Université Catholique de Louvain

Secteur des sciences et technologies
Institut de Recherche en Mathématique et Physique
Centre for Cosmology, Particle Physics and Phenomenology

UCL Université catholique de Louvain

Search for a heavy scalar boson in the $ZZ \to l^+l^-\nu\bar{\nu}$ decay channel and reconstruction of high-energy muons in the CMS experiment

Doctoral dissertation presented by

Alessio Magitteri

in fulfilment of the requirements for the degree of Doctor in Sciences

Jury de thèse

Prof. Vincent Lemaitre (presidént)	UCL, Belgium
Prof. Giacomo Bruno (promoteur)	UCL, Belgium
Prof. Fabio Maltoni	UCL, Belgium
Prof. Pascal Vanlaer	ULB, Belgium
Prof. Petra Merkel	Fermilab, US

April, 2018

Contents

Chapter 1

Theoretical Introduction and Motivation

With the observation of the Higgs Boson by the CMS and ATLAS in 2012 [1, 2] the Standard Model of Particle Physics (SM), which represents the maximum expression of human understanding of the infinitely small, is considered complete since all its predictions are confirmed. However both theoretical and experimental evidence indicates that the SM cannot be the ultimate model of Nature. It should rather be considered as an effective description of a more fundamental theory that should emerge at a higher energy scale. During the last decades a lot of extensions of the SM (BSM) were proposed, but no evidence has yet been found of any of them after nine years of activity of the Large Hadron Collider (LHC) [3]. The present work is driven by the desire to contribute to address this challenge and it is based on the search for a generic heavy scalar resonance using the $ZZ \to 2l2\nu$ topology.

1.1 Phenomenological Description of the Standard Model

The Standard Model of particle physics describes two fundamental interactions of the sub-atomic word: the *electroweak* force and the *strong* force. The first one is obtained via the unification of two forces, the electromagnetic interaction from one side and the weak interaction on the other side. They are responsible for the electromagnetic phenomena and some decays, such as the neutron disintegration (β decay) or the top quark decay, respectively. Their unification in the same theoretical model was proposed independently by S.Glashow, A.Salam, and S.Weinberg during the 60ies [4–6], but not before the 1980s with the direct measurements of the heavy vector bosons ($W^{\pm}$ and Z) its validation was achieved. The experiments involved in the measurements were the UA1 and UA2 experiment all based at CERN [7, 8].

If the study of the electroweak interaction has seen an asynchronous development between the experimental side and the theoretical part, the exploration of the strong interaction has seen a more cogent collaboration between the two sides. The basis of the understanding of the strong interaction had been set during the 50ies through the formulation of a static model of quarks capable of classifying all the known hadrons. The first experiment aiming to the study of the internal structure of the hadrons goes back to R.Hofstader thanks to the electron scattering in atomic nuclei [9]. The experiments during the 1970s laid for the development of the deep-inelastic scattering at the hand of J.I.Friedman, H.W.Kendall, and R.Taylor in the US at SLAC using electrons [10] and the Gargamelle experiment at CERN using deep-inelastic scattering with neutrinos [11]. The inputs coming from the static model of quarks and the results of the deep-inelastic scattering put the basis for the development of a perturbative theory of quarks, the *Quantum CromoDynamics* (QCD) [12–14], whose validity was sanctioned by the e^+e^- experiment Mark-I [15], at SLAC in US, and Tasso [16] at DESY in Germany.

All the particles predicted by the SM and their property are summarized in the portrait 1.1. The entire group of fundamental particles can be classified in two big families according to the value of their intrinsic angular momentum: fermions with spin equal to 1/2 and bosons with spin equal to 1. The fermions can be divided in two other big families: leptons and quarks. The leptons are affected only by the electroweak interaction mediated by the neutral bosons γ and Z and the charged bosons W^+ and W^-. In addition to the electroweak interaction, the quarks are affected by the strong force mediated by the gluons.

All the fermions can be grouped in three different families but a fourth generation made by heavier particles is disfavoured by the recent discovery of the Higgs Boson [17, 18]. Indeed, a fourth generation would increase strongly the gluon fusion rate and

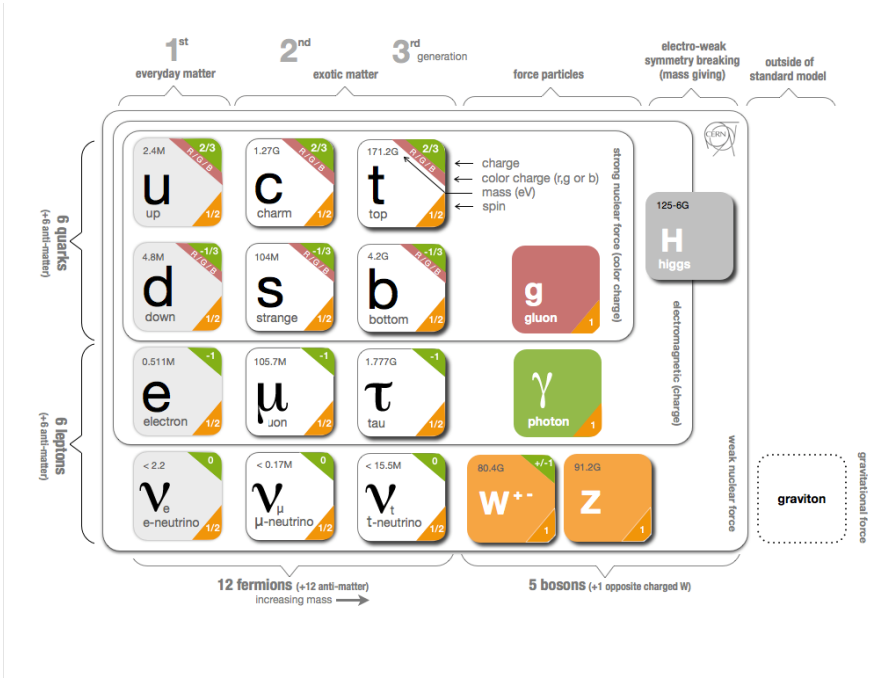

Figure 1.1: Schematic representation of the Standard Model.

decrease the diphoton decay rate. In particular the ZZ and WW decay channels would be enhanced contradicting the current measurement.

All families enter in the SM in the same way but they differ for their mass range. Indeed the particles of each next generation are significantly much heavier then the previous one. The heaviest fermion is the top quark, t, with a mass equal to 173 GeV and it was discovered in 1995 by CDF and $D0$ collaborations at the Tevatron collider in the US [19, 20]. The big variety of the elementary particles is not reflected in the composition of the common matter which is based only on the first generation particles (e^-, ν_e, u, d), the other two families can be produced only in laboratory. Despite terseness of the common matter, the quarks can produce a big variety of particles, called hadrons, this is the only way in which the quarks manifest themselves in nature. Due to the *color confinement* [21], quarks (except for the top quark), can't stay in a *free-particle* mode but first try to get rid of their color charge by radiating gluons, end eventually combine in pairs or triplets in order to create a colorless state, the hadron. The hadrons formed by a couple of quark-antiquark are called *mesons*, the *baryons*, as the proton, are instead composed by a triplet of partons. Hadrons with four quarks are also present in nature as reported in one of the latest results of the LHCb collaboration [22].

The interactions between fermions are mediated by five different vector bosons: the photon, γ, the heavy bosons, $W^{\pm}, Z$ and the gluons, g, which are responsible respectively of the electromagnetic interaction, the weak interaction and the strong interaction. The electromagnetic force is a long range interaction thanks to the massless photon, whereas the weak interaction is mediated by three heavy bosons ($m_W^{\pm}$ = 80.4 GeV, m_Z = 91.2 GeV) which limit the interaction range of the weak force. The gluons are massless and their scale of interaction is of the order of the nucleus. The mass of the elementary particles is obtained by interaction with the Higgs field [23–28], mediated by the Higgs boson which is the only scalar particle of the SM and has mass of 125 GeV. The discovery in 2012 of the Higgs boson [1, 2] completed the particle spectrum of the Standard Model.

1.2 The theory of the Standard Model

The Standard Model is a renormalizable quantum field theory (QFT) where each particle and each interaction is identified by a field defined in each point of the space-time. The model is developed in terms of the Lagrangian formalism and the entire theoretical structure is based on the *Noether Theorem* [29] which claims that for a Lagrangian invariant under a continuous group of transformations corresponds a conservation law and a conserved symmetry. In the context of the QFT, these transformations are known as *gauge* transformations and the associated theoretical model is defined as gauge invariant. Two group of transformation can be defined: a global transformation and a local transformation. The first group acts identically and simultaneously in all the points of the space-time associated to the Lagrangian, instead a local transformation depends on each point of the space-time in which it is applied and the associated variations can be independent by each other.

The SM is a gauge theory based on the local symmetry

$$G_{SM} = SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \tag{1.1}$$

The sub-group $SU(3)_C$ governs the strong interaction described by the *Quantum Cromodynamics* (QCD). The $SU(3)_C$ symmetry group is based on eight generators (Gell-Man matrices) to which the local invariance under the color symmetry associate eight gauge fields: the gluons. The QCD Lagrangian for each quark flavor is

$$\mathscr{L}_{QCD} = \bar{q}_{ij}\big(i\gamma_\mu D^\mu_{ij} - \delta_{ij} m_q\big) q_j - \frac{1}{4} G^a_{\mu\nu} G^{\mu\nu}_a, \tag{1.2}$$

where i and j are the indices of the fundamental representations of the $SU(3)_C$ group. The covariant derivative is defined as

$$D_\mu = \partial_\mu - i g_s \frac{\lambda^a}{2} G^a_\mu, \tag{1.3}$$

g_s is the coupling of the strong interaction, λ^a is the Gell-Man matrices and $G^a_{\mu\nu}$ is the gluon field strength tensor defined as

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc} G^b_\mu G^c_\nu. \tag{1.4}$$

where f^{abc} are the structure constants of the group. A specific feature of the Quantum Cromodynamics is the dependancy of g_s from the energy. Indeed, it decreases as a function of the energy transferred in the process making the strong interaction weaker with the increasing of the energy and stronger when the energy decreases. This property is known as *asymptotic freedom* and it explains the *color confinement* such that the quarks can be always observed bounded together in hadrons. Moreover, the symmetry group $SU(3)$ generated by λ^a, is non-abelian and it translates in the fact

that gluons present both cubic and quartic self-interaction terms. Finally, the gauge boson fields have to be massless to preserve the local symmetry.

The symmetry group $SU(2)_L \otimes U(1)_Y$ governs the electroweak interaction which is described by the following Lagrangian

$$\begin{aligned}\mathscr{L}_{EW} = & i\bar{Q}_L\gamma_\mu D^\mu Q_L + i\bar{u}_R\gamma_\alpha D^\alpha u_R + i\bar{d}_R\gamma_\beta D^\beta d_R + \\ & + i\bar{L}_L\gamma_\phi D^\phi L_L + i\bar{e}_R\gamma_\tau D^{tau} e_R + \\ & - \frac{1}{4}W^i_{\mu\nu}W_i^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}.\end{aligned} \tag{1.5}$$

The covariant derivative is defined as

$$D_\mu = \partial_\mu - ig\frac{\sigma^i}{2}W^i_\mu - ig'\frac{Y}{2}B_\mu \tag{1.6}$$

where g and g' are the couplings for the group $SU(2)_L$ and $U(1)_Y$ respectively. σ^i are the Pauli matrices, and Y is the weak hypercharge of the field on which the derivative acts. The field strength tensors is defined as

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + g\epsilon^{ijk}W^i_\mu W^k_\nu, \tag{1.7}$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{1.8}$$

where ϵ^{ijk} is the Levi-Civita symbol.

As expressed in the Lagrangian 1.5, the $SU(2)_L \otimes U(1)_Y$ group effects differently on the left-handed and the right-handed chiral state which composed each fermion field. $SU(2)_L$ sub-group acts only on the left-handed component of each fermion, which are grouped in doublets

$$Q^i_L = \begin{pmatrix} u^i_L \\ d^i_L \end{pmatrix}, \quad L^i_L = \begin{pmatrix} \nu^i_L \\ e^i_L \end{pmatrix}, \tag{1.9}$$

where the index i runs on the family generation. Differently, the right-handed fermions ν^i_R, e^i_R, u^i_R, d^i_R are not affected by $SU(2)_L$ transformations. On the contrary, $U(1)_L$ sub-group acts on all fermions both left and right-handed except for the right-handed neutrinos, whose $U(1)_Y$ charge is zero. Three gauge bosons correspond to the $SU(2)_L$ symmetries and one field to $U(1)_Y$, which are identified respectively as W^i_μ, $i = 1, 2, 3$, and B_μ. The conserved quantity under $SU(2)_L$ transformation is called isospin, whereas the charged which is conserved under $U(1)_Y$ transformation is called hypercharge.

1.2.1 The Electroweak symmetry breaking

Although from experimental results elementary particles are massive, from a theoretical point of view in order to keep the electroweak sector invariant under the $SU(2)_L \otimes U(1)_Y$ transformation, any explicit mass term is present in the Lagrangian. The Brout-Englert-Higgs (BEH) mechanism [23–28] is the elegant solution adopted to restore the mass term in it. It describes the spontaneous breaking of the $SU(2)_L \otimes U(1)_Y$ symmetry down to $U(1)_{Em}$ and it generates masses for both the weak bosons and the fermions, as discussed below.

The keystone of the BEH mechanism is the introduction of a quadratic potential V($\boldsymbol{\phi}$) defined for a $SU(2)_L$ doublet of complex scalar fields, $\boldsymbol{\phi}$, invariant under the gauge group $SU(2)_L \otimes U(1)_Y$. The field $\boldsymbol{\phi}$ has hypercharge equal to unity and is obtained as a linear combination of four real and scalar fields

$$\boldsymbol{\phi} = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{1.10}$$

whereas the potential is defined as following

$$V(\boldsymbol{\phi}) = \frac{1}{2}\mu^2\boldsymbol{\phi}^2 + \frac{1}{4}\lambda\boldsymbol{\phi}^4. \tag{1.11}$$

and it is dependent from the parameter μ^2 and λ which has to be positive in order to ensure an absolute minimum in the Lagrangian. If μ^2 is greater than zero, the potential has one minimum which corresponds to the vacuum expectation value $\langle 0 \left| \boldsymbol{\phi} \right| 0 \rangle$. On the other hand, if μ^2 is lower than zero, the potential presents an unstable local maximum in $|\boldsymbol{\phi}| = 0$ and a ground state which is not unequivocally defined anymore because it lies on a hypersurface identified by the following equation

$$|\boldsymbol{\phi}|^2 = -\frac{\mu^2}{2\lambda}. \tag{1.12}$$

Without loosing any kind of generality, the minimum can be choosen as

$$\langle \boldsymbol{\phi} \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu \end{pmatrix} \tag{1.13}$$

where ν is the vacuum expectation value, defined by the following relation $\nu = \sqrt{-\mu^2/\lambda}$. Due to the choice of a precise value of the minimum, the ground state is no longer invariant under $SU(2)_L \otimes U(1)_Y$ transformations, but it remains invariant under the electromagnetic symmetry $U(1)_{Em}$. The scalar field on the vacuum state can be parametrized as following

$$\boldsymbol{\phi}(\boldsymbol{x}) = \frac{e^{i\tau\cdot\vec{\psi}(x)/\nu}}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu + h(x) \end{pmatrix} \tag{1.14}$$

in which four new scalar fields emerge. The three fields ψ_i identified the unphysical and massless *Goldstone* bosons, instead $h(x)$ defines the Higgs field. With the choice of the unitarity gauge, the Goldstone bosons can be reabsorbed into the theory, leaving only the Higgs field in the definition of the scalar field $\boldsymbol{\phi(x)}$

$$\boldsymbol{\phi(x)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu + h(x) \end{pmatrix}. \tag{1.15}$$

The BEH mechanism is described by the Lagrangian

$$\mathscr{L}_{Higgs} = (D_\mu \boldsymbol{\phi})^\dagger D^\mu \boldsymbol{\phi} - V(\boldsymbol{\phi}^\dagger \boldsymbol{\phi}), \tag{1.16}$$

where the potential is the one just described and the covariant derivative is defined as for $\mathscr{L}_{EW}$. The mass term for the electroweak gauge boson can be obtained substituting into the Lagrangian the explicit expression of the scalar field $\boldsymbol{\phi}$ after the symmetry break. It can be proven that the mass of the heavy scalar bosons are $gv/2$ and $v\sqrt{g^2 + g'^2}/2$ respectively for the W and Z boson, while the photon remain massless. The mass of the Higgs boson is $v\sqrt{2\lambda}$.

