure 1.14).

The VH subselections “tight ℓ ” (either “tight1” or “tight2”), “loose ℓ ”, “met” and “dijet”, described in the following paragraphs, are performed in this order. Moreover, the categories have to be mutually exclusive, hence the events that are accepted in a class cannot fall in another one. The criteria of selection are listed in table 5.1. The events have been reweighted: both for the weight of the generator at an hard scattering level and for the distribution of pileup events.

The selection “tight1” counts events with one lepton and large missing energy (met), signature of $W \rightarrow \ell\nu$ decay. On the other hand, “tight2” criteria select the two leptons of same flavor and opposite sign (with no constraint on the missing energy), likely coming from $Z \rightarrow \ell^+\ell^-$. The cut on $m_{\ell,\gamma}$ is meant to reject the main background due to $Z\gamma$ production with Z decaying in e^+e^- . The minimal distance $\Delta R(\ell, \gamma)$ is required to cut events with initial or final state radiation. The selection on $m_{\ell\ell}$ is used to select electrons from the Z boson decay. The variable θ^* is the angle of the diphoton system in the center of mass frame of the system diphoton-dijet with respect to the direction of motion of the diphoton-dijet in the lab frame. The distribution of $\cos \theta^*$ for VH events is almost flat, while for background and ggH events it is peaked at $|\cos \theta^*| = 1$. Events with more than three jets that pass the selection in “set 3” are discarded. The diphoton system must also be well isolated in ϕ from the direction of the missing energy ($\Delta\phi(\gamma\gamma, met) > 2.1$).

The fractions of events of the different subcategories obtained in Run2 are shown in table 5.2 (left) and compared to those resulting from Run1 at 8 TeV (19.7 fb^{-1}) (right). On Run2 only the VH selections are applied, while on Run1 the whole classification was performed. The results are compatible considering that in Run2 the VH classes on events from VBF channel suffer of the absence of the “VBF dijet 0-2” selections that should be performed before “VH

		tight1	tight2	loose	met	dijet
set 1	E_T^{miss} [GeV]	> 45			> 70	
	$diphoMVA$	> -0.4	> -0.4	> -0.4	> -0.4	> -0.4
	$p_T^{\gamma_1}$	$> 3m_{\gamma\gamma}/8$	$> 3m_{\gamma\gamma}$	$> 3m_{\gamma\gamma}/8$	$> 3m_{\gamma\gamma}/8$	$> m_{\gamma\gamma}/2$
set 2	p_T^ℓ [GeV]	> 20	> 10	> 20		
	$ \eta_e ^{(*)}$	< 2.5	< 2.5	< 2.5	< 2.5	< 2.5
	$ \eta_\mu $	> 2.4	> 2.4	> 2.4		
	ℓ isolation	< 0.15	< 0.15	< 0.15		
	$\Delta R(e, \gamma)$	> 1.0	> 1.0	> 1.0		
	$\Delta R(\mu, \gamma)$	> 1.0	> 0.5	> 1.0		
set 3	$\Delta R(j, \gamma)$	> 0.5	> 0.5	> 0.5		> 0.5
	$\Delta R(j, \ell)$	> 0.5	> 0.5	> 0.5		
	p_T^j [GeV]	> 20	> 20	> 20		> 40
	$ \eta_j $	< 2.4	< 2.4	< 2.4		< 2.4
set 4	$m_{\ell\ell}$ [GeV]		[70, 110]			
	$ m_{\ell\gamma} - m_Z $ [GeV]			> 10		
	$\Delta\phi(\gamma\gamma, met)$				> 2.1	
	$\Delta\phi(\gamma\gamma, jet_1)$				< 2.7	
	m_{jj} [GeV]					[60, 120]
	$ \cos\theta^* $					< 0.5

Table 5.1: VH selection criteria. ^(*) The range $1.44 < |\eta_e| < 1.57$, covering the gap between EB and EE, is always excluded in the analysis.

met”. Furthermore, in Run1 there was a “b-tag rejection” in “VH dijet” which is not applied here. The cut on the missing energy should probably be slightly increased too, in order to improve the “met” selection. In addition the met corrections implemented in Run1, based on discrepancies between data and simulated events in the direction and magnitude of the E_T^{miss} vector, are not applied here.

subcategory	ggH	VBF	VH
VH tight ℓ	0.9%	3.7%	32.2%
VH loose ℓ	2.5%	15.5%	16.1%
VH met	36.5%	26.0%	31.9%
VH dijet	60.1%	54.8%	19.8%

(a) Fractions at 13 TeV

subcategory	ggH	VBF	VH
VH tight ℓ	0.4%	2.7%	30.0%
VH loose ℓ	2.9%	8.9%	19.0%
VH met	36.4%	43.8%	28.1%
VH dijet	60.3%	44.6%	22.9%

(b) Fractions at 8 TeV

Table 5.2: Fractions of events that pass the different VH selections in the same channel, for the modes ggH, VBF and VH, for Run2 (left) and Run1 (right) simulated events.

The fractions of events for each category in the different channels in Run2 are shown in table 5.3 (top) and compared to those of Run1 (bottom) with reasonable agreement. In table 5.4 other

independent studies on the VH selections in Run2 show pretty similar results using similar cuts [110, 111]. In the table at the bottom met corrections and a simplified VBF selection are accounted. In the table at the top the complete classification is performed.

channel	tight ℓ	loose ℓ	met	dijet
ggH	1.3%	4.9%	32.2%	53.4%
VBF	0.7%	3.6%	2.8%	5.8%
VH	98.0%	91.5%	65.0%	40.8%

(a) Fractions at 13 TeV

channel	tight ℓ	loose ℓ	met	dijet
ggH	0.2%	2.7%	18.4%	31.1%
VBF	0.2%	1.1%	3.0%	3.2%
VH	99.6%	96.2%	78.6%	65.7%

(b) Fractions at 8 TeV

Table 5.3: Fractions of events selected by each VH category that come from the different mechanisms of production (excluded ttH).

channel	tight ℓ	loose ℓ	met	dijet
ggH	1.8%	1.9%	36.0%	51.5%
VBF	0.02%	0.01%	4.9%	3.8%
VH	98.2%	98.1%	59.1%	44.7%

(a) Independent study where also ttH and VBF selections are implemented [110].

channel	tight ℓ	loose ℓ	met	dijet
ggH	1.3%	8.7%	34.1%	80.3%
VBF	0.6%	3.3%	2.8%	2.7%
VH	98.1%	88.0%	63.1%	17.0%

(b) Independent study where met corrections are accounted, but ttH and VBF selections are not applied, as in this work [111].

Table 5.4: Fractions of events selected by each VH category that come from the different mechanisms of production (excluded ttH), obtained by independent studies on Run2 simulated events at 13 TeV [110, 111].

The significance of the “tight” VH selection can be computed, in Gaussian approximation, as the square root of its total number of events N_{tight} . With a 24.5 fb^{-1} integrated luminosity accumulated at 13 TeV, the significance of “VH ℓ tight” selection results to be about 2σ . A significance of $3(5)\sigma$ would be reached with approximately $68(190) \text{ fb}^{-1}$ integrated luminosity. The same computation can be performed on the other categories to find, for the total number

$N = N_{tight} + N_{loose} + N_{met} + N_{dijet}$ of events selected by the complete VH selection, a significance of 4.4σ . The significance of the complete VH selection would reach 5σ with a luminosity of about 31 fb^{-1} . It has to be underlined that these estimates are based on the expected number of events for the Higgs boson signal in the different analysis channels, only. It does not consider the different background processes that take place in proton-proton collisions, important especially for the hadronic selections. Indeed, the significance in the real case that accounts for the background can be expected to be worse, hence larger statistics is needed to reach the sought 5σ .

For the Run2 data taking in 2016 a new version of CMS software (CMSSW framework [112]) has been developed (80x). Several improvements in the reconstruction algorithm have been made with respect to the previous version (76x). A validation of the new version has been performed. The comparison between version 80x and version 76x is shown in the following figures, where the distributions of the variables used in the VH selections are plotted.

The distributions of the variables involved in “tight” are shown: in figure 5.5 including no selection, in figure 5.6 including the selections on criteria in “set 1 - tight”, in figure 5.7 including also selections on lepton properties in “set 2 - tight” and in figure 5.8 including also selections on the number of jets according to “set 3 - tight”.

The distributions of the variables involved in “loose” are shown for events which did not pass “tight”: in figure 5.9 with no selection, in figure 5.10 including selections on criteria in “set 1 - loose”, in figure 5.11 including also selections on lepton properties in “set 2 - loose” and in figure 5.12 including also selections on the number of jets according to “set 3 - loose”.

The distributions of the variables involved in “met” are shown for events which did not pass “tight” and “loose”: in figure 5.13 with no selection and in figure 5.14 including selections based on criteria in “set 1 - met”.

The distributions of the variables involved in “dijet” are shown for events that did not pass all the other selections: in figure 5.15 with no selection, in figure 5.16 including selections on the variables in “set 1 - dijet” and in figure 5.17 including also selections on the number of jets according to “set 3 - dijet”.

The distributions in the two versions 80x and 76x turned out to be compatible, as expected.