BEH mechanism is used also to recover the mass of the fermions by the Yukawa interaction between the fermions and the Higgs field. This term is defined by the following Lagrangian

$$\begin{aligned} \mathscr{L}_{Yukawa} = &-Y^u_{ab} \bar{Q}_{La} \tilde{\boldsymbol{\phi}} u_{Rb} - Y^d_{ab} \bar{Q}_{La} \boldsymbol{\phi} d_{Rb} + \\ &- Y^e_{ab} \bar{L}_{La} \boldsymbol{\phi} e_{Rb} - Y^\nu_{ab} \bar{L}_{La} \boldsymbol{\phi} \nu_{Rb} + h.c. \end{aligned} \tag{1.17}$$

with $\tilde{\boldsymbol{\phi}} = i\sigma_2 \boldsymbol{\phi}^*$, $Y^{u,d,e,\nu}_{ab}$ represents the Yukawa coupling for the different fermions family identified by a and b. As in the case of the gauge bosons, also in this case the mass of the different fermions can be obtained by the explicit substitution of the Higgs field and it can be proven that the mass of the fermions is $y_f v/\sqrt{2}$ where y_f is the corresponding fermion component of the diagonalized Yukawa matrix.

1.3 Beyond Standard Model examples

1.3.1 Motivation to look beyond the SM

Despite the discovery of the Higgs boson and its consistency with all the experimental facts, the SM can't be considered the ultimate theory of Nature. Experimental and theoretical considerations in support of this aspect are:

- The SM has 18 free parameters in total and they can't be justified by the model itself. All the ordinary matter can be constructed out of four fermions of the first family, even if the model predicts three generations which have no obvious role in common life, furthermore there is no explanation both of their masses and the hierarchical mass structure they have (there are 5 orders of magnitude between the top quark mass and the electron mass, and the discrepancies increase further if the neutrino mass is taken into account). Another important aspect related to the parameters is the amount of CP violation: while the CP violation measured in laboratory is properly predicted by the theory, this measurement is not enough to explain the CP violation required in the universe to explain its matter-antimatter asymmetry.

- In order to properly describe both the dynamic, the structure of the universe, and its evolution the modern cosmology has to introduce the dark matter. At the moment there is no way to incorporate dark matter inside the SM except to extend it by the introduction of new symmetries like in Supersymmetry [30].

- Another important aspect is the "naturalness" (sometimes referred to as hierarchy) problem. The Higgs mass in the SM is fixed by mean of an experimental parameter. However, if new physics was to appear first at the Planck scale, then new heavy particles (M « new scales) should be there, coupling to the Higgs field and thus inducing loop corrections to its physical mass; these correction terms would be extremely large so, in order to keep the physical mass of the Higgs to the 125 GeV region, such large corrections should almost exactly cancel in a statistically unlikely fine-tuned manner. This problem is known as the "naturalness" problem, as such level of fine-tuning seems quite unnatural for a fundamental theory.

- Gravity is not unified with the other interactions, and there is no obvious way to generate General Relativity within the standard model context. In addition a hierarchy problem appears between the scale of the EW spontaneous symmetry breaking, about 250 GeV, and the scale at which the quantum gravity should appear, the Planck scale, about 10^{19} GeV.

During the last decades a lot of extensions of the SM were proposed to the high energy physics community in order to solve some of these problems. Among all the available strategies one possibility is to introduce new scalar particles. This is the approach implemented in the Electroweak Singlet Model (EWS) and in the two-Higgs-doublet Model (2HDM). The next lines will be devoted to their description.

1.3.2 Electroweak Singlet Model

The Electroweak Single Model is the simplest extension of the SM Higgs sector, where an additional real scalar field is added. It contains a complex scalar doublet ($\boldsymbol{\phi(x)}$), and an additional real scalar field S which is singlet under the SM gauge group. The Lagrangian which describes the EWS model is

$$\mathscr{L}_{EWS} = (D^\mu \boldsymbol{\phi})^\dagger D_\mu \boldsymbol{\phi} + \partial^\mu S \partial_\mu S - V(\boldsymbol{\phi}, S), \tag{1.18}$$

where the potential is defined as

$$V(\boldsymbol{\phi}, S) = -m^2 \boldsymbol{\phi}^\dagger \boldsymbol{\phi} - \mu^2 S^2 + \lambda_1 (\boldsymbol{\phi}^\dagger \boldsymbol{\phi})^2 + \lambda_2 S^4 + \lambda_3 (\boldsymbol{\phi}^\dagger \boldsymbol{\phi}) S^2, \tag{1.19}$$

Denoting with v and x the expectation value of the Higgs field $\boldsymbol{\phi}$ and S respectively, in the unitarity gauge the scalar boson fields become

$$\boldsymbol{\phi} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \tilde{h} + v \end{pmatrix} \quad S = \frac{h' + x}{\sqrt{2}} \tag{1.20}$$

After the diagonalization of the mass matrix, the EWS presents two mass eigenstates defined by a linear combination of $\tilde{h}$ and h'

$$\begin{pmatrix} h \\ H \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \tilde{h} \\ h' \end{pmatrix} \tag{1.21}$$

where the mixing angle α is given by

$$\sin 2\alpha = \frac{\lambda_3 x v}{\sqrt{(\lambda_1 v^2 - \lambda_2 x^2)^2 + (\lambda_3 x v)^2}}, \tag{1.22}$$

$$\cos 2\alpha = \frac{\lambda_2 x^2 - \lambda_1 v^2}{\sqrt{(\lambda_1 v^2 - \lambda_2 x^2)^2 + (\lambda_3 x v)^2}}. \tag{1.23}$$

The mass values of h and H are

$$m_h^2 = \lambda_1 v^2 + \lambda_2 x^2 - \sqrt{(\lambda_1 v^2 - \lambda_2 x^2)^2 + (\lambda_3 x v)^2}, \tag{1.24}$$

$$m_H^2 = \lambda_1 v^2 + \lambda_2 x^2 + \sqrt{(\lambda_1 v^2 - \lambda_2 x^2)^2 + (\lambda_3 x v)^2}. \tag{1.25}$$

and by convention $m_h \leq m_H$. From the equation 1.21 follows that the light h boson couplings to SM particles are suppressed by $\cos\alpha$ factor and the heavy scalar particle by a factor $\sin\alpha$.

The model is completely defined by five physical parameters: m_h, m_H, $\sin\alpha$, $\tan\beta \equiv v/x$, and v. In the case the light scalar boson h is identified as the SM Higgs, the EWS is totally described by three independent parameters

$$m_H, \sin\alpha, \tan\beta. \tag{1.26}$$

This particular case corresponds to the *decoupling region* which correspond to the case when $\sin\alpha \to 0$.

The Figure 1.2 reports the available phase space for the mixing angle $\sin\alpha$ in the high mass region in the case different theoretical and experimental constrains are taken into account [31]. The direct Higgs searches with the measurement of the Higgs signal strength constrain the most the mixing angle value in the region with $m_H \leq 250$ GeV, whereas the region with 250 GeV $\leq m_H \leq$ 800 GeV is constrained by the measurement on the W boson mass. Finally, the region with $m_H \geq 800$ GeV is constrained by perturbativity of the couplings λ_i up to high energy scale ($\mu = 4 \cdot 10^{10}$ GeV) [32].

1.3.3 The two-Higgs-doublet Model

Although the 2HDM [34] represents the second simplest extension of the SM, its structure is able to answer some open questions mentioned before as the dark matter [35], the asymmetry between matter and antimatter observed in the Universe [36], and both the CP-violation in the Higgs sector [37] and the non-CP violation of the QCD sector [38].

The 2HDM is characterized, as suggested by its name, by two complex doublets (1.27):

$$\phi_1 = \begin{pmatrix} \phi_1^+ \\ \phi_1^0 \end{pmatrix} \quad \phi_2 = \begin{pmatrix} \phi_2^+ \\ \phi_2^0 \end{pmatrix} \tag{1.27}$$

The presence of an invisible Z boson ($Z \to 2\nu$) in the $H \to ZZ \to 2l2\nu$ decay channel, doesn't allow any kind of measurement of the CP property of the heavy scalar resonance. For this reason the following discussion is focused on the particular case of the conservation of the CP-simmetry in the Higgs sector. The most general potential $V = (\phi_1, \phi_2)$ imposing the gauge invariance, has the following form (1.28):

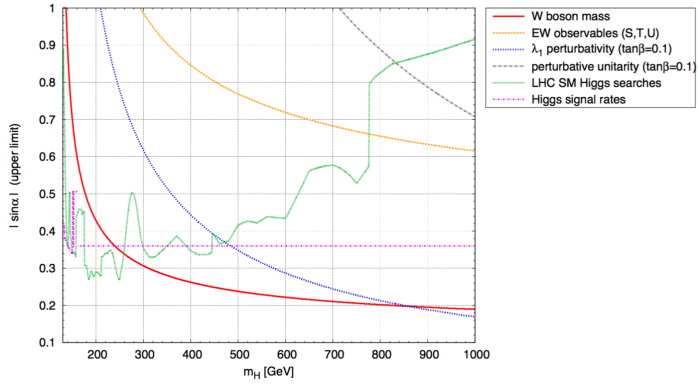

Figure 1.2: Upper limit of the mixing angle value in the mass region between 130 GeV up to 1 TeV combining both theoretical and experimental constrains. The direct LHC Higgs measurements and the Higgs signal strength measurement are drawn in green and magenta respectively. The constrain associated to the measurement of the W boson mass is reported in red, whereas the precision measurements of electroweak observables via S,T and U [33] are reported in orange. Finally the distribution in blue and grey report respectively the allowed region for the mixing angle due to the perturbative requirements on the λ_i couplings and the perturbative unitarity of the $2 \to 2$ S-matrix associated to the model evaluated requiring $\tan\beta = 0.1$ [31].

$$\begin{aligned} V(\phi_1,\phi_2) = {} & m_{11}^2\phi_1^\dagger\phi_1 + m_{22}^2\phi_2^\dagger\phi_2 - m_{12}^2\left(\phi_1^\dagger\phi_2 + \phi_1\phi_2^\dagger\right) + \\ & + \frac{\lambda_1}{2}\left(\phi_1^\dagger\phi_1\right)^2 + \frac{\lambda_2}{2}\left(\phi_2^\dagger\phi_2\right)^2 + \\ & + \lambda_3\phi_1^\dagger\phi_1\phi_2^\dagger\phi_2 + \lambda_4\phi_1^\dagger\phi_2\phi_2^\dagger\phi_1 + \\ & + \frac{\lambda_5}{2}\left[\left(\phi_1^\dagger\phi_2\right)^2 + \left(\phi_2^\dagger\phi_1\right)^2\right] \end{aligned} \tag{1.28}$$

where all the parameters are real and the minimum of the potential is defined by (1.29):

$$\langle\phi_1\rangle_0 = \begin{pmatrix} 0 \\ \frac{v_1}{\sqrt{2}} \end{pmatrix} \quad \langle\phi_2\rangle_0 = \begin{pmatrix} 0 \\ \frac{v_2}{\sqrt{2}} \end{pmatrix} \tag{1.29}$$

With two complex doublets correspond five scalar fields: two charged bosons $H^\pm$, one pseudo-scalar A and finally two neutral scalar boson h, H.

$$\phi_a = \begin{pmatrix} \phi_a^+ \\ (v_a + \rho_a + i\eta_a)/\sqrt{2} \end{pmatrix} \tag{1.30}$$

where $a = 1, 2$, with $v_1 = v\cos\beta$ and $v_2 = v\sin\beta$. In particular, the neutral scalar bosons, *h* and *H*, are an orthogonal combinations of ρ_1 and ρ_2:

$$h = \rho_1 \sin\alpha - \rho_2 \cos\alpha \tag{1.31}$$

$$H = -\rho_1 \cos\alpha - \rho_2 \sin\alpha \tag{1.32}$$

and the SM Higgs boson can be obtained via a linear combination of *h* and *H*:

$$h^{SM} = h\sin(\alpha-\beta) - H\cos(\alpha-\beta). \tag{1.33}$$

In the specific case of the *decoupling region* defined as the phase space region where $\cos(\beta-\alpha) \ll 1$, the $\mathrm{h}^{SM} \approx h$ [39].

The model is fully defined by six free parameters:

$$M_A, M_H, M_{H^\pm}, M_{12}^2, \tan(\beta), \cos(\beta-\alpha) \tag{1.34}$$

where $\tan(\beta)$ is defined as the ratio between v_2 and v_1 and α is the rotation angle which diagonalizes the mass-squared matrices of the neutral scalars.

The 2HDM is constrained by three experimental considerations:

- the measurement of the SM Higgs boson with mass of 125 GeV constrains the parameter space regions to not depart from SM-like condition $\cos(\beta-\alpha) \ll 1$ corresponding to the *decoupling region*.
- v_1 and v_2 have to satisfy the following relation: $v = \sqrt{v_1^2 + v_2^2}$ = 246 GeV which is the vacuum expectation value of the Higgs field.
- the suppression of tree-level flavor-changing neutral current (FCNC) obliged precise interactions between fermions and the Higgs doublet.

In particular due to the last item, four different Types of 2HDM can be distinguished. The interaction between the bosons and the fermions is Type dependent and it is determined by the following Yukawa Lagrangian:

$$\begin{aligned}\mathscr{L} = &-\sum_{f=u,d,l}\frac{m_f}{v}\left(\xi_h^f\overline{f}fh + \xi_H^f\overline{f}fH - i\xi_A^f\overline{f}\gamma_5 fA\right) + \\ &-\left\{\frac{\sqrt{2}V_{ud}}{v}\overline{u}\left(m_u\xi_A^u P_L + m_d\xi_A^d P_R\right)dH^+ + \frac{\sqrt{2}m_l\xi_A^l}{v}\overline{\nu}_L l_R H^+ + H.c.\right\}\end{aligned} \tag{1.35}$$

	Type-I	**Type-II**	**Type-III**	**Type-IV**
ξ_h^u	$\cos\alpha/\sin\beta$	$\cos\alpha/\sin\beta$	$\cos\alpha/\sin\beta$	$\cos\alpha/\sin\beta$
ξ_h^d	$\cos\alpha/\sin\beta$	$-\sin\alpha/\cos\beta$	$\cos\alpha/\sin\beta$	$-\sin\alpha/\cos\beta$
ξ_h^l	$\cos\alpha/\sin\beta$	$-\sin\alpha/\cos\beta$	$-\sin\alpha/\cos\beta$	$\cos\alpha/\sin\beta$
ξ_H^u	$\sin\alpha/\sin\beta$	$\sin\alpha/\sin\beta$	$\sin\alpha/\sin\beta$	$\sin\alpha/\sin\beta$
ξ_H^d	$\sin\alpha/\sin\beta$	$\cos\alpha/\cos\beta$	$\sin\alpha/\sin\beta$	$\cos\alpha/\cos\beta$
ξ_H^l	$\sin\alpha/\sin\beta$	$\cos\alpha/\cos\beta$	$\cos\alpha/\cos\beta$	$\sin\alpha/\sin\beta$
ξ_A^u	$\cot\beta$	$\cot\beta$	$\cot\beta$	$\cot\beta$
ξ_A^d	$-\cot\beta$	$\tan\beta$	$-\cot\beta$	$\tan\beta$
ξ_A^l	$-\cot\beta$	$\tan\beta$	$\tan\beta$	$-\cot\beta$

Table 1.1: Yukawa couplings of u, d, l to the neutral scalar bosons h, H, A in the different 2HDM Types. The couplings for $H^\pm$ boson follow Eq. 1.35.

	Type-I	**Type-II**	**Type-III**	**Type-IV**
u^i	ϕ_2	ϕ_2	ϕ_2	ϕ_2
d^i	ϕ_2	ϕ_1	ϕ_2	ϕ_1
l^i	ϕ_2	ϕ_1	ϕ_1	ϕ_2

Table 1.2: Overview of the four 2HDM Types according on the way the fermions couple to the Higgs doublets. u^i identifies all the quarks with charged 2/3, d^i the quarks with charged -1/3, and finally l^i all the charged leptons. The parameter i runs over all the flavor families. At each 2HDM Type is associated the complex doublet interacting with the corresponding fermion reported in the right colon.

where $P_{L/R}$ are projection operators for left-/right-handed fermions, and the factors ξ are reported in the Table 1.1. The Table 1.2 summarizes all the possible combinations allowed by the model between the fermions and the Higgs doublets.

As it emerges from both the equation 1.35 and the table 1.1, the couplings to the fermions are the same of the SM rescaled by a factor which contains the informations of the mixing angle α and β. The coupling of the neutral scalar bosons to the $W^\pm$ and the Z bosons are the same in all the Types: the couplings of h and H to WW and ZZ are the same of the SM coupling times $\sin(\beta-\alpha)$ and $\cos(\beta-\alpha)$ respectively. The coupling of the pseudo-scalar, A, to the vector bosons vanishes as the one between the neutral scalar bosons and the photons.

Hints of phenomenology of the 2HDM

After this very general description of the 2HDM, it would be obvious wondering about the implications that the recent Higgs discovery has on the model just described.

Actually, a lot of work has been done on this side over the last few years [40–44]. In particular some limits on the 2HDM phase space were set fixing the mass of the light CP-even Higgs h to 125 GeV and taking into account both theoretical considerations (i.e vacuum stability, unitarity) and experimental results (i.e measurements on the Higgs Boson performed both by ATLAS and CMS experiments in different decay channels and the precision measurements for the Z boson production process performed during LEP era [45]). The available region is reported in the Figure 1.3 by the dark red contour in the $\tan(\beta)$ versus $\sin(\alpha-\beta)$ plane. As it can be seen from the picture, three regions are still available: the two narrow region at $\sin(\beta-\alpha)=\pm 1$ and one extended region with $0.55 < \sin(\beta-\alpha) < 0.9$. A similar exclusion plot can be obtained considering the mass of the remaining scalar bosons versus $\sin(\beta-\alpha)$ or $\tan(\beta)$, the corresponding results are reported in figure 1.4.

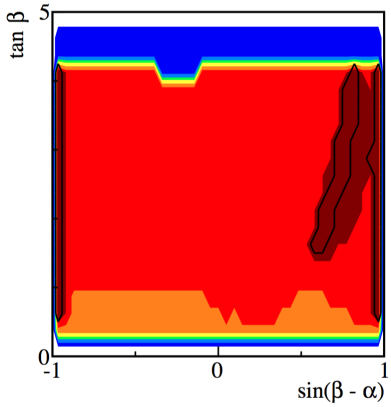

Figure 1.3: The available 2HDM phase space region defined by $\tan(\beta)$ and $\sin(\beta-\alpha)$ in the assumption that the light 2HDM higgs is the one measured by ATLAS and CMS. The results report the specific case of the 2HDM-Type-II scenario. Each color is associated to a different constrain condition: the blue region is associated to the theoretical constrains discussed in the text, the green and orange regions show the available region after the constrains associated to theLEP results on the Z boson measurements, and finally the yellow contour represents the region which is still available after the LHC measurements. The final regions that survive to the combination of all the theoretical and experimental constraints, and the requirement that the light CP-even Higgs (h) has the $gg \to h \to \gamma\gamma, WW/ZZ$ cross-section compatible with the one measured by ATLAS and CMS experiments [40] is the dark red region.

Targeting the heavy CP-even boson, the Higgs search bounds from the LHC removes a large region in negative $\sin(\beta-\alpha)$ space, instead the positive side of this variable is less constrained since $gg \to H \to WW/ZZ$ is much more suppressed [40].

Requiring further to *h* to fit the observed Higgs signal, the available region is further narrows down the favored as shown in dark red bands. For $\sin(\beta - \alpha) = \pm 1$, m_H could be as large as 650 GeV whereas for $0.55 < \sin(\beta - \alpha) < 0.9$, m_H is constrained to be less than 300 GeV (Fig. 1.4). The Fig. 1.4 reports the exclusion region as a function of m_H and $\tan(\beta)$. In this case the excluded $\tan(\beta)$ region is below 0.5 and above 4. The correlation between the mass of the heavy scalar boson and $\sin(\beta - \alpha)$ indicates that if it is discovered to be between 300 GeV and 650 GeV, then $\sin(\beta - \alpha)$ is constrained to be very close to ± 1, indicating the light Higgs has SM-like couplings to the gauge sector.

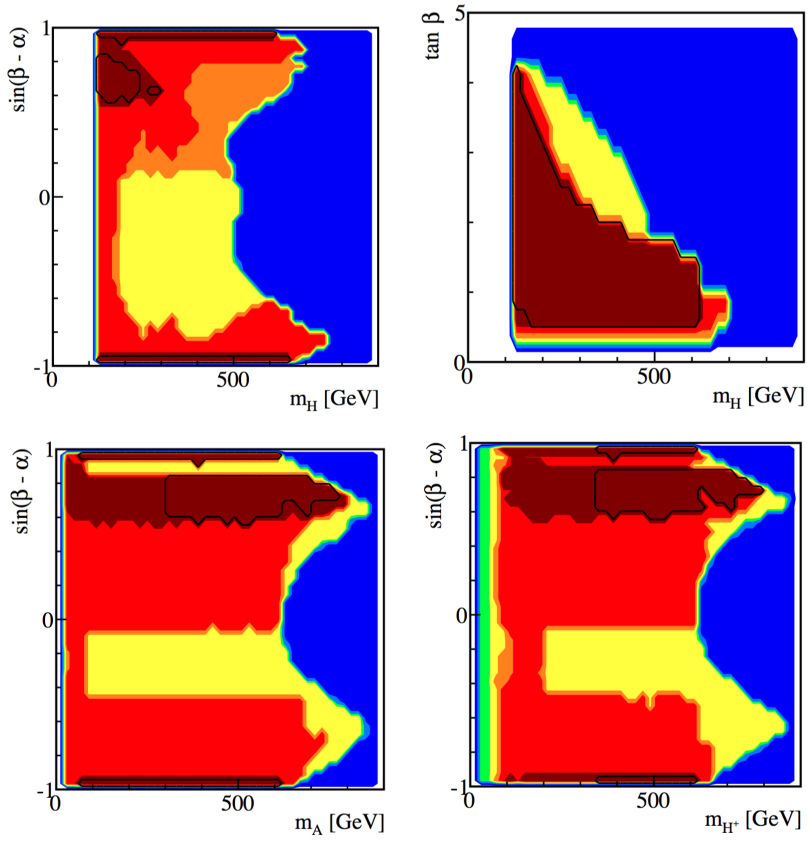

Figure 1.4: The distributions present the phase space region excluded as a function of $\sin(\beta - \alpha)$ (*top left*) and $\tan(\beta)$ (*top right*) versus the mass of the heavy CP-even Higgs (H), the CP-odd Higgs (A) (*bottom left*), and charged Higgs ($H^{\pm}$) (*bottom rigth*) in the 2HDM Type-II scenario. The results as a function of $\tan(\beta)$ (*top right*) are obtained assuming $\sin(\beta - \alpha) \geq 0$. The meaning of the colors is the same as in the Fig. 1.3 [40].

Focusing now a bit more in the region still allowed by the previous results, it could be interesting to analyze some properties that characterize the production of the CP-even Higgs in the ZZ channel. The results reported in the Figure 1.5 show the width range associated to a heavy scalar boson with the mass between 200 GeV up to 1 TeV. As can be seen, the heavy scalar boson width stretch a big range of values from a dozens of GeV up to hundreds of GeV for each mass point.

Another important aspect to consider is the rate at which this boson is expected to be produced in the ZZ channel. To have a better idea of the cross-section at stake,

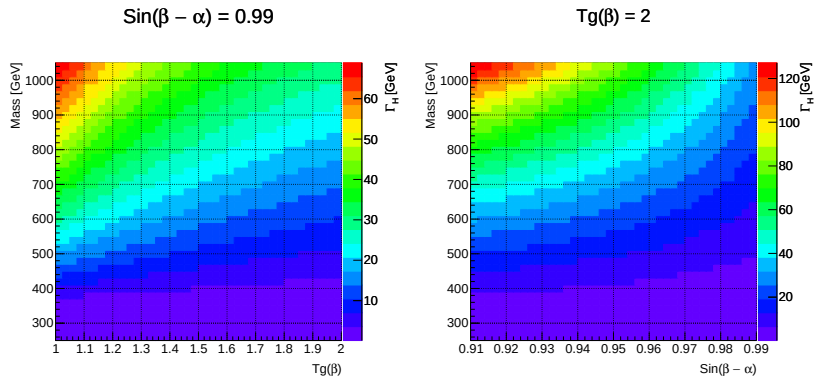

Figure 1.5: Distributions showing the range of the width for a heavy scalar boson *H* as a function of its mass, $\sin(\beta - \alpha)$, and $\tan(\beta)$ in the 2HDM Type-II scenario. The result on the *left* is obtained in the case the $\sin(\beta - \alpha) = 0.99$, whereas the plot on the *right* reports the case where the $\tan(\beta)$ is fixed to 2. All these results were obtained privately using *SusHi* code [46] assuming that $m_A = m_H + 30$ GeV, $m_{H^\pm} = m_H + 50$ GeV. These studies were used as phenomenological support to the analysis performed in $ZZ \to 2l2\nu$ using the data collected by CMS experiment during 2015 and targeting 2HDM. These results are not reported in this manuscript but are available in this article [47].

it could be interesting to look to both the distributions reported in Fig. 1.6. They report the gluon fusion cross section and the fraction of time the heavy scalar boson decays in two Zs. In both of them the contour distributions define the allowed region still available in the $\tan(\beta)$-$\sin(\beta - \alpha)$ plane, assuming for the light scalar boson *h* the same $h \to ZZ$ branching ratio estimated by the "Higgs Cross Section working group" [48] and measured by CMS experiment [49] in the context of the SM Higgs. As can be seen from the distribution of the cross-section, its value is tightly connected to both the m_H and the mixing angles α and β. Focusing on the allowed region, the cross section varies from a fistful of picobarn as for 350 GeV or 650 GeV, up to a dozens of femtobarn for masses close to 1 TeV. Comparing now this numbers with the different cross sections at 13 TeV reported in Fig. 2.1, it is evident the magnitude of the discrepancy between the total proton-proton cross section or the W boson cross section with the gluon fusion cross section in the 2HDM: about 11 orders of magnitude separate the total cross section of a proton-proton collision and the gluon fusion in the 2HDM, whereas 5 orders of magnitude considering the W boson production. Roughly speaking, this means that each 10^{11} *p-p* collisions there will be a single CP-even Higgs boson *H* and 10^5 W vector bosons. The discrepancies are further increased as soon as the proper decay chain is taken into account. All these considerations draw the

conclusion that two ingredients are needed in order to be sensitive to a possible 2HDM signature: a huge amount of collisions at the energy of tens of TeV. The Large Hadron Collider [3] (LHC) was built to fulfill both the requirements in order to prove the SM Higgs and to discover hints of new physics. For a more detailed description of the LHC can be found in the second Chapter of this thesis.

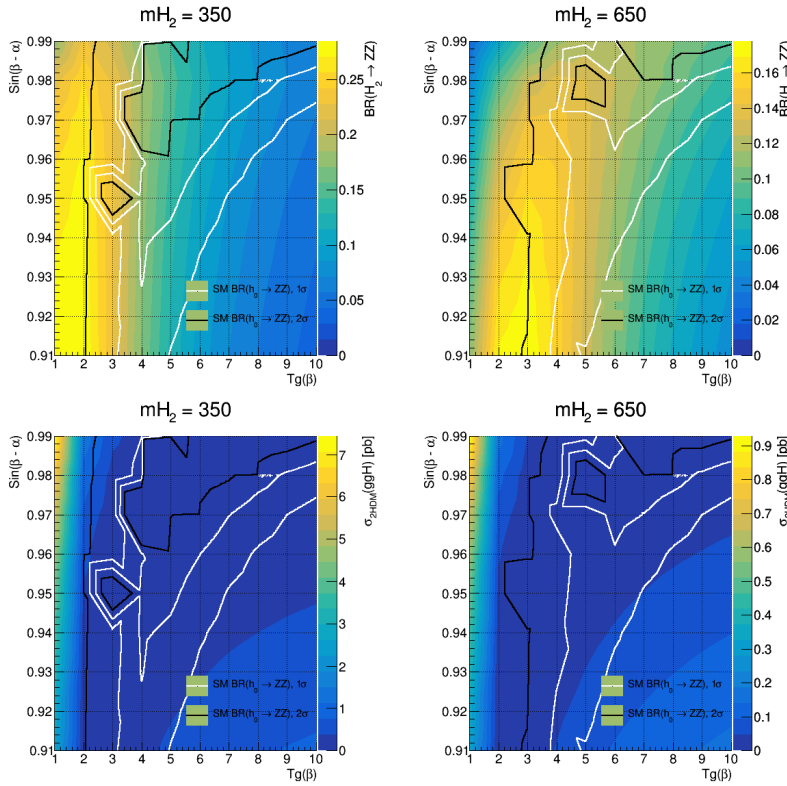

Figure 1.6: Gluon fusion cross section (*bottom*) and branching ratio in ZZ channel (*top*) for a heavy scalar boson in 2HDM Type-II scenario as a function of $\tan(\beta)$ and $\sin(\beta - \alpha)$ variables. The contour distributions report the allowed region in the case the decay in ZZ of the light higgs *h* is constrained to be compatible with the SM Higgs boson branching ratio considering both theoretical considerations [48] and experimental measurements [49]. All these results were obtained privately using *SusHi* code [46] assuming that $m_A = m_H + 30$ GeV, $m_{H^\pm} = m_H + 50$ GeV. These studies were used as phenomenological support to the analysis performed in $ZZ \to 2l2\nu$ using the data collected by CMS during 2015 and targeting 2HDM. These results are not reported in this manuscript but are available in this article [47].

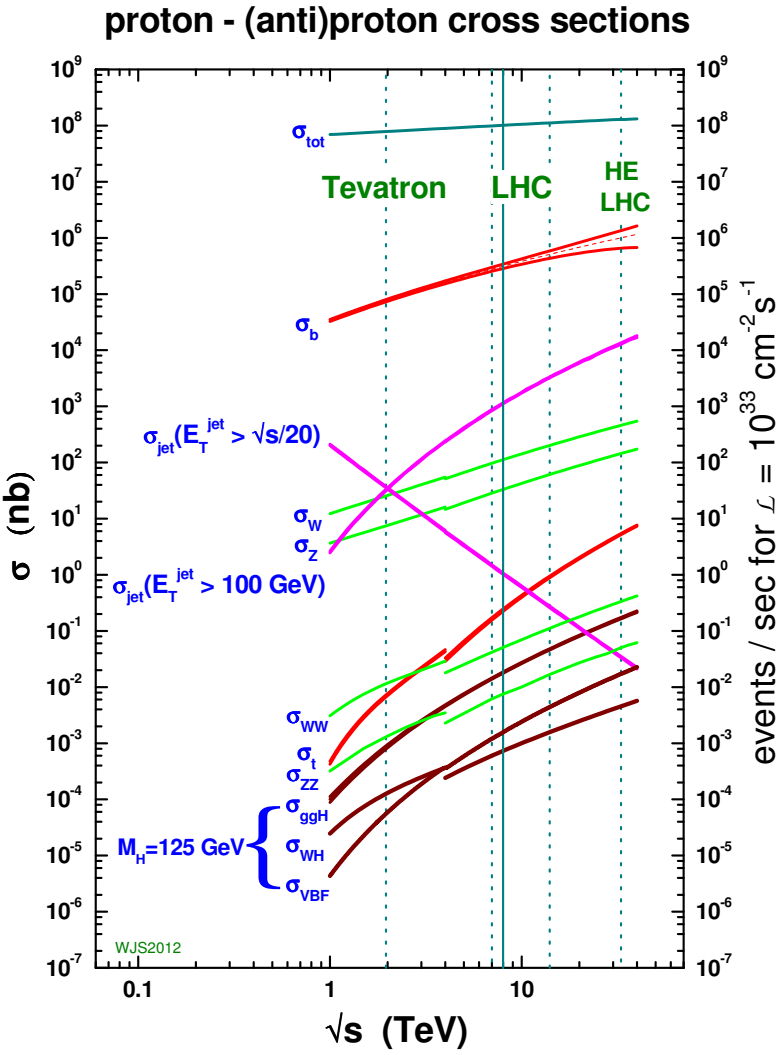

Figure 1.7: Production cross sections and corresponding event rates for different processes at LHC reported as a function of the center-of-mass energy $\sqrt{s}$ and in correspondence of a instantaneous luminosity $\mathscr{L}(t) = 10^{33} cm^{-2} s^{-1}$ [50].