After showing that the versions 80x and 76x are consistent and the VH selections work properly on simulated events, these studies will continue and be improved with an optimization of the diphoton MVA estimator and implemented in the complete analysis. The whole categorization, including also the selections not accounted for in this specific study, in the future will be applied on data in the diphoton mass window $115 < m_{\gamma\gamma} < 135 \text{ GeV}$ to complete the search for the peak corresponding to the mass of the Higgs boson.

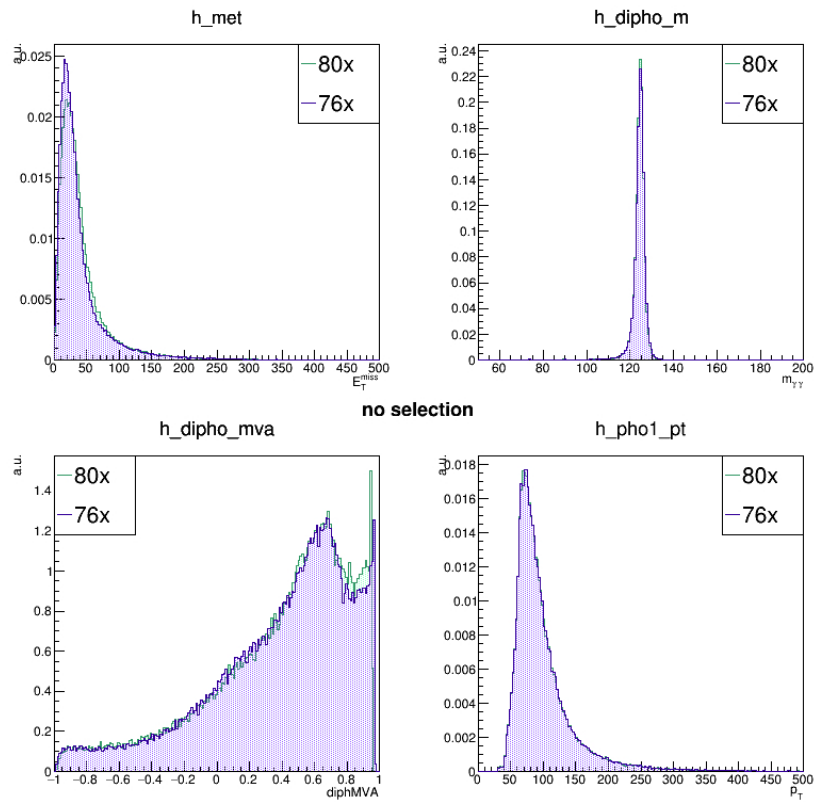


Figure 5.5: Distributions of E_T^{miss} , $m_{\gamma\gamma}$, *diphoMVA* and $p_T^{\gamma_1}$, to be used in “tight” selection. No selection is applied.

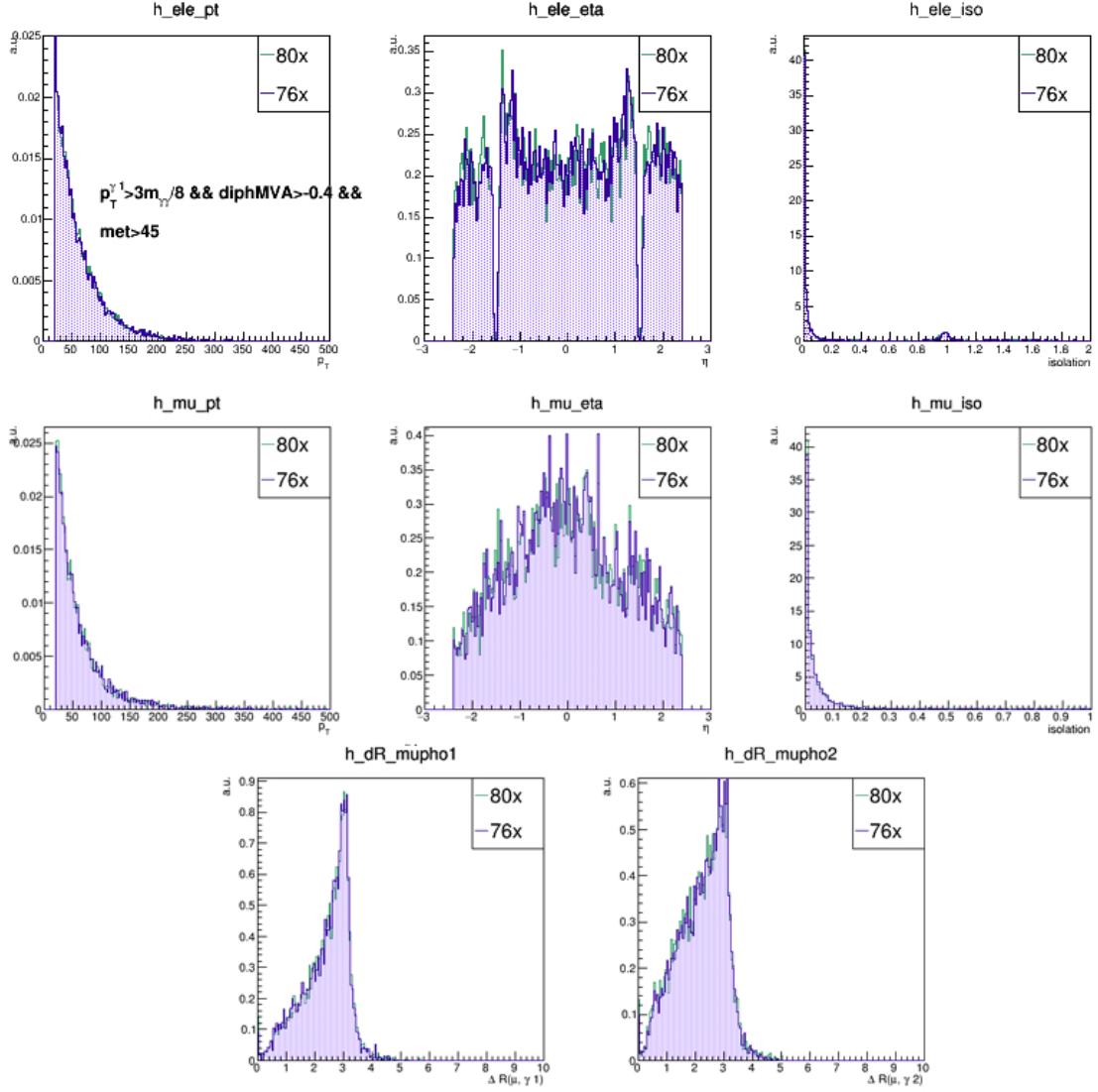


Figure 5.6: Distributions of $p_T^{(e,\mu)}$, $\eta_{(e,\mu)}$, lepton isolation and $\Delta R(\mu, \gamma_{(1,2)})$, to be used in "tight" selection. The selections in "set 1 - tight" are applied.

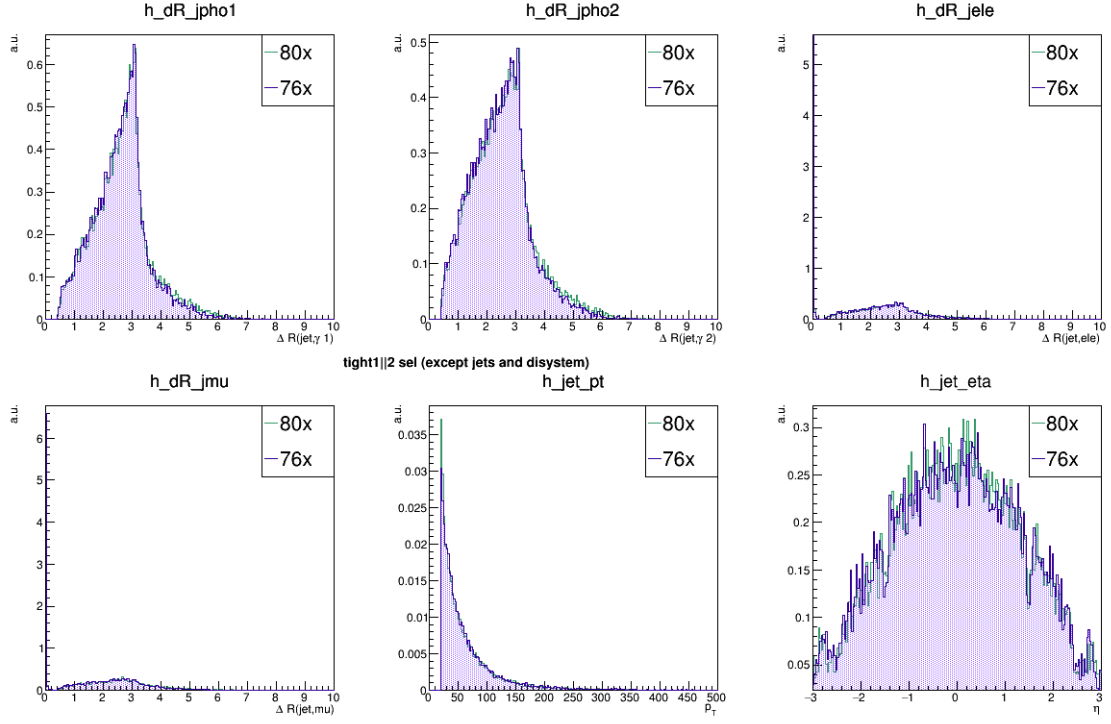


Figure 5.7: Distributions of $\Delta R(jet, \gamma_{(1,2)})$, $\Delta R(jet, \ell)$, p_T^j and η_j , to be used in "tight" selection. The selections in "set 1 - tight" and "set 2 - tight" are applied.