1.4 A Slice of Phenomenology at the LHC

As described in the previous paragraph, the protons are not fundamental particles but they are constituted by three valence quarks of the first generation, uud. With the increase of the energy in the collision point a new contribution appears inside the proton and during the collision, the quantum vacuum characterized by a continuous creation and annihilation of quark-antiquark pairs and gluons which actively participate in the interaction. The most exhaustive description of the proton structure is given by the *Parton Distribution Functions* (PDF) [51–54], which describes the probability density to find a partons (quark or gluon) with a precise fraction of momentum of the proton inside the proton itself. The figure 1.8 shows that for low fractions of momentum (x) the biggest contribution inside the protons comes from the gluons, instead the valence quarks carries the biggest fraction of the proton momentum. The energies reached by LHC are such that allow to probe regions of low x.

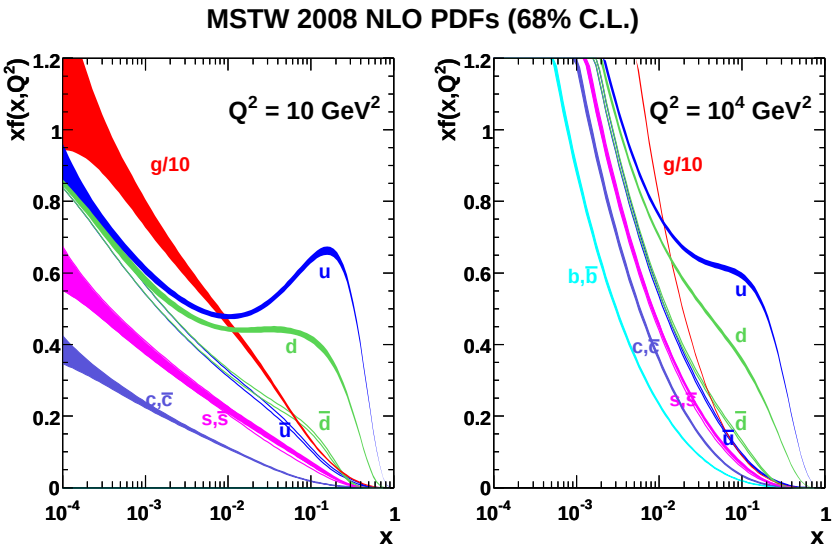

Figure 1.8: Parton distribution function as function of the fraction carried by partons (x) and the collision energy Q. The x-axis reports the fraction of momentum carried by the parton, whereas on the y-axis reports the mean value of the parton momentum [55].

The PDFs are a crucial ingredient to estimate the cross section of a proton-proton collision. Indeed, according to the QCD *factorization theorem* [56], the total cross section for a hadronic process $pp \to X$ assuming a scalar resonance can be written as:

$$\sigma_{pp\to X} = \sum_{a.b} \int\int dx_a dy_b f_a(x_a, \mu_F^2) f_b(x_b, \mu_F^2) \hat{\sigma}_{ab\to X}(x_a, x_b, \mu_F, \alpha_S(\mu_R)). \tag{1.36}$$

The probability to produce a particle X depends on the cross section of the subprocess $\hat{\sigma}_{ab\to X}$ mediated by the initial particles a and b weighted by the PDF functions, $f_a(x_a, \mu_F^2)$ and $f_b(x_b, \mu_F^2)$. The sum takes into account all the free degrees of freedom of each parton (the flavor, the color and the spin), while the integration is performed over the momentum fraction of the two partons. Finally, μ_R, the *renormalization scale*, represents the energy at which the strong coupling constant α_S is measured, instead the *factorization scale*, μ_F, is the energy scale at which the decoupling of short and long range physics happens.

In the particular case of a resonance, $\hat{\sigma}_{ab\to X}$ can be factorized as:

$$\hat{\sigma}_{ab\to X} = \hat{\sigma}_{ab\to Y} \frac{m_Y \Gamma^2}{(s - m_Y^2)^2 + m_Y^2 \Gamma(s)^2} \Gamma_{Y\to X} \tag{1.37}$$

where $\hat{\sigma}_{ab\to Y}$ identifies the mechanism to generate the resonance Y, instead $\hat{\sigma}_{Y\to X}$ represents its decay process.

Both $\hat{\sigma}_{ab\to Y}$ and $\hat{\sigma}_{Y\to X}$ are derived directly from the theoretical model under investigation. In the following lines the particular case of the SM will be described aware that this choice is not exhaustive of the entire phenomenological spectrum that could populate the physics at the LHC. This approach, moreover, lies outside the goals of this thesis.

At the same time, however, as described in the previous section, the vertex interactions between the neutral scalars and the other particles can be properly obtained rescaling the SM ones. Therefore, if from one side this difference creates a more rich phenomenology, on the other side the interaction building blocks between the particles are shared between the SM and the 2HDM. Finally, it has to be underlined that these conclusions can be drawn in first approximation also for other BSM models like the Electroweak Singlet Model [57, 58] and the Minimal Supersymmetric Standard Model [59].

Higgs Production Mechanism

According to the particular case of the SM, a scalar resonance can be produced through four different mechanisms: gluon fusion (ggF), vector boson fusion (VBF), associated production with a vector boson (VH), and finally an associated production with heavy quarks (ttH, bbH and tH). Their contributions to the total cross section production for the heavy scalar resonance is related to both the energy in the collision

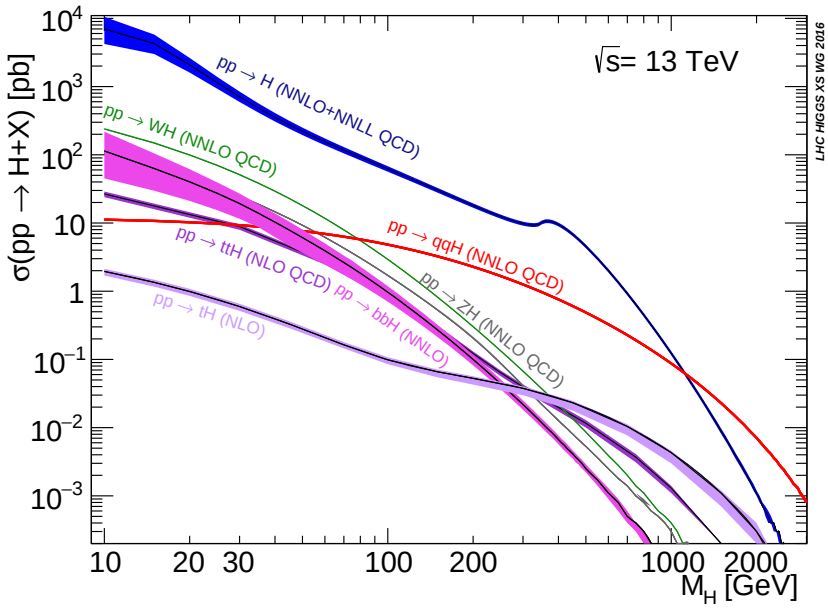

Figure 1.9: Theoretical cross sections of the main SM Higgs production mechanisms reported as a function of the Higgs boson mass. The computation has been performed considering p-p collisions at a center-of-mass energy $\sqrt{s}$ equal to 13 TeV [48].

point and the mass of the produced scalar. The figure 1.9 shows the cross section of each mechanism as a function of the mass of the heavy scalar boson at 13 TeV.

As illustrated in the picture, the dominant contributions are represented by ggF and VBF. In particular the gluon fusion is the dominant process up to the region where the mass of the heavy scalar boson SM-like reaches the value of 1 TeV whereas the VBF process dominates the region above 1 TeV. The others contributions ($VH, t\bar{t}H, b\bar{b}H$ and tH) are almost one or two order of magnitude below the dominant cross-section.

The dynamics of gluon fusion, whose its Feynman diagram is reported in figure 1.10, is controlled by the strong interaction. Its main contribution arises from the top loop thanks to the coupling with the scalar boson. In recent years a big effort was put to reduce the uncertainties related to the missing order of the perturbation theory. Recently, an effective approach was proposed to estimate the gluon fusion total cross-section at the $N^3LO + N^3LL$ precision in QCD expansion. With this procedure the uncertainties associated to the missing higher order is reduced to about 5% [60]. In the following analysis the gluon fusion is estimated using Powheg generator, [61–65], at NLO precision in QCD expansion and the events are subsequently reweigthed by NNLO scale factors computed using $hNNLO$ Monte Carlo code [66–68].

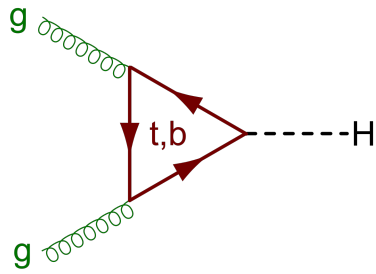

Figure 1.10: Feynman diagram for the gluon process at LO in α_S.

The production of the heavy boson with two additional jets quoted as vector-boson-fusion (Feynman diagrams available in figure 1.11) is the second mechanism that contributes to its production up to 1 TeV and the first in the above kinematic region. This process is characterized by two hard jets with a strong tendency to be forward-backward in contrast to the other mechanisms and present a high jet invariant mass. These two features offer a good background suppression. At the state of the art the best prediction of the VBF process are at the NNLO precision in the QCD expansion and its corresponding theoretical uncertainties is about 1% [60]. In the analysis this process is modeled by Powheg at the NLO precision in α_S without further corrections applied.

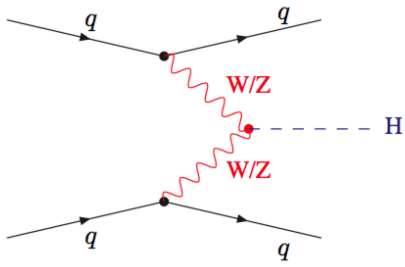

Figure 1.11: Feynman diagram for Vector Boson Fusion at LO in α_S.

The others three processes are not used in the final $ZZ \rightarrow 2l2\nu$ analysis due to their relative tiny cross-section. For this reason they will not be described further.

Higgs decays

The second part of the equation 1.37 describes the decay of the resonance Y into the X particle. In the case more then one decay channel is accessible, an important

physical quantity that described the probability of the resonance Y to decay either in the X particle or in the other channels is the *Branching ratio* (BR). The Figure 1.12 shows all the decay channels available for the Higgs boson. The Higgs particle can decay in both the fermions and the bosons of the standard model and in particular the decay is related to the coupling between them. Indeed, the couplings between the Higgs boson and the vector bosons depends quadratically by the vector boson mass, in the other case the coupling present a linear dependency to the mass of the fermions [69].

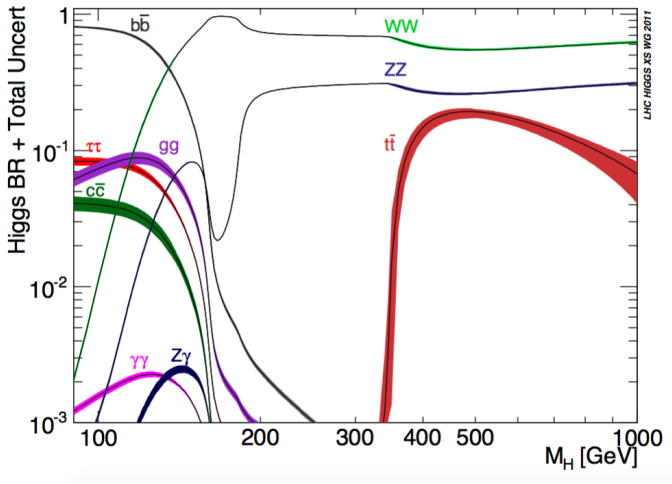

Figure 1.12: The different branching ratios for a heavy scalar boson SM-like as function of its mass from 90 GeV up to 1 TeV [70].

It can be observed that, depending on the Higgs mass boson, some decay channels are favored with respect to the others. This behavior can be explained taking into account that the increasing of the Higgs mass allows to access new decay mechanisms involving heavier SM particles. The fact that the total sum of the BR has to be equal to one at a fixed Higgs mass, constrains the branching ratios to compensate each others.

In the region below the WW threshold, the most dominant decays are those to $b\bar{b}$ and $\tau\tau$ pairs. Despite being experimentally clean, the $H \rightarrow \gamma\gamma$ decay is highly theoretically suppressed in this region (such a process can only happen through loops of a W boson). The decays to WW and ZZ pairs start to be dominant for M_H of 200 MeV, which is the region of interest of our analysis. In particular even if the WW decay channel is the most promising, the ZZ is, in the end, the most sensitive process as can be seen in the Fig. 1.13. The picture reports the combination results of the ZZ and WW channels using the data collected during Run-I using 7-8 GeV of energy for

the range between 200 GeV up to 1 TeV by CMS [71]. As reported in the plot the combined limits are guided by $ZZ \to 4l$ in the region below 500 GeV, whereas in the region above 500 GeV the channel $ZZ \to 2l2\nu$ is the most sensitive.

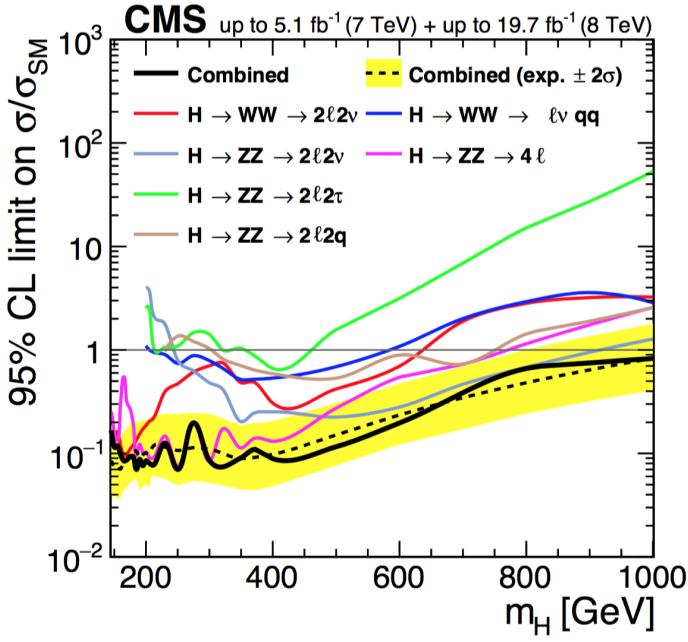

Figure 1.13: Upper limits at the 95% CL for each WW and ZZ decay channel. The observed and expected limits of the individual decay channels are compared with each other and with the combined results [71].

1.5 Interference Effects in $H \to 2l2\nu$ channel

One aspect that needs to be considered in the $H \to ZZ \to 2l2\nu$ analysis is the interference between the signal and the background. In this specific case, two different sources of interference have to be considered (Fig. 1.14): the one between $H \to ZZ \to 2l2\nu$, the SM-Higgs ($M_h = 125$ GeV), and the continuum background ($ZZ \to 2l2\nu$) and the interference between the $H \to ZZ \to 2l2\nu$ and the $H \to WW \to ll\nu\nu$ channels.

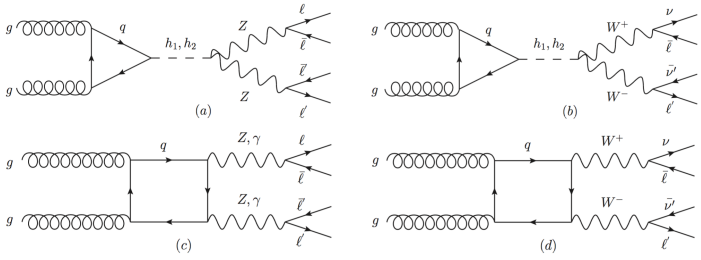

Figure 1.14: Representative Feynman graphs for the specific case of gluon fusion in ZZ and WW channel decay. The heavy Higgs (h_2) graphs define the signal process, which interferes with the light Higgs (h_1) graphs (*a*,*b*). They also interfere with the gluon-induced continuum background graphs (*c*,*d*). In the case where the two Zs decay in the $2l2\nu$ final state the interference between the ZZ and the WW decay chain can occur [72].

If the second contribution can be substantially reduced applying some kinematical cuts, the first interference can range from O(10%) to O(1) effects for integrated cross sections [60, 73]. Indeed, the heavy scalar signal is affected by both the interference with the continuum background, and a non-negligible interference with the off-shell tail of the light Higgs boson. In this case, the interference effects increase significantly with the increasing of heavy scalar mass and width. Moreover the interference spreads more and more around the peak as the width increases, while its effects are totally limited close to the region around the peak for lower values of the width. It has to be finally underlined that in the heavy scalar resonance region a strong cancellation occurs when both interference contributions (continuum and light Higgs) are added together because the heavy scalar boson-continuum background interference is negative above M_H and positive below M_H, while the heavy scalar-light Higgs interference has the opposite behavior [73] as documented by the Figure 1.15 for the case of an heavy scalar boson with mass of 450 GeV and width 46.8 GeV.