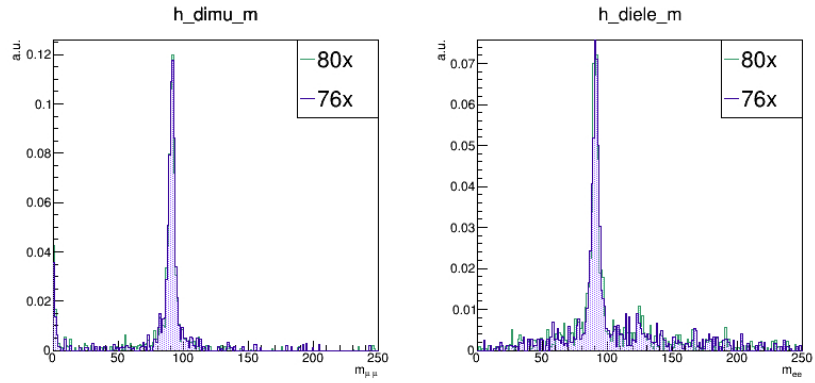


Figure 5.8: Distributions of $m_{\ell\ell}$, to be used in "tight" selection. The selections in "set 1 - tight", "set 2 - tight" and "set 3 - tight" are applied.

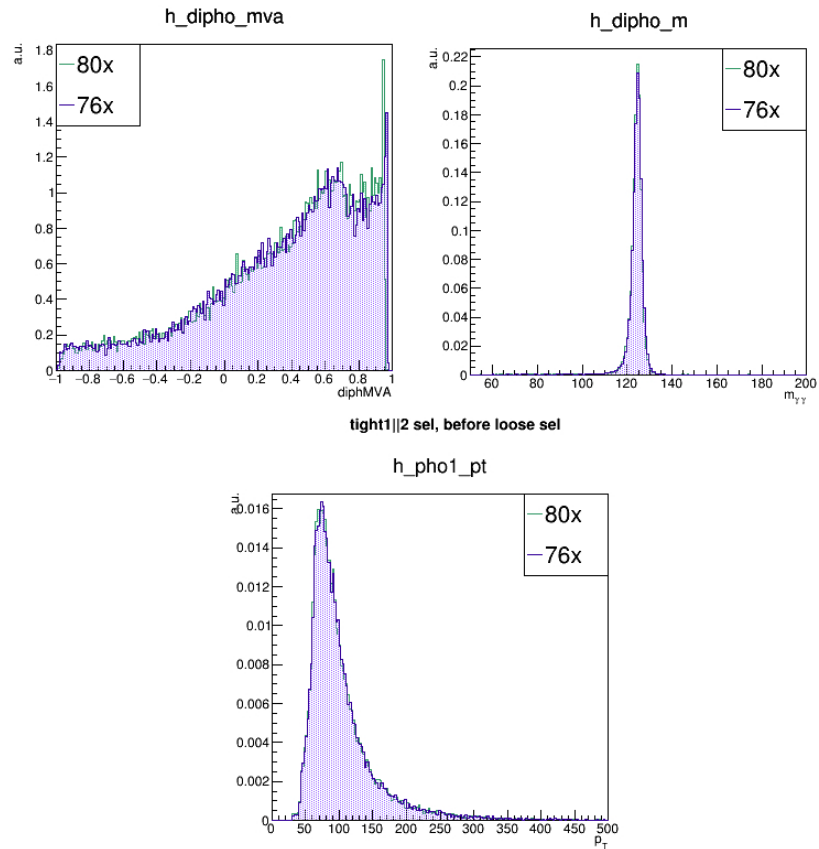


Figure 5.9: Distributions of $m_{\gamma\gamma}$, $diphoMVA$ and $p_T^{\gamma_1}$, to be used in “loose” selection, performed on data that did not fulfill the “tight” criteria. No selection is applied.

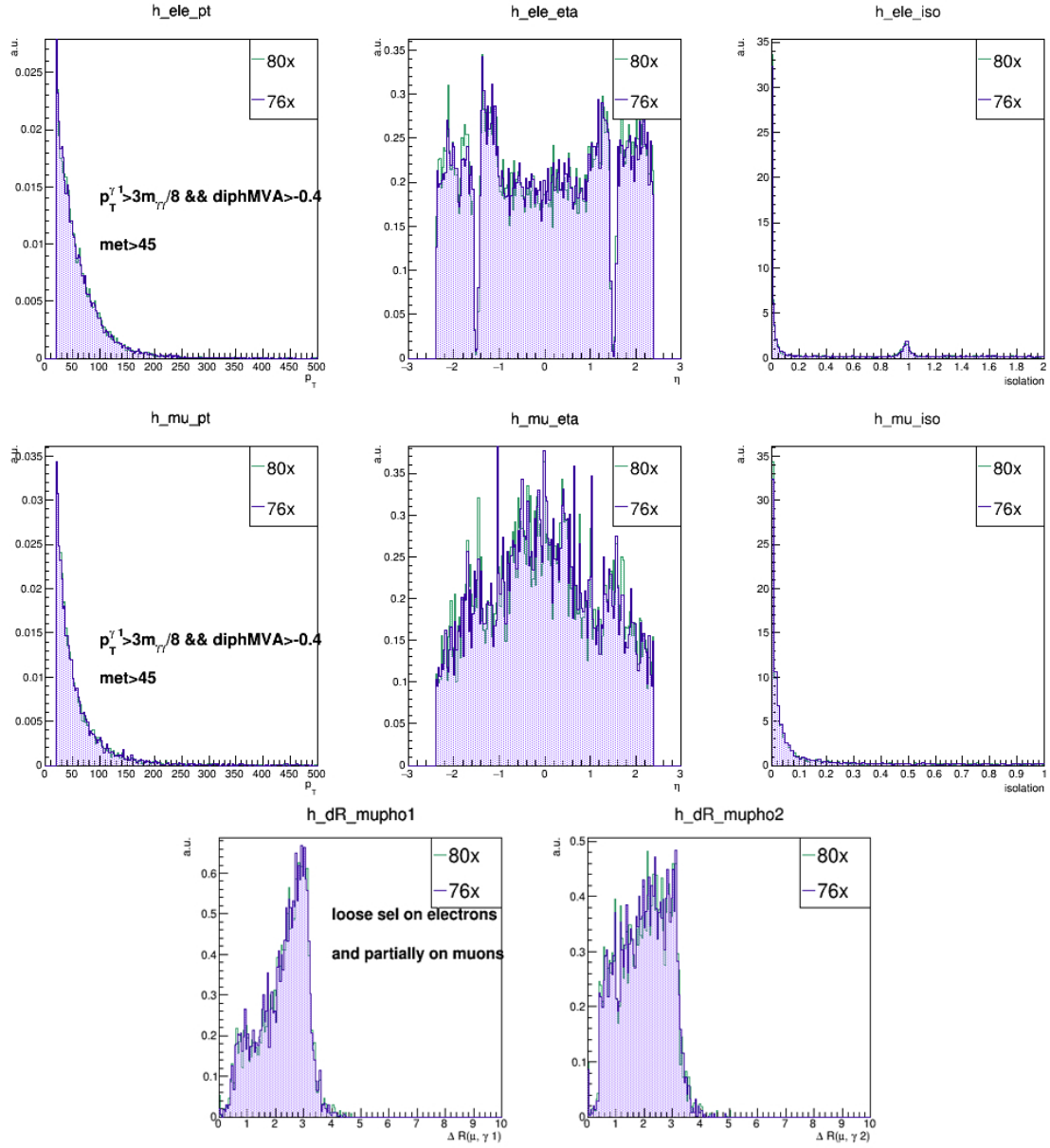


Figure 5.10: Distributions of $p_T^{(e,\mu)}$, $\eta_{(e,\mu)}$, lepton isolation and $\Delta R(\mu, \gamma_{(1,2)})$, to be used in “loose” selection. The selections in “set 1 - loose” are applied.

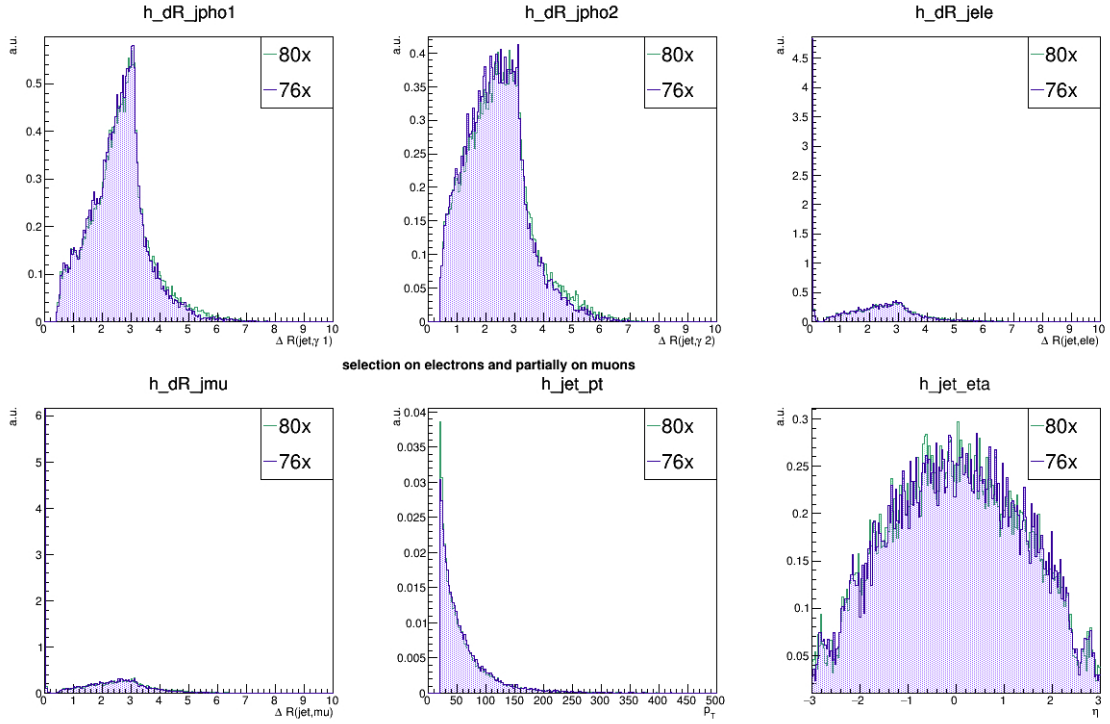


Figure 5.11: Distributions of $\Delta R(jet, \gamma_{(1,2)})$, $\Delta R(jet, \ell)$, p_T^j and η_j , to be used in “loose” selection. The selections in “set 1 - loose” and “set 2 - loose” are applied.

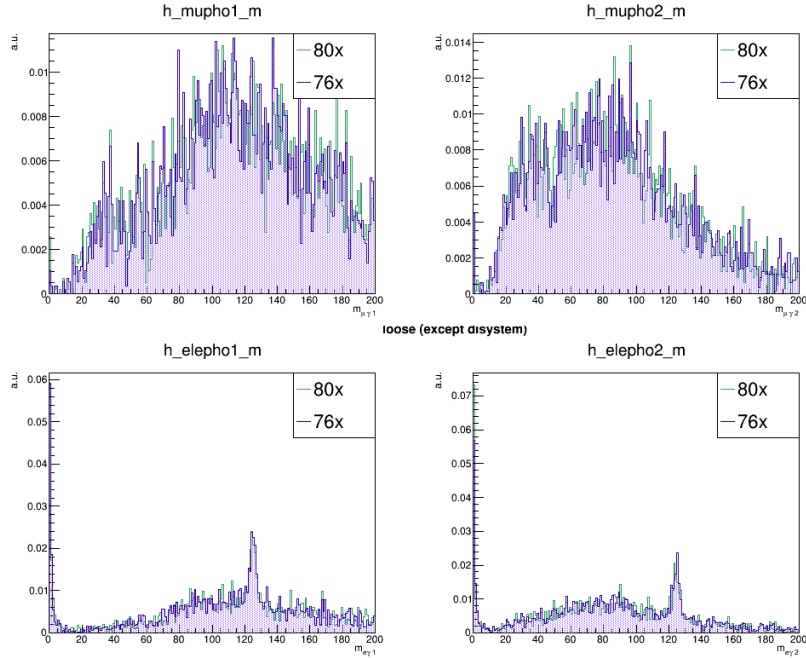


Figure 5.12: Distributions of $m_{\ell\gamma_{(1,2)}}$, to be used in “loose” selection. The selections in “set 1 - loose”, “set 2 - loose” and “set 3 - loose” are applied.

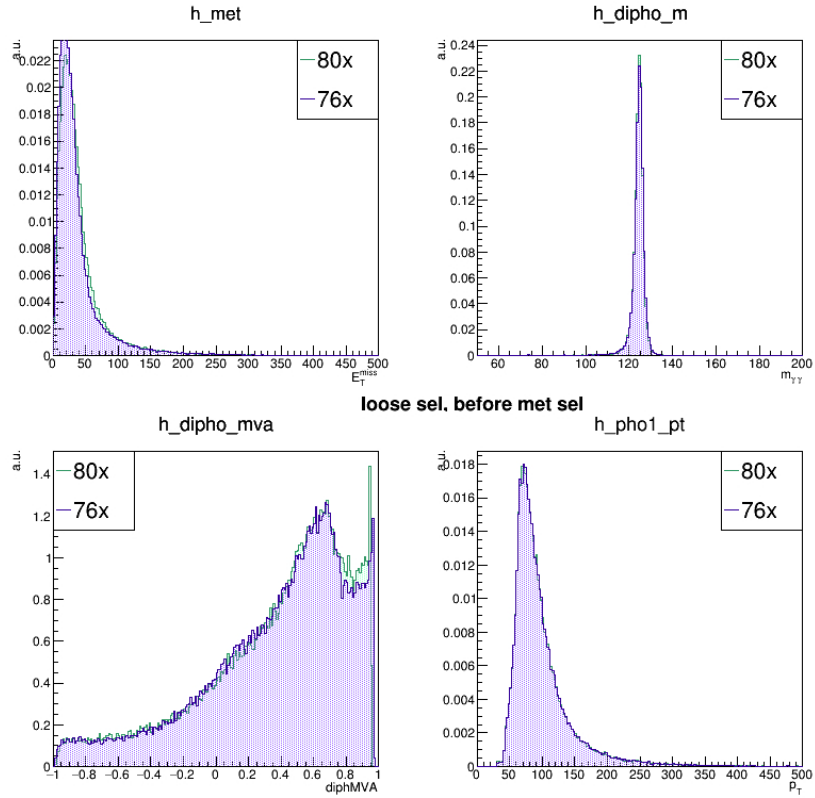


Figure 5.13: Distributions of E_T^{miss} , $m_{\gamma\gamma}$, $diphoMVA$ and $p_T^{\gamma^1}$, to be used in "met" selection, performed on data that did not fulfill the "tight" and "loose" criteria. No selection is applied.

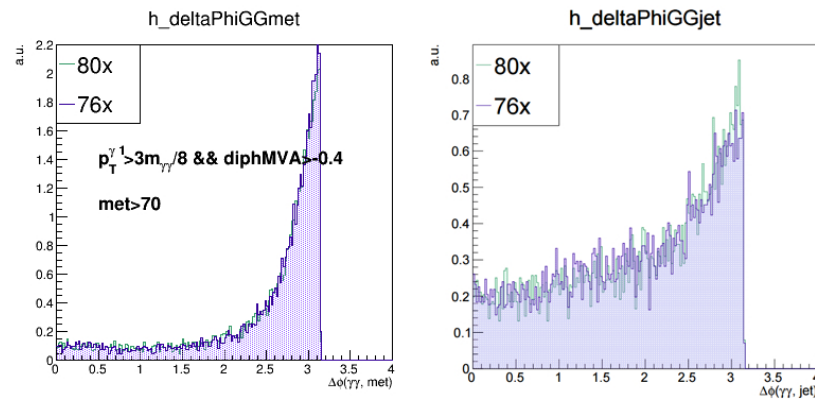


Figure 5.14: Distributions of $\Delta\phi(\gamma\gamma, met)$ and $\Delta\phi(\gamma\gamma, jet)$, to be used in "met" selection. The selections in "set 1 - met" are applied.

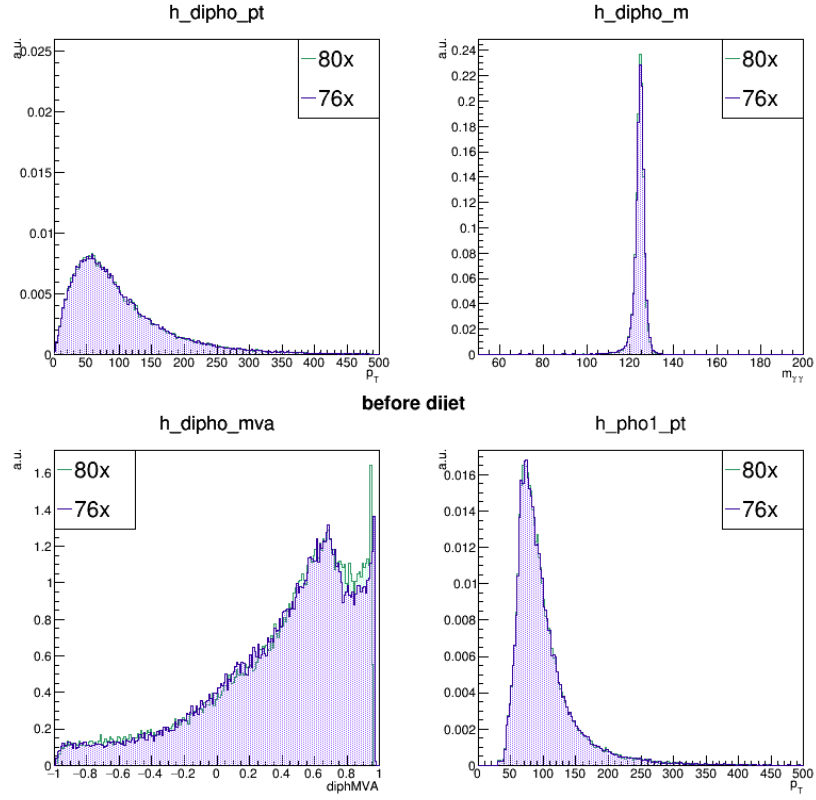


Figure 5.15: Distributions of $p_T^{\gamma\gamma}$, $m_{\gamma\gamma}$, *diphoMVA* and $p_T^{\gamma^1}$, to be used in “dijet” selection, performed on events that did not fulfill the “tight”, “loose” and “met” criteria. No selection is applied.