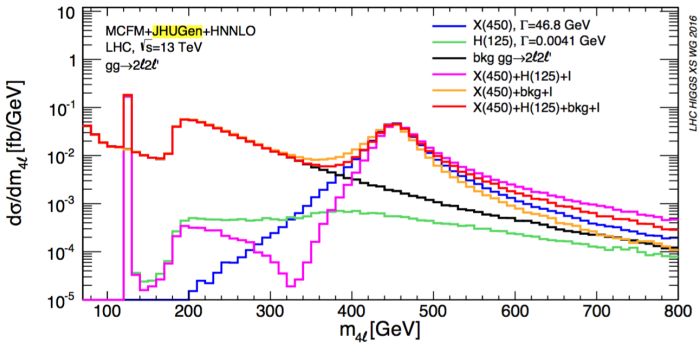

Figure 1.15: Differential cross section of the process $gg \to (X(450)/H(125) \to) ZZ/Z\gamma^*/\gamma^*\gamma^* \to 2l2l'$ (where l, $l' = e$, μ, or τ) as a function of the final four lepton invariant mass. The cross section is computed with MCFM [74, 75] and JHUGen generator [76–79], including the NNLO QCD K-factors. The different distributions show the impact on the invariant mass shape of the different processes in the case the heavy scalar boson, the SM higgs and the continuum background are considered isolated or combined in the total cross section computation [60].

1.6 Latest Results for BSM physics in di-boson channels

Experimental constraints on a scalar boson have been set in the context of searches performed exploiting the data collected by CMS during 2015 and 2016. In this section an overview of the main results in the di-boson channels will be presented. Such searches refer to ZZ $\rightarrow$ 2l 2ν [47] (ZZ2l2ν), ZZ $\rightarrow$ 4l [80] (ZZ4lep), and ZZ $\rightarrow$ 2l 2q [81] (ZZ2l2q) for the ZZ channel, WW $\rightarrow$ lν lν [82] (WWlνlν), X $\rightarrow \gamma\gamma$ [83] ($\gamma\gamma$), and finally X $\rightarrow Z\gamma$ [84] (Zγ).

In all of them any deviation from the expected background was measured.

Although each analysis has its own peculiarity, they can be classified in three different groups according to the signal modeling strategy adopted.

Most of them have used a model independent approach targeting a generic heavy resonance. This is the case of $ZZ \rightarrow 4l$, $ZZ \rightarrow 2l2q$, and finally $X \rightarrow Z\gamma$. The range in mass covers the region from 130 GeV up to 4 TeV, whereas the width hypothesis is analysis dependent. The ZZ4lep used a fixed width from 0 GeV up to 40 GeV, in the $Z\gamma$ analysis the width changes as a fraction of the associate mass point, and finally the ZZ2l2q analysis uses a narrow width approximation approach.

The $WW \rightarrow l\nu l\nu$ has targeted the ElectroWeak Singlet Model (EWS) as ZZ2l2ν. The first analysis has looked for a new resonance in the mass range between 200 GeV up to 1 TeV, testing only four width hypothesis 9%, 25%, 49%, and 100% the width of the heavy scalar boson SM-like. For $ZZ \rightarrow 2l2\nu$ the range in mass is slightly different because the upper value reaches 1.5 TeV, whereas for the width the analysis performed a scan from 1% up to 100% the width of the heavy scalar boson SM-like.

Finally, the $ZZ \rightarrow 2l2\nu$ has tested also the 2HDM model in both the Type-I and the Type-II scenario in the *decoupling region*. The limits are set as a function of the CP-even Higgs boson (*H*) mass and $\tan(\beta)$. Also $X \rightarrow \gamma\gamma$ targets the 2HDM but its limits are set as a function of the mass and the width. The results of the $ZZ \rightarrow 2l2\nu$ channel are reported in the Fig. 1.16.

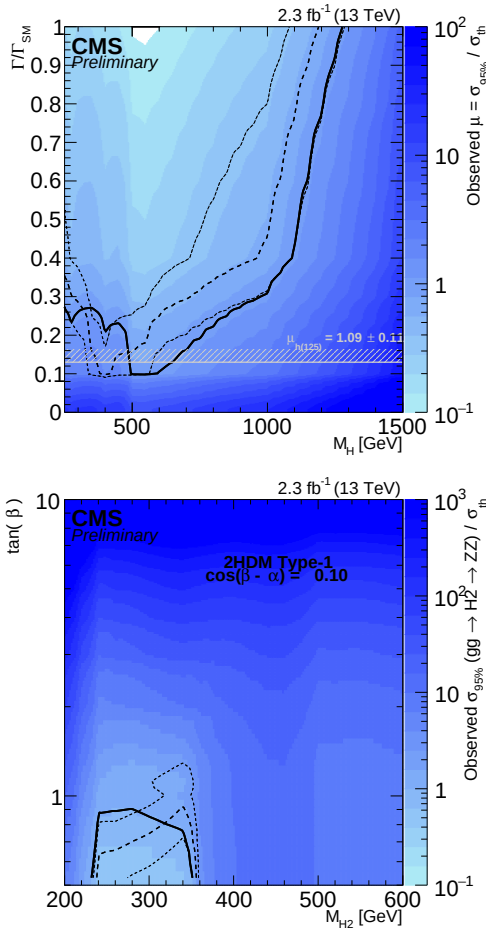

Figure 1.16: Limits set in $ZZ \rightarrow 2l2\nu$ in EWS (*top*) and 2HDM (Type-I) (*bottom*) [47].

Chapter 2

Experimental setup

To test the SM, one needs to probe the interactions among elementary particles up to the highest energy scales. One way to achieve this purpose is to let a particle accelerator to produce collisions between known particles. The particles produced in the collision produce electric signals in particle detectors. These signals allow the particles to be identified and measured thanks to dedicated software [85]. The next two chapters will be devoted entirely to the description of the experimental setup used in this thesis. Chapter 2 is focusing on the description of the two instruments involved in the experiment: the Large Hadron Collider (LHC) [3] and the Compact Muon Solenoid (CMS) [86] detector. Chapter 3 describes in more detail how the particles are reconstructed and identified starting from the informations recorded by CMS.

2.1 The Large Hadron Collider

The LHC is currently the largest and the most powerful proton-proton collider in the world. It was built by European Organization for Nuclear Research (CERN) between 1998 and 2008 and it lies in a tunnel of 26.7 km as deep as 175 m at the border between France and Switzerland near Geneva, previously occupied by the LEP collider.

The LHC is a synchrotron designed to operate at a centre-of-mass energy of $\sqrt{s}$ = 14 TeV with an *Instantaneous Luminosity* of $10^{34}\ cm^{-2}s^{-1}$.

The instantaneous luminosity is defined as [87]:

$$\mathscr{L}(t) = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi\epsilon_n \beta^*} \cdot F \tag{2.1}$$

and it quantifies the amount of particles that are brought into collision for unit of time and surface. The equation contains all the properties of one bunch of proton in one LHC beam (the beam is composed of separate bunches of protons): γ_r is the Lorentz factor of the colliding protons, f_{rev} is the revolution frequency and n_b is the number of bunches per beam; N_b and ϵ_n, are respectively the number of protons per bunch and the normalized transverse beam emittance[1]. The betatron function β^* quantifies the beam focusing by the magnetic optics at the interaction point[2].

In the design configuration of the LHC parameters, there are 2808 bunches per beam (n_b) each containing about $1.15 \cdot 10^{11}$ protons and covering 11245 turns of the LHC ring each second ($f_{rev}[Hz]$). The emittance ($\epsilon_n[mm \cdot mrad]$) is 3.75 mm · mrad, while the β^* = 0.55 m. All theses parameters combined together give an instantaneous luminosity of $\mathscr{L}(t) = 10^{34} cm^{-2}s^{-1}$ which means that in a certain amount of time, the average number of events N for a process defined by the cross-section σ is given by N = $\sigma \cdot \int \mathscr{L}(t)dt$. In the vicinity of each collision point both beams share the same

[1] The transverse emittance is the area in the phase space (x,x') or (y,y') containing the beam particles and it is given by the following expression:

$$\epsilon = \sqrt{\langle z^2 \rangle \langle z'^2 \rangle - \langle z \cdot z' \rangle^2} \tag{2.2}$$

where $z = x, y$ and z' is defined as $z' = dz/ds$ where s is the direction along the beam. The beam emittance decreases with increasing beam energy during acceleration and a convenient quantity for the operation of a hadron storage rings is the normalized emittance, ϵ_n, which stays constant along this period. It is defined as $\epsilon_n = \epsilon\gamma_r\beta_r$, where γ_r and β_r are the relativistic gamma and beta actors [88].

[2] The β function is determined by the accelerator magnet configuration (basically, the quadrupole magnet arrangement) and powering. When expressed in terms of the cross-sectional size of the bunch (σ) and the transverse emittance (ϵ), it becomes:

$$\beta = \pi\sigma^2/\epsilon \tag{2.3}$$

Beta gives the width of the beam squared divided by the emittance. In particular in the interaction point it is identified by β^*.

beam pipe of about 130 m in length. If the beams were collinear, this would result in about 30 parasitic collisions along the common beam pipe. In order to prevent this, the beams are crossed at an angle $\theta_c \sim 300$ μrad. This non-null crossing angle causes a reduction in the instantaneous luminosity that is described by the factor F in Eq. 2.1:

$$F = 1/\sqrt{(1 + \left(\frac{\theta_c \sigma_z}{2 r_{beam}}\right)^2)}, \tag{2.4}$$

where σ_z is the bunch length and $r_{beam} = \sqrt{(\epsilon_n \beta_* / \gamma_r)}$ is the beam radius at the interaction point.

These designed parameters are crucial to address all the questions which the accelerator was built for: the probe of the existence of the Higgs boson, the discovery of new resonances or the measurement of some deviations from the SM predictions require both high energy (order of TeV) and a huge amount of collisions. Figure 2.1 shows how the cross section of various processes varies with the center of mass energy of the proton-proton collisions.

The LHC ring consists of eight circular arcs and eight straight sections. The straight sections, long 0.5 Km, accomodate the collision points, the beam injection, the extraction facilities, the collimation system and the radio-frequency systems used to accelerate particles. Along the collider there are four different interaction points to which four different experiments are assigned: two general purpose detectors, ATLAS [89] and CMS [86], in addition to the LHCb [90] and ALICE [91] detectors, the first one is designed specifically to study the physics of the b-quark whereas the second studies heavy ions collisions.

To accelerate and bend the two proton beams, the collider uses respectively superconducting radio-frequency cavity system, and approximately 9000 different magnets of about 50 different types. The cavities are made of copper and niobium maintained at a temperature of 4.5 K. Each beam has eight cavities which operate at a frequency of 400 MHz. The core of the magnet system is represented by the dipoles and the quadrupoles. The LHC has 1232 dipole magnets, used to bend the trajectory of the circulating beams. All the dipoles have a cylindrical shape with a diameter of 1 m and a length of 16 m and weight about 35 t. They are made of superconducting niobium-titanium coils and generates a magnetic field of 8.3 T; the operational temperature is 1.9 K. In addition to the dipoles, the collider hosts 392 superconducting quadruple magnets used to focus the beams, the magnetic optics of the machine is further tuned using other multipole magnets. Other groups of magnets are used for beam injection, the beam extraction and instrument the straight sections with collision points, where the two beams are guided into a single beam pipe and squeezed to provide the desired luminosity. In particular, at the two general purpose detectors collision points, where

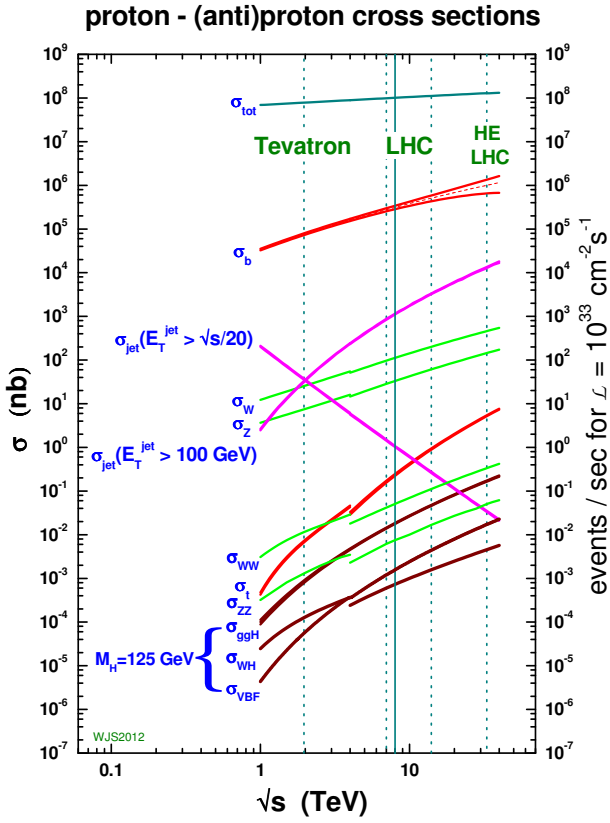

Figure 2.1: Production cross sections and corresponding event rates for different processes reported as a function of the center-of-mass energy $\sqrt{s}$ [50].

high luminosity is required, the beams are squeezed by quadrupoles such that their radius is decreased from about 1 mm to 17 μm

The LHC can't accelerate particles that are initially at rest. Because of this reason, protons are pre-accelerated through the complex accelerator chain hosted at CERN, before being injected in the LHC (Fig. 2.2). Protons, which are obtained by ionization of hydrogen, are first accelerated up to 50 MeV in a linear accelerator (LINAC2), and secondly are injected into the Proton Synchrotron Booster (PSB) to increase their energy to 1.4 GeV. The Proton Synchrotron (PS) is used to increase further the proton energy to 26 GeV and to form the LHC bunch structure: it captures six bunches from two consecutive cycles of the PSB and splits them in 72 bunches. At this point the LHC bunch spacing of 25 ns is formed. Finally, bunches in the PS are 120

cm long. The bunches from PS are then accumulated in the Super Proton Synchrotron (SPS), and accelerated to an energy of 450 GeV and shortened to 50 cm. Twelve cycles of the SPS are used to fill the LHC, in total the Large Hadron Collider contains 2808 bunches organized into 39 groups. Within the group the bunch spacing is 24.97 ns, but the groups are separated by larger gaps. After the LHC is filled, protons are accelerated to the target energy and the bunches are further shortened to 30 cm.

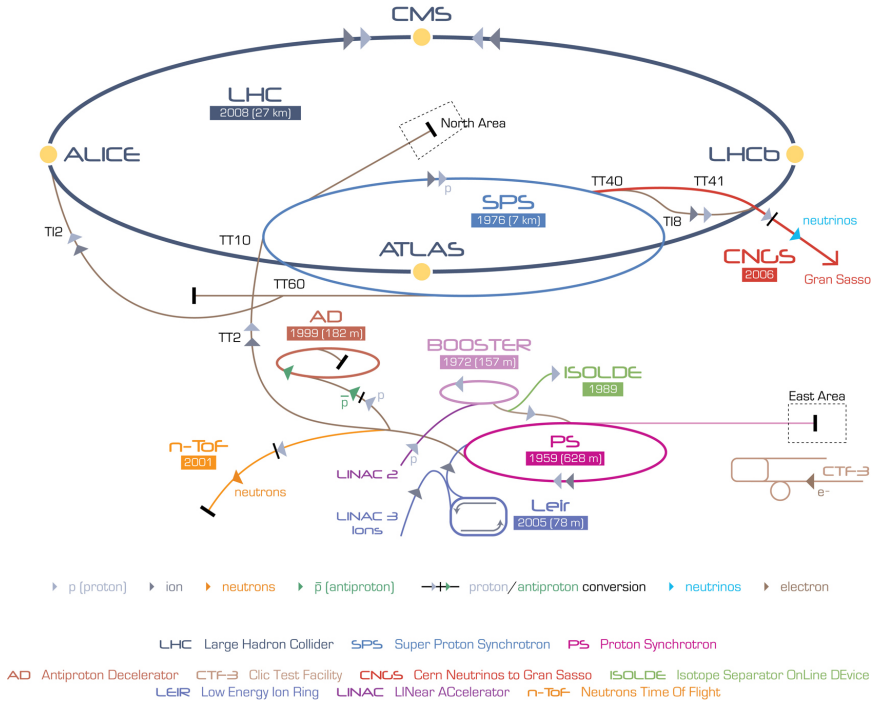

Figure 2.2: Schematic view of the CERN acelerator complex [92].