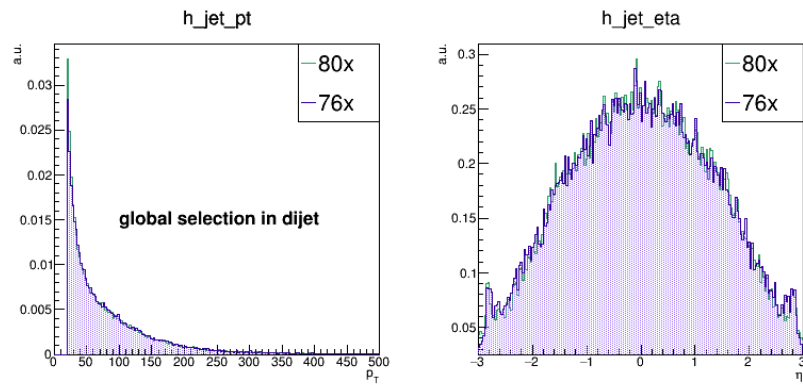


Figure 5.16: Distributions of p_T^j and η^j , to be used in “dijet” selection. The selections in “set 1 - dijet” are applied.

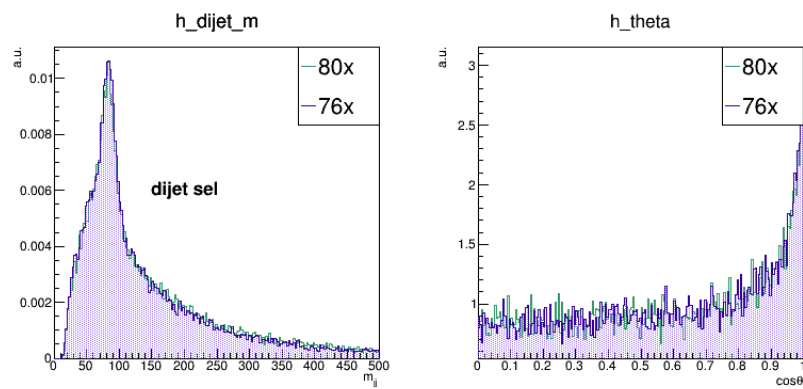


Figure 5.17: Distributions of m_{jj} and $\cos\theta$ to be used in “dijet” selection. The selections in “set 1 - dijet” and “set 3 - dijet” are applied.

Conclusions

This thesis is focused on preliminary studies for the search of the Higgs boson in the VH production channel with decay in two photons in Run2 at 13 TeV. It develops around several aspects that can affect the estimate of the Higgs boson mass resolution: the effect of the material in front of ECAL, the shape of the showers in ECAL at different conditions, the validation of the MC generators and the optimization of the classification of good candidates for the Higgs boson.

The mismodeling of the material of the tracker upstream ECAL was checked to be limited to a 5% in the confidence region up to $|\eta| = 2$, growing to the 20% for higher $|\eta|$, and confirms an improvement of 5% with respect to the Run1 data to simulation agreement (where the mismodeling in central barrel reached the 10%). The description of some features, like the gap between barrel and endcaps or the boxes of cables in the preshower, can still be improved. Overall there is good agreement between data and simulations and the incidence of the mismodeling on the systematics on the measurement of the mass of the Higgs boson have been reduced with respect to Run1.

The study of the shower shape variables showed a discrepancy between data and MC which can be only partially explained with a mismodeling of the material. Other causes of the disagreement were studied and excluded as the most relevant factors: clustering effects are not significant and the choice of the method of signal reconstruction improves the agreements only in the distribution of the η -development of the shower. As a future development, a more detailed study of the influence of the material budget could be performed with a comparison of data and MC with and without the 3.8 T magnetic field. In fact, it was observed that in absence of the field the influence of the tracker material on the development of the shower is less significant. In addition, the amplitude of the signal from out-of-time pileup reconstructed with the “multifit” method was checked in any sample and expected to be the same on average. A different behavior was observed and could not be explained with a dependence on η or on the number of primary vertices. In future, a study of the timing dependence could be performed exploiting the information about the bunch crossing which was not available at the time of this work.

The HRes generator, able to compute cross sections up to the resummed NNLL+NNLO order, was studied and compared to the two fixed order generators that are normally used in CMS, interfaced with parton shower MC: POWHEG and aMC@NLO. Despite their differences, both

the event generators were observed to work reasonably. The p_T distributions of the Higgs boson computed by aMC@NLO for the ggH and VH modes ended up to be unreliable due to the presence of negative weights that are not well treated by the diphoton MVA, but this behavior does not affect the analysis, based on other variables.

A new version of the CMS reconstruction and analysis software for the year 2016 was validated. The classification of promising events for the search of the Higgs boson was performed on simulations, in order to define four categories related to the VH channel. The selections used in Run1 were applied in Run2 to check if they are a good starting point for the updated analysis. Results in agreement with those of Run1 at 8 TeV were obtained, and they were positively compared with other studies on Run2. The classification resulted to be valid, considering that in the final analysis the VBF selection must be performed before the VH selection, the diphoton MVA estimator has to be optimized, the b-tag rejection must be added and the met filters should be implemented.

This preliminary work is just the first step of the final analysis of the Higgs boson that will be implemented by the CMS Collaboration in the near future.

Acronyms

$\gamma\gamma$ = diphoton system

$\ell\ell$ = dilepton system

ADC = analogue-to-digital converter

ALICE = A Large Ion Collider Experiment

APD = avalanche photodiode

ATLAS = A Toroidal LHC ApparatuS

BDT = boosted decision tree

BPix = barrel of the pixel tracker

BX = bunch crossing

CB = Crystal Ball

CERN = European Organization for Nuclear Research

CKM = Cabibbo-Kobayashi-Masakawa

CMS = Compact Muon Solenoid

CSC = cathode strip chamber

DGLAP = Dokshitzer-Gribov-Lipatov-Altarelli-Parisi

DIS = deep inelastic scattering

DS = dispersion suppressor

DT = drift tube

EB = barrel of the electromagnetic calorimeter

ECAL = electromagnetic calorimeter

EE = endcaps of the electromagnetic calorimeter

ES = preshower of the electromagnetic calorimeter

FPix = endcaps of the pixel tracker

ggH = gluon gluon fusion

GSF = Gaussian Sum Filter

HB = barrel of the hadron calorimeter

HE = endcaps of the hadron calorimeter

HF = forward hadron calorimeter

HCAL = hadron calorimeter

HLT = high-level trigger

HPD = hybrid photodiode

HO = outer hadron calorimeter
ID = identification
IP = interaction point
IR = infrared
KF = Kalman Filter
L1 = level-1 trigger
LEP = Large Electron Positron collider
LHC = Large Hadron Collider
LHCb = Large Hadron Collider beauty experiment
LEIR = Low Energy Ion Ring
LINAC = LINear ACcelerator
MC = Monte Carlo
MET = missing transverse energy
MVA = multivariate analysis
NLL = next-to-leading logarithm
NNLL = next-to-next-to-leading logarithm
NLO = next-to-leading order
NNLO = next-to-net-to-leading order
OOT = out-of-time (pileup)
PDF = parton distribution function
PF = particle flow
pp = proton-proton
PS = Proton Synchrotron
PSB = Proton Synchrotron Booster
PU = pileup
QCD = Quantum Chromodynamics
QED = Quantum Electrodynamics
RF = radiofrequency
RMS = root mean square
RPC = resistive plate chamber
SC = supercluster
SL = superlayer
SM = Standard Model
SPS = Super Proton Synchrotron
TEC = tracker endcaps
TIB = tracker inner barrel
TID = tracker inner disk
TOB = tracker outer barrel
TPS = trigger primitive generator
trk = track

ttH = associated production of the Higgs boson with top quarks

UV = ultraviolet

VBF = vector boson fusion

VEV = vacuum expectation value

VH = associated production of the Higgs boson with a vector boson (Higgsstrahlung)