A direct consequence of the high luminosity is the high probability of having multiple interactions within the same bunch crossing. The average number can be found as

$$\mu = \frac{\sigma_{tot}}{\mathscr{L} n_b f_{rev}} \tag{2.5}$$

where σ_{tot} is defined as the total cross section of inelastic scattering in $p-p$ collisions. This effect gives rise to the so-called *in-time pileup*. The mean number of interactions per bunch crossing for the 2016 data taking was about 27 (Fig. 2.4). There is also the *out-of-time pileup* which is caused by particles that are produced in the previous or following bunch crossings and produce signals in the current bunch crossing

time window either because of detector timing inaccuracy or because they are still in the volume of the detector. The overall effect is a degradation of the performance of the event reconstruction.

Commissioning of the LHC with beams was started in 2008 but it was followed nine days later by big damage to 50 superconducting magnets, their mountings and the vacuum pipe caused by the magnet quenching. Recovery from the incident delayed the start of the machine by one year. In the 2010 the centre-of-mass energy was $\sqrt{s} = 7$ GeV and the integrated luminosity was 45 pb^{-1}, whereas 6.1 fb^{-1} and 23.3 fb^{-1} were accumulated in 2011 and in 2012, respectively. The first period of data taking (Run-I) ended in 2012, which was followed in 2013 by the first LHC long shutdown (LS1). In 2015 the LHC started the second period of data taking (Run-II) with an energy of $\sqrt{s} = 13$ TeV. The integrated luminosity delivered in 2015 was 4.2 fb^{-1} and 41.1 fb^{-1} in 2016. The results presented in this thesis are obtained analyzing the data collected by CMS during 2016.

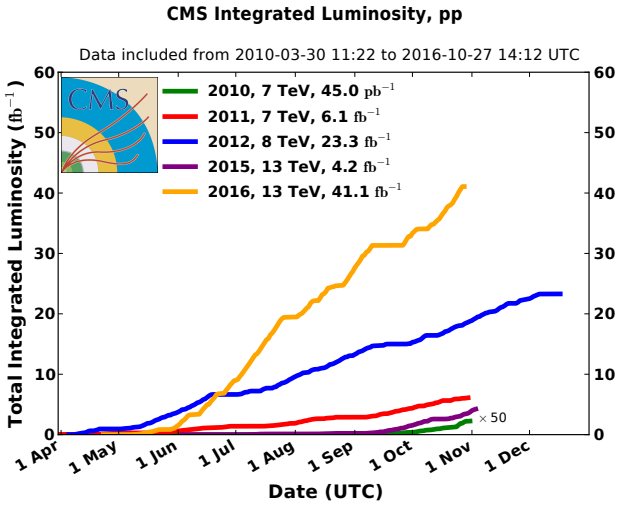

Figure 2.3: Comparison of the luminosity delivered by the LHC during the different years of data taking until 2016 [93].

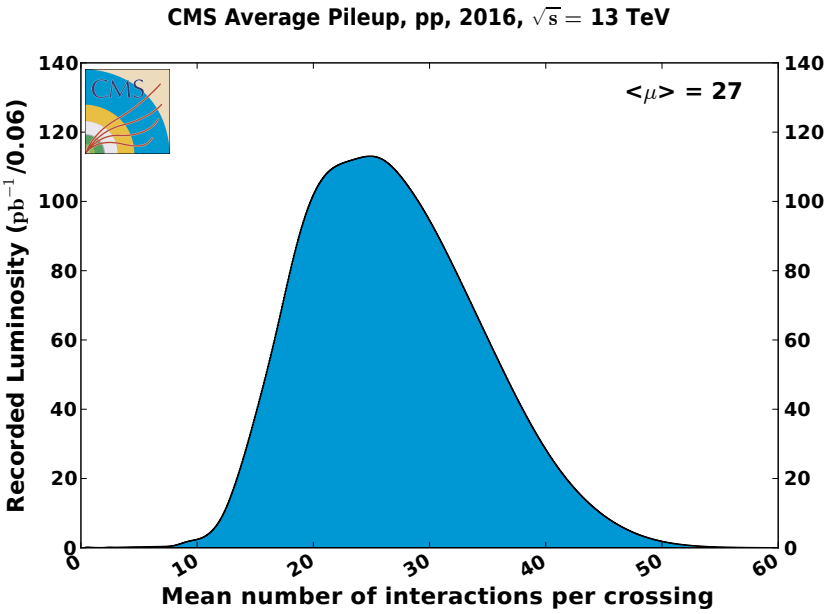

Figure 2.4: Mean number of interactions per bunch crossing for the 2016 *p-p* run at 13 TeV [93].

2.2 The CMS Detector

The primary design goal of the Compact Muon Solenoid experiment is to be able to discover the Standard Model Higgs boson. Additionally, the precise measurements of standard model processes (QCD, electroweak, flavor physics, etc.), the search for supersymmetric particles and new massive vector bosons, complete the challenging physics program the CMS detector has been built for. The experimental signatures from these processes are numerous and include involve high energy muons, electrons, taus, photons, neutrinos and jets. In order to cleanly detect these signatures the identification and precise energy measurement of these objects over a large energy range and at high luminosities is essential.

CMS is a general purpose proton-proton detector designed to run at the highest luminosity ($\mathscr{L}(t) > 10^{34} cm^{-2} s^{-1}$) at the LHC. The presence of about 27 (Fig. 2.4) inelastic events every 25 ns causes a severe increase of the detector occupancy and requires high-granularity subdetectors.

The CMS detector distinctive features include a high solenoidal magnetic field (4 T) coupled with a multilayer muon system, a homogeneous scintillating crystal electromagnetic calorimeter and an all-silicon inner tracking system based on fine-grained microstrip and pixel detectors. Additionally, the 25 ns bunch spacing constraints the trigger system, the time-response of each subdetector and the readout to cope with a collision rate of 40 MHz. The high radiation levels, caused by copious flux of particles, requires radiation-hard detectors and front-end electronics.

The experimental apparatus has a cylindrical shape with a length of 21.6 m and a diameter of 14.6 m and it can be roughly divided in a central part called *barrel* and in two caps called *endcaps*. The CMS experiment adopts two different coordinate systems. The cartesian system has its origin in the nominal collision point inside the detector: the x axis points radially inwards towards the centre of LHC ring, while the y axis aims vertically up, and finally the z axis points towards the Jura mountains in the direction of the counter-clockwise beam. The cylindrical coordinate system is also used. The radial distance is defined as $r = \sqrt{x^2 + y^2}$, and the ϕ angle is measured in the plane defined by the x and y axe. The polar angle θ is measured relative to the beam axis and is used to define the pseudorapidity, $\eta = \log(\tan\theta/2)$.

The detector has an onion-like structure: the sub-detector closest to the interaction point is the *tracker* which is used to measure the trajectory of the charged particles and it is wrapped by the *calorimeter* part aimed to measure the energy carried by the photons, electron and hadrons. The outermost detector is the *muon system* which permits to measure the trajectory and the momentum of the muons. Both the tracker and the calorimeter are accommodated inside the *magnet*.

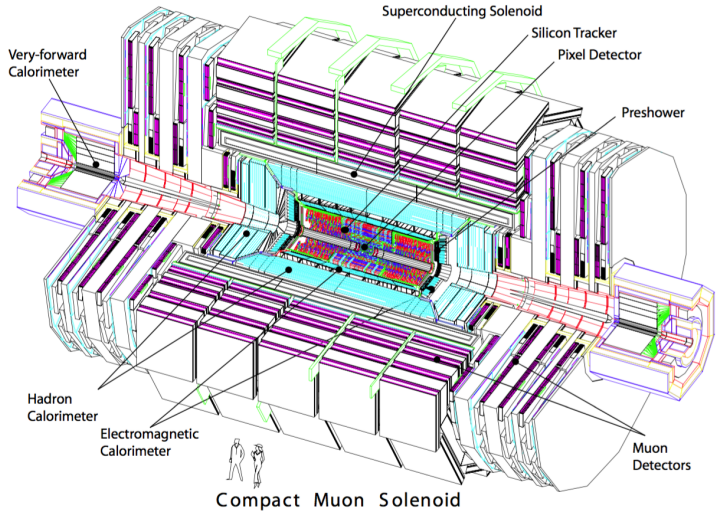

Figure 2.5: Tridimensional view of the CMS detector [86].

2.2.1 Magnet

In order to achieve a high momentum and position resolution in a very compact detector, the CMS Collaboration has adopted a superconducting coil to generate a 3.8 T solenoidal magnetic field over the entire tracking region. Such an intense magnetic field is achieved with a current of 18 kA and by cooling the magnet at the temperature of -268.5°C. The inner diameter of the coil is 6.22 m, and its length is 13.48 m. The tracking and calorimetry subsystems are completely enclosed within the field. The flux of the solenoid is returned by a set of iron disks in the endcaps and concentric twelve-sided cylinders in the barrel. Figure 2.6 shows the intensity of the magnetic field and its field lines on a longitudinal cross section of the CMS detector. In the central region of the detector, inside the solenoid, the magnetic field is uniform and mostly directed along the beam direction, on the contrary outside the solenoid and in the endcap the field is no longer uniform and no longer axial. Figure 2.7 shows the radial contribution of the Magnetic field considering one quadrant of the detector. Finally, the effects of the magnetic field propagates also outside the iron plates: at a radius of 8-9 m, the field si about 0.05 T, whereas at a distance of 35 m (roughly the location of the underground control room) the field is still 0.0005 T.

The magnetic flux density in the central region of the CMS detector which accommodates the tracker and electromagnetic calorimeter is known with a precision of

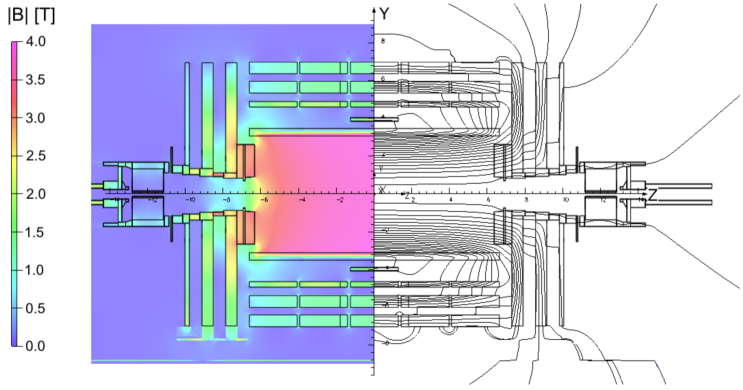

Figure 2.6: Value of Magnetic field (*left*) and field lines (*right*) predicted on a longitudinal section of the CMS detector, for the underground model at a central magnetic flux density of 3.8 T [94].

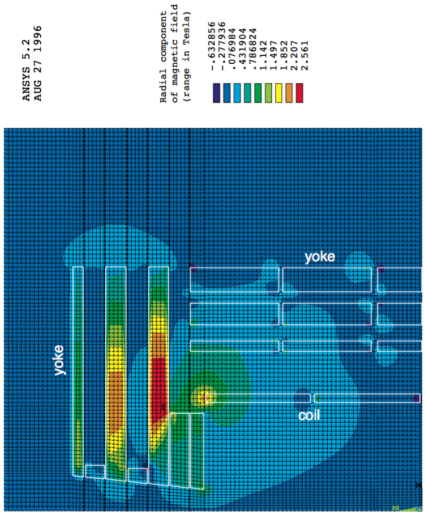

Figure 2.7: Radial component of the solenoidal magnetic field in the CMS detector [95].

less than of 0.1%, whereas it degrades to 2.5% in the iron yoke incorporated in the muon spectrometer [94].

2.2.2 Inner Tracking System

The tracker is the sub-detector closest to the interaction point. Its aim is the measurement of the trajectories of all the charged particles produced in the collisions. This information can be used further to measure the position of the primary and secondary vertices which is crucial to identify pile-up particles and the jets produced by a b quark.

To fit with the needs of high granularity, speed and radiation hardness, the tracker is built entirely using the silicon technology. In the region closer to the beam a Pixel sub-detector is used, while a Silicon Microstrip sub-detector is used in the rest of the tracker volume.

To ensure reliable charge assignment for muons and electrons up to the TeV scale, the design resolution on the transverse momentum of a track, $\delta p_T/p_T$, is $\approx 0.15 \cdot p_T$ (TeV) $\oplus$ 0.5% in the barrel. This relative resolution gradually degrades to $\approx 0.6 \cdot p_T$ (TeV) $\oplus$ 0.5% as $\mid \eta \mid$ approaches to 2.5. Figure 2.8 shows a schematic view of the inner tracker.

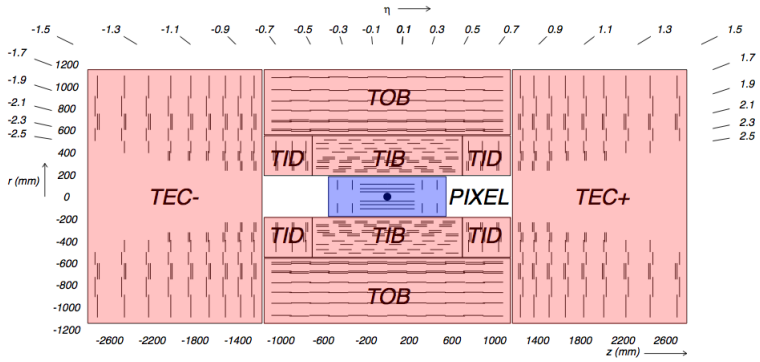

Figure 2.8: View of the inner tracking system. The violet part shows the pixel detector, instead the pink region the strip modules [96].

The pixel is composed by three concentric detection layers in the barrel and two detection disks in each endcap. Every layer is composed by matrices of 100 $\times$ 150 μm^2 of silicon pixels. The distance of the layers from the beam line in the barrel is r = 4.4, 7.3 and 10.2 cm, while in the endcap the disks are located at $|z| = 34.5$ cm and 46.5 cm. The spatial resolution is 10 μm for the r-ϕ measurement and 20 μm for the z measurement.

The silicon strip tracker is divided in two sub-structures characterized by a different granularity in the silicon strip arrangement. The innermost part is composed by the TIB in the barrel region and the TID in the endaps, whereas the outer part consists of the TOB and the TEC. The TIB is composed by four layers while the TID consists

of three disks of trapezoidal modules for each endap. The TIB covers the region up to 55 cm in r-direction and up to 65 cm in the z-direction. The silicon strip sensor used in both the TIB and the TID have a thickness of 320 μm. The pitch varies from 80 μm in the innermost barrel layer to 120 μm in the outermost disks. The TOB surrounds the TIB and the TID and it is composed by six layers of silicon strips modules of 500 μm of thickness and pitch of 122 μm for the innermost four layers and 183 μm for the remaining. The endcap is completed with the TEC sub-detector which is composed by nine different disks on each side of the tracker, covering the regions 124 cm $< |z| <$ 282 cm and 22.5 cm $< |r| <$ 113.5 cm. Each disk is divided into seven rings of trapezoidal modules having a thickness of 320 μm in the first four rings and 500 μm in the last two rings, the average pitch is between 97 μm and 184 μm according to the distance to the beam line.

The strip technology provides a poor (resolution of O(5 cm)) measurement of the particle crossing point in the direction along the strips. To mitigate this limitation, the first two TIB and TOB layers, the first two TID rings, and the first, the second and the fifth rings of the TEC, are characterized by the so-called "stereo" structure. These components present a second microstrip module mounted back-to-back and turned of an angle by 100 mrad. In the barrel, this configuration permits to have a resolution of about 23-34 μm in the r-ϕ plane and 230 μm in the z direction in the TIB, whereas for the TOB the resolution is 35-52 μm in the r-ϕ plane and 530 μm in the z direction. In the endcap any measurement of the tracker resolution was provided by the CMS experiment and for this reason it is not quoted in this thesis.

2.2.3 Electromagnetic Calorimeter

To measure the energy of electrons and photons the CMS detector takes advantage of the Electromagnetic Calorimeter (ECAL) placed immediately after the tracker detector. The overview of this subdetector is reported in Figure 2.9.