VPT = vacuum phototriode

vtx = vertex

WLS = wavelenght shifter

Bibliography

- [1] David Griffiths. *Introduction to Elementary Particles*. John Wiley & sons, 1987.
- [2] Particle Data Group. Particle physics booklet, 2014.
- [3] S.M. Bilenky and B. Pontecorvo. Lepton mixing and neutrino oscillations. *Phys. Rep.* 41, 225-261, 4 May 1978. doi:10.1016/0370-1573(78)90095-9.
- [4] J. Einasto. Dark Matter. [*astro-ph.CO*], 26 Sep 2011. <http://arxiv.org/abs/1109.5580v1>.
- [5] R.D. Peccei. Matter-Antimatter Asymmetry in the Universe and an Arrow for Time. [*hep-ph*], 20 Aug 2006. <http://arxiv.org/abs/hep-ph/0608226v1>.
- [6] J. F. Donoghue. General relativity as an effective field theory: the leading quantum corrections. *UMHEP-408*, 25 May 1994. <https://arxiv.org/pdf/gr-qc/9405057v1.pdf>.
- [7] W. de Boer. Grand Unified Theories and Supersymmetry in Particle Physics and Cosmology. [*hep-ph*], 7 Feb 2001. <https://arxiv.org/abs/hep-ph/9402266>.
- [8] Mark Thomson. *Modern Particle Physics*. Cambridge University Press, 2013.
- [9] Enrico Fermi. *Zs. f. Phys.* 88, 161, 1934.
- [10] C. S. Wu et al. Experimental Test of Parity Conservation in Beta Decay. *Phys. Rev.* 105, 1413, 15 Jan 1957.
- [11] J. H. Christenson et al. Evidence for the 2π Decay of the K_2^0 Meson. *Phys. Rev. Lett.* 13, 138, 27 Jul 1964.
- [12] Xinyi Zhang and Bo-Qiang Ma. Correlating lepton mixing angles and mixing matrix with Wolfenstein parameters. [*hep-ph*], 13 Nov 2012. <http://arxiv.org/pdf/1206.0519v2.pdf>.
- [13] J. Leite Lopes. Forty Years of the First Attempt at the Electroweak Unification and of the Prediction of the Weak Neutral Boson Z. *Brazilian Journal of Physics*, vol. 29, no. 3, Sep 1999.

- [14] F. Halzen and A. D Martin. *Quarks and Leptons*. John Wiley & sons, 1984.
- [15] Alan D. Martin. Proton structure, Partons, QCD, DGLAP and beyond. [*hep-ph*], 1 Feb 2008. <https://arxiv.org/pdf/0802.0161v1.pdf>.
- [16] Gordon Krane. *Modern Elementary Particle Physics*. Addison-Wesley, 1993.
- [17] T. Lappi F. Gelis and R. Venugopalan. High energy scattering in quantum chromodynamics. [*hep-ph*], 1 Aug 2007. <https://arxiv.org/pdf/0708.0047.pdf>.
- [18] Taizo Muta. *Foundation of Quantum Chromodynamics: an Introduction to Perturbative Methods in Gauge Theories*. World Scientific Lecture Notes in Physics, 1998.
- [19] R. Akhouri et al. The KLN Theorem and Soft Radiation in Gauge Theories: Abelian Case. [*hep-ph*], 6 Feb 1997. <http://arxiv.org/abs/hep-ph/9702270>.
- [20] Davison E. Soper John C. Collins and George Sterman. Factorization of Hard Processes in QCD. [*hep-ph*], 27 Sep 2004. <https://arxiv.org/pdf/hep-ph/0409313.pdf>.
- [21] G. Altarelli and G. Parisi. Asymptotic freedom in parton language. *Nucl. Phys. B* 126, 298, 8 Aug 1977. doi:10.1016/0550-3213(77)90384-4.
- [22] A. Salam J. Goldstone and S. Weinberg. Broken Symmetries. *Phys. Rev.* 127 965970, 1962.
- [23] Abdelhak Djouadi. Higgs Physics: Theory. [*hep-ph*], 19 Mar 2012. <https://arxiv.org/abs/1203.4199v1>.
- [24] Abdelhak Djouadi. The Anatomy of ElectroWeak Symmetry Breaking - Tome I: The Higgs boson in the Standard Model. [*hep-ph*], 3 May 2005. <http://arxiv.org/abs/hep-ph/0503172v2>.
- [25] Abdelhak Djouadi. Higgs production at the LHC. [*hep-ph*], 2 Dec 2010. <http://arxiv.org/abs/1012.0530v5>.
- [26] Robert V. Harlander. Virtual corrections to $gg \rightarrow H$ to two loops in the heavy top limit. [*hep-ph*], 25 Jul 2000. <http://arxiv.org/pdf/hep-ph/0007289.pdf>.
- [27] S. Forte et al. Update on Higgs production in gluon fusion. *9th Workshop of the LHC Higgs cross section Working Group*, 22-24 Jun 2015. <http://indico.cern.ch/event/331452/contributions/1721496/>.
- [28] S. Dawson et al. Associated Top Quark-Higgs Boson Production at the LHC. [*hep-ph*], 28 Nov 2002. <https://arxiv.org/pdf/hep-ph/0211438v1.pdf>.
- [29] M. Lethuillier et al. Search for a neutral Higgs boson with WH / ZH, $H \rightarrow \gamma\gamma$ channel. *CMS NOTE 2006/110*, 30 May 2006. http://inspirehep.net/record/726816/files/NOTE2006_110.pdf?version=1.

-
- [30] C. Delaere. Study of associated WH production with $H \rightarrow WW^{(*)}$ in the 3 leptons final state. *CMS NOTE 2006/053*, 4 May 2006. http://inspirehep.net/record/721285/files/NOTE2006_053.pdf?version=1.
- [31] C. Oleari T. Figy and D. Zeppenfeld. Next-to-leading order jet distributions for Higgs boson production via weak-boson fusion. [*hep-ph*], 12 Jun 2003. <https://arxiv.org/pdf/hep-ph/0306109v1.pdf>.
- [32] P. Bolzoni et al. Higgs production at NNLO in QCD: the VBF channel. [*hep-ph*], 29 Jun 2010. <http://arxiv.org/pdf/1006.2323.pdf>.
- [33] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. [*hep-ex*], 31 Jul 2012. <https://arxiv.org/pdf/1207.7235v1.pdf>.
- [34] CMS Collaboration. Search for the standard model Higgs boson decaying into two photons in pp collisions at $\sqrt{s} = 7$ TeV. [*hep-ex*], 7 Feb 2012. <http://arxiv.org/pdf/1202.1487v1.pdf>.
- [35] CMS Collaboration. Search for the standard model Higgs boson in the decay channel $H \rightarrow ZZ \rightarrow 4\ell$ in pp collisions at $\sqrt{s} = 7$ TeV. [*hep-ex*], 9 Feb 2012. <http://arxiv.org/pdf/1202.1997v1.pdf>.
- [36] CMS Collaboration. Search for the standard model Higgs boson decaying to W^+W^- in the fully leptonic final state in pp collisions at $\sqrt{s} = 7$ TeV. [*hep-ex*], 7 Feb 2012. <http://arxiv.org/pdf/1202.1489v1.pdf>.
- [37] CMS Collaboration. Search for neutral Higgs bosons decaying to tau pairs in pp collisions at $\sqrt{s} = 7$ TeV. [*hep-ex*], 12 Nov 2014. <http://arxiv.org/pdf/1202.4083v2.pdf>.
- [38] CMS Collaboration. Search for the standard model Higgs boson decaying to bottom quarks in pp collisions at $\sqrt{s} = 7$ TeV. [*hep-ex*], 12 Nov 2014. <http://arxiv.org/pdf/1202.4195v2.pdf>.
- [39] ATLAS and CMS Collaborations. Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV. [*hep-ex*], 7 Jun 2016. <https://arxiv.org/pdf/1606.02266v1.pdf>.
- [40] ATLAS Collaboration. Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS detector at the LHC. [*hep-ex*], 31 Aug 2012. <http://arxiv.org/abs/1207.7214v2>.
- [41] Eilam Gross and Ofer Vitells. Trial factors for the look elsewhere effect in high energy physics. [*physics.data-an*], 7 Oct 2010. <http://arxiv.org/pdf/1005.1891v3.pdf>.

- [42] ATLAS and CMS Collaborations. Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments. [*hep-ex*], 26 Mar 2015. <http://arxiv.org/pdf/1503.07589v1.pdf>.
- [43] CMS Collaboration. The CMS experiment at the CERN LHC. *JINST* 3 (2008) S08004, 14 Aug 2008. doi:10.1088/1748-0221/3/08/S08004.
- [44] ATLAS Collaboration. The ATLAS experiment at the CERN Large Hadron Collider. 2008 *JINST* 3 S08003, 14 Aug 2008. <http://iopscience.iop.org/1748-0221/3/08/S08003>.
- [45] ALICE Collaboration. The ALICE experiment at the CERN LHC. *JINST* 3 (2008) S08002, 14 Aug 2008. doi:10.1088/1748-0221/3/08/S08002.
- [46] LHCb Collaboration. The LHCb experiment at the CERN LHC. *JINST* 3 (2008) S08005, 14 Aug 2008. doi:10.1088/1748-0221/3/08/S08005.
- [47] Lyndon Evans and Philip Bryant. LHC machine. *JINST* 3 S08001, 14 Aug 2008. doi:10.1088/1748-0221/3/08/S08001.
- [48] The LHC Study Group. LHC - The Large Hadron Collider.
- [49] Oliver S. Bruning et al. The LHC main ring. *CERN libraries, LHC Design Report, vol. 1*, 4 Jun 2004.
- [50] Oliver S. Bruning et al. The LHC infrastructure and general services. *CERN libraries, LHC Design Report, vol. 2*, 5 Nov 2004.
- [51] J.P. Blewett. 200 gev intersecting storage accelerators. *Proceedings of the 8th International Conference on High-Energy Accelerators*, 20-24 Sep 1971.
- [52] K. M. Potter. The Large Hadron Collider (LHC) project of CERN. *LHC Project Report 36*, 19 Jul 1996. <https://cds.cern.ch/record/308243/files/lhc-project-report-36.pdf>.
- [53] E. Keil. The CERN Large Hadron collider LHC. *LHC Project Report 83*, Nov 1996. <https://cds.cern.ch/record/321820/files/lhc-project-report-83.pdf>.
- [54] S. Peggs. Parasitic beam-beam collisions and crossing angles in rhic. *RHIC/AP/66*, Aug 1995.
- [55] M. Benedikt et al. The LHC injector chain. *CERN libraries, LHC DESIGN REPORT, vol. 3*, 15 Dec 2004.
- [56] A. Beuret et al. The LHC Lead Injector Chain. *Proceedings of EPAC 2004*, 2004. <http://accelconf.web.cern.ch/AccelConf/e04/PAPERS/TUPLT011.PDF>.