ECAL is an homogeneous calorimeter composed by crystal scintillator made of lead tungstate ($PbWO_4$) characterized by a very high containment of electromagnetic showers thanks to its radiation length of X_0 = 0.89 cm and its small Moliere radius, R_M = 2.2 cm. In the barrel ($|\eta|$ < 1.476) the crystals have a length of 23 cm, and a section of 22 $\times$ 22 mm^2, corresponding to a granularity of $\Delta\eta \times \Delta\phi$ = 0.0174 $\times$ 0.0174. The endap is composed by crystals of a bigger section, 30 $\times$ 30 mm^2. The crystals are arranged according a η - ϕ grid in barrel, while the disposition is organized along *x*-*y* in the endcaps.

A quality of the lead tungstate crystal is its fast response delivering about 80% of the light within 25 ns. The crystals are also radiation resistant. However the biggest part of the stored energy is thermally dissipated, then, in order to cope with the relatively low light yield (30 γ/MeV), photodetectors with intrinsic high gain, and able to operate in a high magnetic field, are used. Silicon avalanche photodiodes, APDs, are

used in the barrel, while the vacuum phototriodes, VPTs, are employed in the endcap. Moreover, considering that the sensitivity of the crystals and the APDs changes with the temperature, a cooling circuit is used to keep the operating temperature constantly at 18 ± 0.05 °C. The energy resolution of the detector for electrons with an energy of 20 GeV is about 1% and it improves to below 0.5% for energies above 100 GeV.

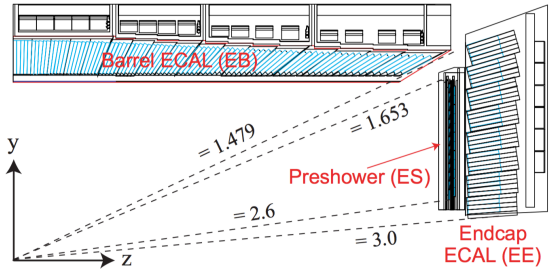

Figure 2.9: Longitudinal view of the ECAL of the CMS detector [86].

The Electromagnetic Calorimeter presents an additional sub-detectors placed in the forward region covering the pseudorapidity between 1.653 and 2.6, called ECAL Preshower (ES). This sampling substructure, composed by two layers of lead radiator, interleaved with silicon strip detector, permits a better discrimination between genuine prompt photons and the ones coming from the decay on flight of the neutral pions.

2.2.4 Hadronic Calorimeter

The last sub-detector enclosed by the solenoid, is the hadron calorimeter (HCAL) whose goal is to allow the measurement of the energy of the charged hadrons, the neutral hadrons, and the missing transverse energy (MET). The structure of the HCAL is more complex compared to the ECAL. Indeed it is composed by four different sub-detectors: HCAL Barrel (HB) in the barrel region, the HCAL EndCap (HE) in the endcaps, an outer hadronic calorimeter in the barrel around the solenoid to contains the tails of the showers (HO) and finally the very forward region is covered by the HCAL Forward (HF).

The HB is a sampling calorimeter built with brass as the passive material and a scintillator used as the active part. It covers the central region of the detector ($|\eta| <$

1.3) and it occupies a radial distance from 177 cm to 295 cm. It consists of 36 identical azimuthal wedges, which are constructed out of flat absorber plates parallel to the beam axis. The innermost and outermost absorber plates are made of stainless steel to provide an additional structural support and present a thickness of 40 mm and 75 mm respectively. The intermediate absorber plates, 14 in total, are made of brass and have a thickness of 50.5 mm or 56.5 mm. This translates into a total absorber thickness of 5.8 nuclear interaction lengths λ_I at η = 0, which increases to 10.6 λ_I at η = 1.3 due to the diagonal trajectory of the particle. The sensitive material is made of plastic scintillator organized in 17 sensitive cylindrical layers, divided into 32 segments along the z direction. The final segmentation in the η-ϕ plane is 0.087 $\times$ 0.087. The readout of the barrel is implemented using embedded wavelength-shifting fibers and hybrid photodiodes.

The endcap region ($1.3 \leq |\eta| \leq 3.0$) of the hadronic calorimeter consists of a sampling detector characterized by 17 layers of brass absorber with a thickness of 79 mm and 18 layers of plastic scintillator. The total nuclear interaction length is about 10 λ_I. The granularity is 0.087×0.087 in the η-ϕ if $|\eta| \leq 1.6$ and 0.17×0.17 when $1.653 < |\eta| \leq 3.0$.

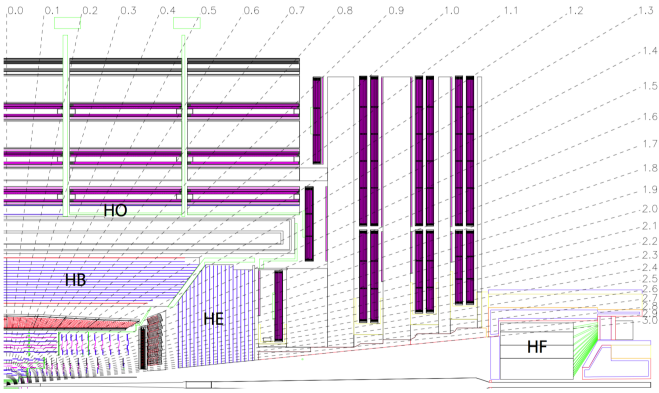

Figure 2.10: Longitudinal profile of CMS detector. In the picture the different components of the hadronic calorimeter are highlighted [86].

The HO is placed in the the central region ($|\eta| \leq 1.3$) between the solenoid and the innermost layer of the muon detector. The total nuclear radiation length obtained is 11.8 in the central region. Its granularity is also 0.087 $\times$ 0.087 in the η-ϕ plane.

The last sub-detector, belonging to the HCAL family, is the HF. It is located at a distance of 11.2 m from the center of the detector in the $|z|$ direction and it covers the very forward region between $3.0 \leq |\eta| \leq 5.2$; it presents a cylindrical shape with an external radius of 130 cm and internal one of 12.5 cm. The HF designed was

totally driven by its forward location. Indeed, the detector experiences a large flux of particles and thus must be able to operate under very harsh radiation conditions. The depth of the steel absorber is 1.65 m, for an equivalent nuclear interaction length of 10 λ_I, penetrated by quartz fiber parallel to the beam. The fibers form a rectangular grid in the (x, y) plane with a step size of 5 mm. They are bundled to provide 0.175×0.175 segmentation in the (η, ϕ) plane. By using a different disposition of half of the fibers (half of them start at 22 cm from the front face of the detector) the HF is able to distinguish between hadronic and electromagnetic showers. Indeed the hadronic shower presents a more uniform energy distribution whereas the electromagnetic ones are mostly contained within the first 22 cm.

The energy resolution is σ/E = 65%/$\sqrt{E}$ (GeV) $\oplus$ 5% in the barrel, σ/E = 83%/$\sqrt{E}$ (GeV) $\oplus$ 5% in the endcaps and σ/E = 100%/$\sqrt{E}$ (GeV) $\oplus$ 5% in the HF.

2.2.5 Muon System

The description of the following detector will be more accurate in order to give all the most interesting and useful informations to appreciate better the peculiarity of the *Dynamic Truncation Algorithm* (DYT) that will be presented in the Chapter 5.

The muon system, represented in the Fig. 2.11, is the last and outermost detector that composes CMS. It has to provide an independent muon tracking in the momentum range from 1 GeV up to a few TeV, and provide a robust and independent trigger system. To reach this goal the muon spectrometer is composed by three different sub-detectors, each built with a different technology. The barrel ($|\eta| < 1.2$), where the residual magnetic field (Fig. 2.6) is mostly contained in the iron yoke and the total hit rate is low (about 1 Hz/cm^2), is covered by *drift-tubes* (DT) and *resistive-plate-chambers* (RPC). *Cathode-strip-chambers* are used in the $0.9 < |\eta| < 2.4$ region characterized by a strong and non uniform magnetic field (up to 3.5 T) and a high total hit rate (about 1 kHz/cm^2). Each of these three types of Muon chambers is an independent, modular tracking detector capable of determining the position and direction of a passing muon from the measurements of the muon crossing points, which will be called "hits" in the following.

The resolution on the transverse momentum measurement is estimated to be 9-10% at 10 GeV and it degrades to 20-40% for muons with a transverse momentum of 1 TeV. When the tracker is also used in the p_T measurement, the resolution improves 1-1.5% at 10 GeV and 5-6% at 1 TeV in the barrel and the endcap respectively.

In the next three sub-sections a more detailed description of the three components is presented.

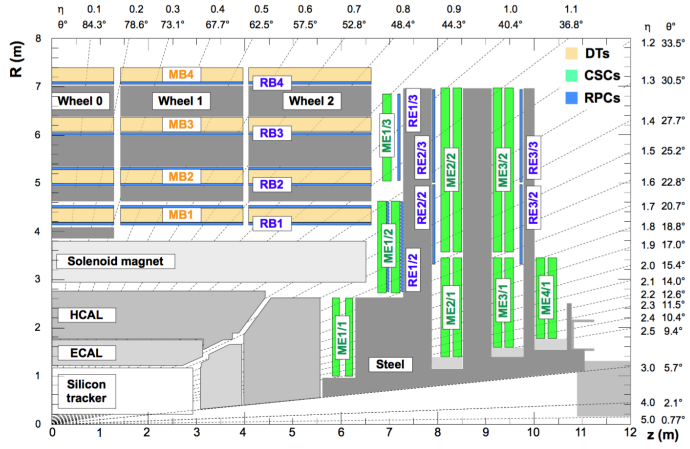

Figure 2.11: Muon Spectrometer profile. The dashed lines correspond to fixed η values. The DTs are in orange, the CSC in green, and finally the RPC in blue [97].

DT chambers

The DT sub-detector is composed by five concentric wheels, with three layers of iron are interleaving four stations of drift-tube chambers (MB1, MB2, MB3 and MB4). Each wheel is divided in 12 sectors covering 30° in the r-ϕ plane; the DT detector has a total of 250 chambers. Each DT chamber is composed of three super-layers (SLs), two of them presenting wires parallel to the beam direction in order to measure the position in the transverse plane of the detector, and a third, which is present only in the first three stations of each wheel, placed orthogonally to the other two SLs to measure the $|\eta|$ position. The first SL is separated from the other two, by a 128 mm thick aluminum honeycomb spacer. Each SL consist of 4 layers of DT tubes which have a rectangular section of 42 $\times$ 13 m^2 and a wire of length 2-3 m. The tube is filled by a mixture of Ar/CO_2 (85% / 15%) which provides a drift velocity of about 55 μm/s. The resolution of a single tube in the r-ϕ plane is about 200 μm in the first station and degrades to about 300 μm in the station four, whereas the resolution of position of the single local segment (track stubs reconstructed with a straight-line fit of the hits in the different measurement layers) fitted in r-ϕ layers is about 80 μm in the first station and about 100 μm in the last station. The poorer resolution of the r-ϕ layers in the last station compared to the others is because in this station, where a measurement of the longitudinal coordinate is missing, it is not possible to correct for the actual muon time-of-flight and signal propagation time along the wire.

The local reference frame of the DT chamber is defined by the z-axis pointing to the center of the CMS detector, the y-axis parallel to the beam line and finally the x-axis perpendicular to the y-z plane along the ϕ direction.

At the operating value of B = 3.8 T for the magnetic field inside the solenoid, typical values of the magnetic field inside the steel return yokes of the magnet structure in the barrel region range between 1.2 T and 1.8 T. The magnetic field lines are parallel to the beam in a direction opposite to the one in the region inside the solenoid.

CSC chambers

Due to the high B field in the volumes occupied by the detectors and the high rate of background in the forward region, the drift-tube chambers are not usable in this region of the detector, which is thus instrumented with *chatode-strip-chambers* (CSC).

The CSC detector is a multi-wire proportional chamber with cathode strips running orthogonally to the wires. Each endcap is composed by four stations (ME1, ME2, ME3 and ME4). Within a station the chambers are grouped into rings. With the exception of the first station, which has three rings, all the remaining stations have two rings. Each ring is composed by 18 or 36 trapezoidal chambers placed in the r-ϕ plane and overlapping in the ϕ direction. Each chamber consists of 6 gas gaps 9.5 mm thick and each of them composed by a plane of radial cathode strips and a plane of anode wires running perpendicularly to the central strips. The gas is a mixture of Ar/CO_2/Cf_4, in proportion of 40% / 50% / 10%. The wires have a diameter of 50 μm and are spaced by 3.2 mm, and they are read out in groups of 5 to 16 in order to reduce the number of channels, finally the strips have a width of 6.7-16.0 mm. The wires permit to measure the r coordinate and the strips the ϕ direction. The precisely measured coordinate in the CSCs is the one measured by the strips, since it is in this direction that muons are deflected by the CMS magnetic field. The spatial resolution of the CSC strip measurement depends on the relative position at which a muon crosses a strip: it is better for a muon crossing near a strip edge than at the center because then more of the induced charge is shared between that strip and its neighbor, allowing a better estimate of the center of the charge distribution. To benefit from this fact, alternate layers in a CSC are staggered by half a strip width. The resolution in the direction orthogonal to the strips is of about 200 μm. This value translates into a resolution of 10 mrad in the ϕ coordinate.

Each station is divided in two rings with the exception of the first station which has three rings. In particular the first ring of the first station, ME1/1 is located inside CMS solenoid in the gap between the hadron endcap calorimeter and the return iron joke and it operates in the strong coaxial magnetic field of the solenoid. The ME1/1 unites 36 identical trapezoidal chambers covers an angle in $\phi = 10.8^\circ$ to allow an overlap between two neighboring chamber and avoiding the appearance of dead zones

in the ME1/1 station. To reduce the rate in the channels of the cathode electronic circuits, the ME1/1 is divided in two parts: the first one is ME1/1B characterized by a strip length of 1065 mm and covering the pseudorapidity region $1.6 < |\eta| < 2.1$, the second part is ME1/1A with a strip length of 440 mm and covering the pseudorapidity region $2.1 < |\eta| < 2.4$. The magnetic field of the CMS degrades the spatial resolution of the chamber because the electrons are deflected by the Lorentz force while they drift toward the anode wire. This effect is compensated by rotating the wires by 29° with respect to the the chamber axis. Each anode wire readout channel corresponds to a group of 11-12 wires. The resolution in the ϕ angle is about 2.5 mrad.

The CSC local reference frame is defined in the following way: the local z-axis is parallel to the beam, whereas the y-axis is parallel to the r-axis direction and finally the x-axis is perpendicular to the y-z plane.

RPC chambers

The main goal of the RPC system is to provide a trigger response system complementary to the one provided by DT and CSC. The detector are installed both in the barrel and in the endcap, and are characterized by a fast response and a high accuracy of the time measurement. The RPC in the barrel mimic the DT installation. There are six layers of RPC detectors in every wheel: two layers are installed in the first and second station and one in each of MB3 and MB4. In the barrel the strips run parallel to the beam with a pitch that increases with the layer, 2.1 cm in the first one and 4.1 cm in the last one. In the endcap the RPC system covers the region up to $|\eta| \leq 1.6$ using chambers with trapezoidal shape, as for the CSC the strips run radially with a trapezoidal shape such that the angle covered by each strip is constant. The dimensions of the strips change with the layer as in the barrel: in ME1 they are 25 cm long and 0.7 cm thick, whereas in the ME2, ME3 and ME4 stations the strips have a length of 80 cm and a pitch of about 3 cm. Each chamber has two gas gaps, each made out of a two planes of bakelite coated by graphite for the voltage distribution. Each gas gap has a thickness of 2 mm and it is filled with a mixture of gases composed of 95% of Freon ($C_2H_2F_4$) and 5% of Isobutane (iC_2H_{10}). The RPCs are operated in *avalanche mode*, which allows rates as high as a few kHz to be withstood with high efficiency.

Magnetic field in Muon Detector

The Muon spectrometer is the sub-detector affected the most by the magnetic field and its flux return region with the iron plates. Large forces on the endcap disks appear as a result of these magnetic fields. The overall magnetic force on the first endcap disk is roughly 7000 metric tons for an object that weighs about 900 metric tons, so the magnetic forces dwarf the gravitational forces even for such heavy disks

meaning that the iron structure in which the chambers are mounted is not fixed and continue to change.

The full field is present in the region in which the innermost endcap CSCs, the ME1/1 chambers, must operate. However, the field at this position is uniform and almost entirely axial. At the next endcap station going out radially, called ME1/2, the field has fallen off to a considerable degree, but it is no longer uniform and no longer axial as shown in the left plot of Fig. 2.12, where the radial component is the same or even bigger then the longitudinal one. The same problem affects all the chambers in the second station of the CSC detector.