-
- [57] CMS Collaboration. The CMS experiment at the CERN LHC. [*physics.ins-det*], 14 Aug 2008. <http://stacks.iop.org/1748-0221/3/i=08/a=S08004>.
- [58] CMS Collaboration. The Magnet Project Technical Design Report. *CERN-LHCC-97-10*, May 1997.
- [59] CMS Collaboration. Precise Mapping of the Magnetic Field in the CMS Barrel Yoke using Cosmic Rays. [*physics.ins-det*], 4 Jan 2010.
- [60] CMS Collaboration. Tracker Technical Design Report. *CERN-LHCC-98-006 ; CMS-TDR-5*, 15 Apr 1998.
- [61] CMS Collaboration. Addendum to the CMS tracker Technical Design Report. *CERN-LHCC-2000-016 ; CMS-TDR-5-add-1*, 21 Feb 2000.
- [62] CMS Collaboration. The Electromagnetic Calorimeter Technical Design Report. *CERN/LHCC 9733, CMS TDR 4*, 15 Dec 1997.
- [63] CMS Collaboration. The Hadron Calorimeter Technical Design Report. *CERN/LHCC 97-31; CMS-TDR-2*, 20 Jun 1997.
- [64] CMS Collaboration. The Compact Muon Solenoid Technical Design Report. *CERN/LHC 97-32; CMS-TDR-3*, 15 Dec 1997.
- [65] CMS Collaboration. Cms the tridas project technical design report, volume 1: The trigger systems. *CERN/LHCC 2000 - 38, CMS TDR 6.1*, 15 Dec 2000.
- [66] CMS Collaboration. Cms the tridas project technical design report, volume 2: Data acquisition and high-level trigger. *CERN/LHCC 02-26, CMS TDR 6*, 15 Dec 2002.
- [67] Ildefons Magrans de Abril et al. Conceptual Design of the CMS Trigger Supervisor. *IEEE Transactions on Nuclear Science*, VOL. 53, NO. 2, Apr 2006.
- [68] Claus Grupen and Boris Shwartz. *Particle Detectors*. Cambridge Monographs, 2008.
- [69] Heinrich J.Volk and Konrad Bernlohr. Imaging Very High Energy Gamma-ray Telescopes. [*astro-ph*], 22 Dec 2008. arXiv:0812.4198v1.
- [70] M. Anfrevilleb et al. Laser monitoring system for the CMS lead tungstate crystal calorimeter. *CMS NOTE-2007/028*, 15 Nov 2006.
- [71] Federico Ferri. Monitoring and Correcting for Response Changes in the CMS Lead-tungstate Electromagnetic Calorimeter. *CMS CR -2012/259*, 8 Oct 2012.
- [72] CMS Collaboration. Energy calibration and resolution of the CMS electromagnetic calorimeter in pp collisions at $\sqrt{s} = 7$ TeV. *JINST 8 P09009*, 4 Oct 2013. <http://arxiv.org/abs/1306.2016>.

- [73] CMS Collaboration. Performance and operation of the CMS electromagnetic calorimeter. *JINST 5 T03010*, 19 Mar 2010. doi:10.1088/1748-0221/5/03/T03010.
- [74] Paolo Meridiani. The CMS Electromagnetic Calorimeter Calibration strategy - 9th ICATPP Conference 17-21 Oct 2005.
- [75] 2015 ECAL detector performance plots - in CMS-DP-2015/057. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/EcalDPGResultsCMSDPS2015057>.
- [76] CMS Collaboration. Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV. [*physics.ins-det*], 1 Jul 2015. <https://arxiv.org/abs/1502.02701>.
- [77] CMS Collaboration. Performance of photon reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV. [*physics.ins-det*], 21 Aug 2015. <https://arxiv.org/abs/1502.02702>.
- [78] CMS Collaboration. Updated measurements of the Higgs boson at 125 GeV in the two photon decay channel. *CMS PAS HIG-13-001*, 21 Mar 2013.
- [79] CMS Collaboration. First measurements of the Higgs boson production in the diphoton decay channel at $\sqrt{s} = 13$ TeV. *CMS PAS HIG-15-005*, 22 Mar 2016.
- [80] Florian Beaudette. The CMS Particle Flow Algorithm - Proceedings of CHEF 2013.
- [81] Susanna Cucciarelli. Track and Vertex Reconstruction with the CMS Detector - Proceedings of the 10th International Conference in B-Physics at Hadron Machines, BEAUTY 2005.
- [82] A. Hoecker et al. TMVA 4 Toolkit for Multivariate Data Analysis with ROOT - Users Guide. [*physics.data-an*], 7 Jul 2009. <http://arxiv.org/pdf/physics/0703039v5.pdf>.
- [83] CMS Collaboration. Updated measurements of Higgs boson production in the diphoton decay channel at $\sqrt{s} = 13$ TeV in pp collisions at CMS. *CMS PAS HIG-16-020*, 4 Aug 2016.
- [84] Fengwangdong Zhang. Performance of the CMS jets and missing transverse energy trigger for the upgraded LHC RunII. *PoS(EPS-HEP2015)290*, 22-29 Jul 2015. http://pos.sissa.it/archive/conferences/234/290/EPS-HEP2015_290.pdf.
- [85] CMS Collaboration. Further measurement of $h \rightarrow \gamma\gamma$ at $\sqrt{s} = 13$ tev. *CMS AN-16-209*, 16 Jul 2016.
- [86] CMS Collaboration. A search using multivariate techniques for a standard model Higgs boson decaying into two photons. *CMS PAS HIG-12-001*, 19 Apr 2012.

-
- [87] F Distribution Tables - Statistics Online Computational Resource. http://www.socr.ucla.edu/applets.dir/f_table.html.
- [88] Fabrice Couderc. Higgs to 2 photons in CMS - LHC France 4 Apr 2013.
- [89] Tracker Material Budget plots. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/TrackerMaterialBudgetplots>.
- [90] Arabella Martelli. fBrem studies to estimate the material budget - Weekly Tracker DPG Meetings 8-12 Jul 2013.
- [91] Stéphanie Baffioni. About “material”: electrons occupancy, fbrem, nhits - ECAL $e\gamma$ meeting.
- [92] Giulia Negro. Data-MC Studies at 0T and 3.8T - ECAL Joint DPG/MoCa Meeting 13 Apr 2016.
- [93] The MadGraph5_aMC@NLO homepage. <http://madgraph.physics.illinois.edu/>.
- [94] E. Di Marco and J. Bendavid. ECAL local reconstruction with pileup - joint ECAL/ $e\gamma$ meeting 21 Aug 2014.
- [95] J. Bendavid and E. Di Marco. Ecal OOT Pileup Mitigation with Template Fits in Pulse Reconstruction - ECAL DPG Meeting 28 Aug 2014.
- [96] E. Di Marco and J. Bendavid. Local reconstruction - ECAL calibration workshop 15 Oct 2015.
- [97] M. Grazzini D. de Florian, G. Ferrera and D. Tommasini. Higgs boson production at the LHC: transverse momentum resummation effects in the $H \rightarrow \gamma\gamma$, $H \rightarrow WW \rightarrow \ell\nu\ell\nu$ and $H \rightarrow a\ell$ decay modes. [*hep-ph*], 28 Mar 2012. <http://arxiv.org/pdf/1203.6321.pdf>.
- [98] Massimiliano Grazzini and Hayk Sargsyan. Heavy-quark mass effects in Higgs boson production at the LHC. [*hep-ph*], 27 Sep 2013. <https://arxiv.org/pdf/1306.4581v2.pdf>.
- [99] C. Oleari. The POWHEG box. [*hep-ph*], 22 Jul 2010. <https://arxiv.org/pdf/1007.3893v1.pdf>.
- [100] J. Alwall et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. [*hep-ph*], 21 Jul 2014. <https://arxiv.org/pdf/1405.0301.pdf>.
- [101] Hayk Sargsyan. Higgs pt: Hres - ggf pt and jet-bin modelling workshop (cern) 7 May 2015.

- [102] Massimiliano Grazzini. Higgs cross section at NNLL+NNLO (current recommendation) - meeting on the ggF inclusive cross section 18 Nov 2014.
- [103] Mikhail Rogal Alexey Pak and Matthias Steinhauser. Finite top quark mass effects in NNLO Higgs boson production at LHC. [*hep-ph*], 24 Nov 2009. <https://arxiv.org/pdf/0911.4662v1.pdf>.
- [104] Massimiliano Grazzini. The Higgs pT spectrum in the SM and beyond - Higgs Couplings 2014 (turin).
- [105] Julien Baglio and Abdelhak Djouadi. Higgs production at the LHC. [*hep-ph*], 2 Jun 2012. arXiv:1012.0530v5.
- [106] David Hall. *Discovery and Measurement of the Higgs Boson in the WW Decay Channel*. Springer International Publishing, 2015.
- [107] CMS Collaboration. Measurement of differential cross sections for Higgs boson production in the diphoton decay channel in pp collisions at $\sqrt{s} = 8$ TeV. [*hep-ex*], 15 Jan 2016. <https://arxiv.org/pdf/1508.07819.pdf>.
- [108] P. Nason and B. Webber. Next-to-leading order generators. [*hep-ph*], 6 Feb 2012. <http://arxiv.org/pdf/1202.1251.pdf>.
- [109] Simone Alioli et al. NLO Higgs boson production via gluon fusion matched with shower in POWHEG. *JHEP04(2009)002*, 1 Apr 2009. <http://iopscience.iop.org/article/10.1088/1126-6708/2009/04/002/pdf>.
- [110] N. Marinelli M. Planer and C. Jessop. VH tag update - Hgg meeting 11 Jul 2016.
- [111] Federico Ferri. Some checks of the VH tag - Higgs Working Meetings 4-8 Jul 2016.
- [112] CMSSW Reference Manual. <https://cmssdt.cern.ch/SDT/doxygen/>.

Ringraziamenti

Ringrazio, innanzitutto, Federico per avermi seguita in tutti questi mesi, insegnandomi un sacco di cose e sopportando i miei frequenti momenti ansiogeni.

Ringrazio Shervin per avermi fornito i dati e avermi sopportata anche la duecentesima volta che gli chiedevo con insistenza un nuovo Monte Carlo.

Ringrazio Fabio, non solo per aver condiviso le sue nozioni di fisica, ma anche per le chiacchierate politico-culturali durante le giornate svizzere di inizio estate alla CMS week.

Ringrazio Franz perché ero tanto preoccupata all'idea del controrelatore e invece lui è stato di una gentilezza incredibile, ha mostrato molto interesse verso l'argomento di tesi e ha apprezzato il mio lavoro.

Ringrazio Vazzo per aver movimentato le pause dallo studio in ufficio, perdendoci in siti psichedelici così come discutendo animatamente di cucina e questioni sociali.

Ringrazio i coinquilini di questi anni, in particolare Ewa, la prima persona che ho incontrato a Trieste e che mi ha fatta subito sentire a casa con le serate Prison Break e quelle in cui si affogava la tensione dell'università e della convivenza in stuzzichini e vodka polacca, Luigi che in così poco tempo è diventato un pilastro importante della mia vita quotidiana condividendo la mia nerdaggine e la mia passione per la cucina e la fotografia, Margherita che ha portato un po' del calore di giù in questa città fredda e ventilata.

Ringrazio lo YoCus, sempre in piena, un'oasi di allegria elettrizzante!

Ringrazio tutti i compagni di corso con cui ho condiviso questa bellissima sofferenza chiamata magistrale. Li ringrazio per avermi traumatizzata con la lubianska quando ero a Trieste da tipo appena due giorni, per le cene etniche, il pessimismo ontologico, le passeggiate tra boschi e torri sperdute, le serate a teatro (alcune piacevoli, altre da strapparsi occhi e timpani), la fonduta di cioccolato all'ultimo sangue, i progetti di viaggi improbabili (che sì, ci credo, riusciremo a fare! 3 donne in giro per l'Iran è la versione 2.0 di Thelma e Louise, magari senza il tuffo dal Canyon xD), le discussioni in lab, le sessioni di ripasso di QFT all'ultimo momento, i tuffi a Miramare per scaricare l'ansia dopo un esame, i caffè condivisi, i pranzi ai tavolini arancioni e gli sguardi di comprensione in aula informatica, in agosto, deserta.

I want to thank all the people who shared with me my Erasmus experience, six unforgettable months that made me understand which is the direction that I want to take in my life. I thank who brought me to their far and exotic countries through stories, images, sounds, smells and flavors. Who helped me with the French and also who laughed at my weird pronunciation

(it's only motivational!). Who spent weekends and nights at the IPN, studying together and solving endless exercises. Who shared the post-GR beer at the ignorant bar near Diderot. Who followed me into improbable events, from the pseudo-gipsy Italian group to the Sudanese concert in a squat house. Who followed me into the improvised trips to Caen and Strasbourg. Who followed me into the cultural journeys at the museums or into those more “bohémien” with a bottle of cider along the Canal. Who worried for me in the moments of darkness and fear. Who laughed and smiled with me in the moments of happiness.

I want to thank all the professors who wrote recommendation letters for my PhD application and wished me to find my way.

I thank my future PhD supervisor for compensating my pessimism and for believing that I could achieve my goal... and he was right, finally I did it!

Ringrazio, col senno di poi, anche chi in questi anni, più o meno consapevolmente, mi ha messo i bastoni tra le ruote o ha infranto per un istante i miei sogni e le mie speranze perché mi ha ricordato che ho la forza per andare avanti. Sempre. Una forza che mi ha portata dove sono ora, contro ogni aspettativa.

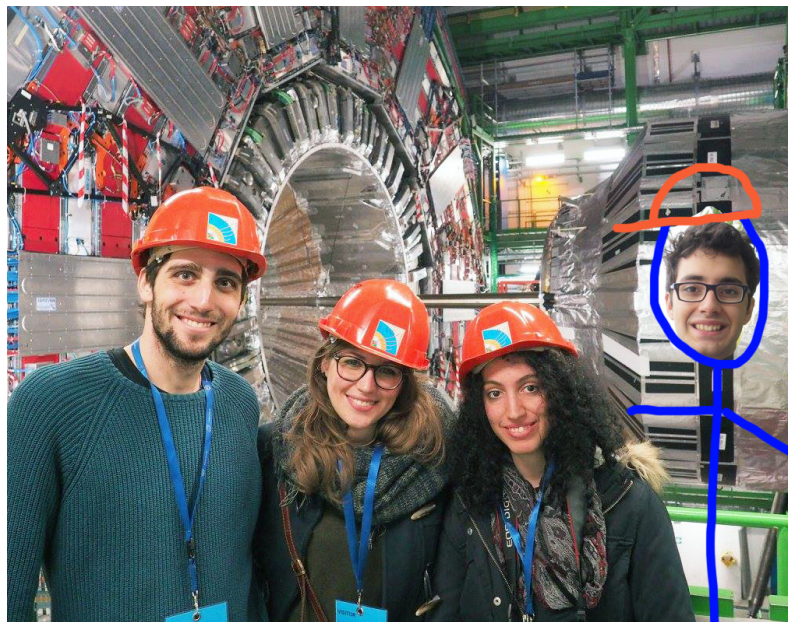


Figure 5.18: Che fisici, eh?!

Ringrazio l'Italian corner (vi siete pure meritati una foto photoshoppata con tanto di CMS sullo sfondo :p), un angolo di casa a mille chilometri di distanza, che mi ha ricordato gli obiettivi e perché vale la pena di sudare, faticare e piangere, che è sempre lì nel momento del bisogno o per farsi una risata, che mi ha accompagnata negli ultimi cinque anni e si è riunito ancora una volta oltralpe, che mi ha ricordato che “la strada giusta è sempre quella in salita” e “ogni fiume deve svoltare almeno tre volte per raggiungere il mare”.

Ringrazio Benedetta perché anche se non ci vediamo per periodi lunghi ogni volta che ci sentiamo è tuttora come la sorella che non ho mai avuto, che sa dire sempre la cosa giusta per farmi sorridere...e spero che un giorno le nostre strade s'incrocino nuovamente, "hermana querida".

Ringrazio gli amici di Ravenna, in particolare Romina, che nonostante le difficoltà e la distanza, c'è e ci sarà sempre.

Ringrazio i miei scout, che sono sempre in giro, ma resteranno sempre una delle fette più importanti della mia vita e che mi ha resa quella che sono.

Ringrazio mia nonna perché mi vuole un bene che non si può descrivere a parole, sente la mia mancanza, ma le brillano gli occhi quando le parlo dei miei piani futuri con termini incomprensibili.

Ringrazio, infine (last but not least!), i miei genitori che non mi hanno fatto mai mancare nulla pur insegnandomi il valore delle cose. Che mi hanno sempre lasciata spiccare il volo anche quando faceva paura. Che sono parte di me anche quando non ci vediamo per mesi. Che riescono a capirmi sempre, anche quando non parlo, anche quando sono così incasinata che non mi capisco neppure io. Li ringrazio perché la famiglia è il posto in cui imparare cosa significa l'amore incondizionato, al di là dello spazio e del tempo. E io l'ho imparato!

*“Celui qui n’ aime nul être ignore
la joie de vivre”.*

Saadi