A radial component is present also in the first station of the DT sub-detector, indeed at the end of the coil and in the iron gaps there are large stray fields in the chamber area as reported in the right plot of Fig. 2.12. Near the end of chambers in the first station second wheel (ME2/1) the radial component reaches 0.8 T. Fortunately, these regions are small with respect to the entire area covered by the DTs

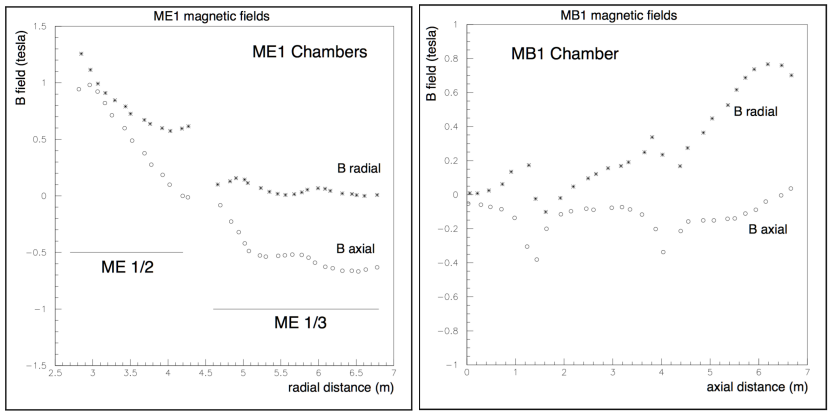

Figure 2.12: Axial and radial components of the magnetic field in the ME1 (*left*) and in the MB1(*right*) stations . [95].

2.2.6 Trigger

One of the most challenging tasks for the CMS detector is the online selection of the interesting events. Three constraints make this task highly demanding. The bunch crossing rate of the LHC is 40 MHz, but only about 1kHz on average can be saved on permanent storage. The rate of the interesting events with production of on-shell heavy bosons, top quarks or new physics, is very small compared to the total rate of the inelastic p-p scattering, which is about 10^9 Hz at the LHC nominal luminosity. The bunch crossing separation of 25 ns constrains the time available to analyze online

all raw data collected by the detector. CMS performs the online event selection with a double level trigger: the *Level 1* (L1) trigger is built using custom hardware, and reduces the event rate down to 100 kHz. The *High-Level-Trigger* uses software run on a processor farm and exploits all the available detector information resulting in a more accurate selection of the events passing the L1 triggers. After the HLT selection, the event rate is reduced to about 1kHz on average, which is what is permanently stored on disk for further offline analysis.

Level 1 Trigger (L1)

The L1 trigger has the aim to apply the first and rough skim to the events produced in the LHC collision. Part of the L1 trigger electronics is on the detectors, while the rest operates in the so-called service cavern located next to the underground detector cavern and it is built from custom and programmable electronics. The time needed to decide whether an event could be of interest or has to be discarded, and transfer back this decision to the sub-detectors, which keep the high resolution data in pipeline memories in the meantime is 3.3 μs.

The L1 uses only the information coming from the calorimeter, *Global Calorimeter Trigger*, (GCT) and from the muon spectrometer, *Global Muon Trigger*, GMT. The GCT and GMT identify *trigger objects* of different flavor ($e/\mu/\tau/\gamma$ and jets) and assign to them the measured direction and energy or the transverse momentum. The final decision is taken by the *Global Trigger*, which can run in parallel up to 128 trigger algorithms based on the event trigger objects.

In the following the description will focus on the muon L1 trigger.

All three sub-detectors of the muon spectrometer are involved in the trigger. Both in the DT and in the CSC chambers local segments are reconstructed starting from hits recorded in individual detector layers. In each DT, a segment is constructed in the (r,ϕ) and (r,z) planes in order to obtain a full three dimensional trajectory. In particular in the (r,ϕ) plane the so-called DT track finder (DTTF) combines all the segments found in all the stations to identify trajectories compatible with muons produced close to the interaction point. The same procedure is done in the CSC subdetector where the segments exploited by the CSC track finder (CSCTF) come from the independent reconstruction done within an individual chamber using wire and strip hits. The RPC trigger system doesn't reconstruct any local segment but it searches for pattern of hits matching predefined ones corresponding to the expected trajectories of muons produced close to the interaction vertex.

Each sub-component gives to the global muon trigger (GMT) up to four candidates ranked by the transverse momentum; the GMT combines the candidates based on the spatial information and the duplicates are removed. The GMT receives information about energy deposits in each calorimeter from the GCT. This information is

used to apply isolation criteria to the muon candidates. If the event satisfies at least one trigger algorithm the event is passed to the HLT.

High Level Trigger (HLT)

The High Level Trigger has about 100 ms average time to analyze each event. More time allows to use more sophisticated algorithms which can exploit the entire information coming from the detectors including the tracker, which represents almost 80% of the event size.

The HLT is composed by about 400 triggers paths. Every path is designed as a combination of intermixed reconstruction and filter modules. Once some part of the event is reconstructed, a filter module decides if it satisfies or not some thresholds and in the first case, the event is passed to the next reconstruction step. To save time the steps are ordered by CPU time. If an event is not accepted by a path, it can be accepted by another one.

Some triggers paths are characterized by very loose thresholds mostly for control and calibration purposes. In order to keep their accept rate within acceptable limits. These trigger paths are pre-scaled.

Chapter 3

Particle Reconstruction and Simulation

The previous chapter was entirely devoted to the metering equipment adopted by the CMS collaboration, the next one will be used to describe the offline software data reconstruction steps. Event reconstruction starts from the "hits" recorded by the detectors of the CMS experiment. Hits carry the information on the particle crossing point and time, and in some cases also on the energy deposited by the particle in the detector. The reconstruction process transform the hits information into physical objects like electrons, photons, muons or jets. The first part of the chapter will focus on the reconstruction of the two main physical quantity needed to identify a particle: its momentum and its energy. Instead, the second part will be used to describe how the particles are reconstructed focusing on the main algorithm used by the collaboration, the Particle-Flow Algorithm (PF).

3.1 Detector Basic Reconstruction

The tracker, thanks to the bending of the trajectory of charged particles due to the magnetic field, allows the measurement of the momentum and the charge of each particle passing through it. The calorimeters sub-detectors permit the determination of the energy released by photons, electrons and hadrons. Finally, muon tracks are identified and reconstructed by the muon spectrometer.

The next section is devoted to describe how the CMS collaboration manages both the reconstruction of the track and the energy deposit in the calorimeters starting from the hits building block.

3.1.1 Track Reconstruction in the tracker

Track reconstruction refers to the process of estimating of the momentum and the position parameters of each charged particle that is produced in the collision [98].

Reconstructing the trajectories of charged particles is a computationally challenging task that is tackled with an iterative approach (*iterative tracking*). The tracking algorithm adopted by CMS is commonly referred to as the Combinatorial Track Finder (CTF) [98], which is an adaptation of the Kalman filter [99] that simultaneously performs the determination of the set of detector position measurements believed to originate from the same particle (*track finding*) and the estimate of the set of parameters describing the kinematical state of the particle at some point in space (*track fitting*). The basic idea of the iterative tracking is that the initial iteration searches for tracks with a high transverse momentum and produced at the primary collision point. These tracks are easier to identify. After each iteration all the hits associated to a new track are removed from the hit collection, in this way the combinatorial complexity is reduced and more complex track topologies can be reconstructed.

Each iteration is characterized by four different steps: seed generation, track finding, track fitting and track selection.

Seed Generation

A charged particle inside a uniform magnetic field traces, in the absence of interactions with matter, a helicoidal trajectory. The state of the track on a given two-dimensional surface can be defined by five different parameters. The Seed generation aims to determine some estimate of the track parameters and their associated uncertainties using the information of three 3-D hits or two 3-D hits constrained by some requests on the origin of the trajectory itself. In order to reduce the number of seeds and obtain the highest possible accuracy in the track parameters seeds are built using the inner part of the tracker made of pixels.

Track Finding and Track Fitting

The Track Finding step aims to collect the hits associated with each seed. The track fitting step performs the estimate of the particle trajectory starting from the seed and the set of hits associated to it by the track finding step. The track fitting results in the determination of the track parameters and associated uncertainties at all detecting surfaces intercepted by the reconstructed particles trajectory. These two steps are clearly distinguished, but in the CMS track reconstruction algorithm, called Combinatorial Kalman Filter (CKF), they use the same basic tool: the Kalman Filter (KF) [99]. In the CKF the seed track parameters are propagated a first time to the following detecting layers that could possibly carry measurements produced by the particle. This first propagation assumes a uniform magnetic field and does not take into account neither Coulomb scattering nor energy loss (*analytical propagator*). Once candidate track hits are identified, a more accurate propagation that takes into account energy loss and Coulomb scattering is performed in order to estimate the track parameters on the detection surface of each candidate hit. These track parameters are combined with the hit information and the procedure is iterated until the outermost detection layer is reached. The procedure is then reproduced by moving inward in order to use full track information for the estimate of the track parameters on each layer. In case more than one compatible hit is found on a detection layer, new track candidates are created and processed independently. If no hits are found a conventional "ghost" hit, with no position information, is attached to the track candidate.

The track of a single particle can be reconstructed more than once, either starting from a different seed or when one seed develops in more than one possible candidate. To remedy to this problem a cleaning procedure is applied: the track candidate with the lowest number of hits or worst χ^2 is discarded for each couple of tracks that share an amount of hits greater then a precise threshold.

The track finding produces a collection of track candidates. In order to obtain the most accurate estimate of the track parameters on all detecting layers, the KF procedure outlined above is repeated on each track candidate in both the direction of the momentum, outwards and inward. This refitting step is based on the Runge-Kutta propagator [100] taking into account both the material effects and the in-homogenity of the magnetic field.

Track Selection

The last part of the track reconstruction consists of a cleaning procedure of all the track candidates with the aim to reduce the contributions of the fake tracks that are generated during the previous steps. The cleaning is performed applying requirements like a minimum number of layers with hits, a minimum number of 3-D hits, a

maximum value for the χ^2 and a compatibility with the beam spot[1]. The amount of fake tracks decreases exponentially with the number of layers with hits.

Using a $t\bar{t}$ sample simulated at $\sqrt{s}$ = 7 GeV and assuming 2011 LHC pileup conditions, the tracking efficiency for prompt particles[2] with p_T greater than 0.9 GeV, is estimated to be about 94% in the Barrel region and 85% in the EndCap (Fig. 3.1). The deterioration of the performance in the forward region is caused by hadron-nuclear interaction in the tracker material that are not properly handled by the tracking algorithm. The fake rate of the tracking algorithm is of the order of few percents for particles with a transverse momentum between 1 GeV and 10 GeV. In the lower and upper region of the momentum range the fake rate tends to increase. At low momentum values the multiple scattering increases the window to be used when searching for additional hits associated to the track candidates increasing the probability of wrong assignment of hits to the track. Instead, in the high p_T region there is a higher possibility to assign incorrectly the hits to the track in case collimated secondary particles are produced following a nuclear interaction between the particle and the tracker material.

The relative resolution on the transverse momentum, and the absolute resolution on the two components of the impact parameter for the $t\bar{t}$ sample[3] is reported in figure 3.2. At high momentum ($p_T \geq 100$ GeV), the resolution of both the components of the impact parameter are affected by the position of the innermost hit in the pixel detector, whereas the lower momentum region is affected by multiple scattering. In the high momentum region, the material in the tracker accounts for between 20% and 30% of the transverse momentum resolution, whereas at lower momenta, the resolution is dominated by multiple scattering [98]. The impact of the pileup on the track resolution is negligible.

3.1.2 Energy Reconstruction in the calorimeters

The reconstruction of the energy in the calorimeters represent the second essential step to reconstruct and identify charged particles in an efficient way, and the only one for neutral particles like photons and neutral hadrons. The Electromagnetic Calorimeter is dedicated to the electrons and photons reconstruction, and to unambiguously detect and measure the energy and the direction of stable neutral and charged hadrons (given that hadronic showers can start in the ECAL volume). The Hadron

[1] The beam spot represents a 3-D profile of the luminous region, where the LHC beams collide in the CMS detector. For more details about its determination and the relative performance refer to [98].

[2] A particle is defined as prompt if it is produced during a physics process in the collision point. A particle is defined as non-prompt if it is produced at a significantly displaced vertex ($>\sim$100 μm).

[3] The transverse impact parameter (d_0) is defined as the distance of closest approach to the primary vertex. The longitudinal component (d_z) is the z coordinate of the point on the track that determines d_0.

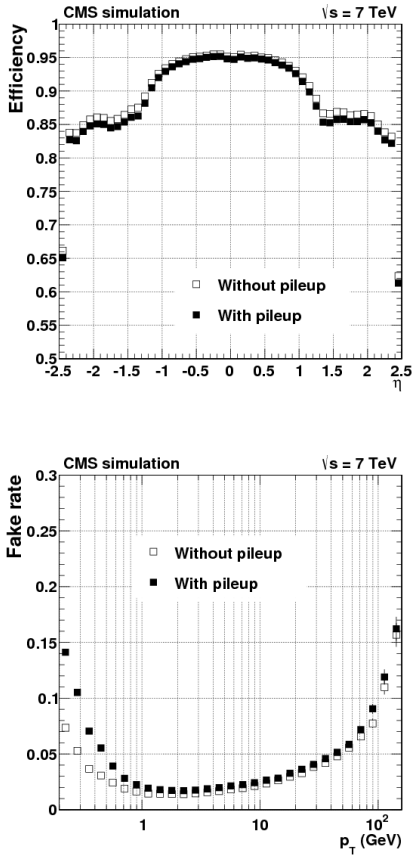

Figure 3.1: Tracking efficiency (*left*) and fake rate (*right*) for $t\bar{t}$ events simulated with and without superimposed pileup collisions. The number of pileup interactions superimposed on each simulated event is generated assuming 2011 LHC conditions. The efficiency results are for charged particles produced less than 3 cm in r direction and 30 cm in z direction from the centre of the beam spot and with $p_T > 0.9$ GeV [98].

Calorimeter, in conjunction with the ECAL, is used to measure the remaining hadron energy deposits.

In the ECAL, the energy is determined thanks to a clustering algorithms, which is different in barrel and in endcap. The goal of these algorithms is to collect all the energy deposits of a given electromagnetic shower and recover also the radiated energy. After the identification of the seed, defined as a cell with a local maximum of energy deposit, the different clustering algorithms identify the energy clusters. The group of energy deposits must be compatible with an electromagnetic shower caused by an

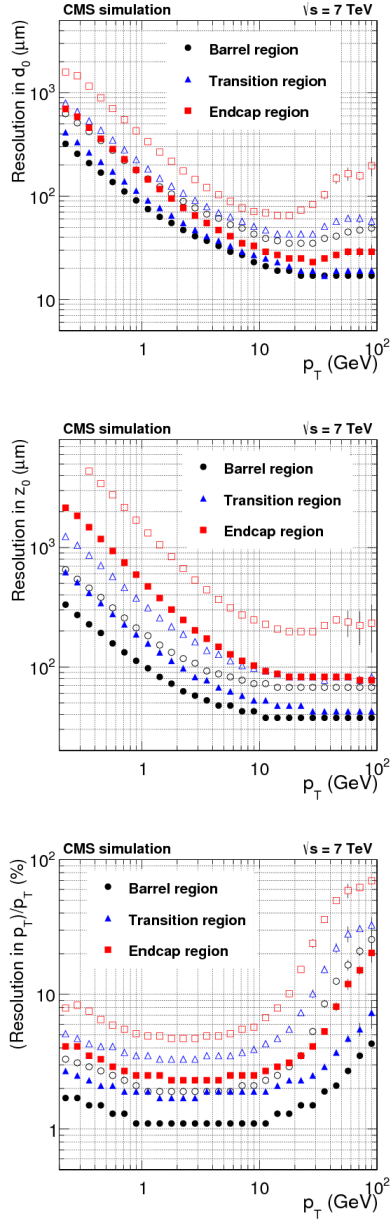

Figure 3.2: Resolution as a function of p_T and η for tracks produced in $t\bar{t}$ sample assuming 2011 LHC pileup condition. The following variables are reported: transverse (*top right*) and longitudinal (*top left*) impact parameters, and p_T (*bottom*). For each distribution, the solid (open) symbols correspond to the half-width of the 68% (90%) intervals centered on the mode of the distribution in residuals [98].

electron or a photon. Subsequently, nearby clusters are collected into a supercluster (SC), in order to recover the energy irradiated by bremsstrahlung by electrons. More details about the different clustering algorithms can be found in [85]. Following the same strategy used in the electromagnetic calorimeter, in the HCAL the energy deposit above a certain threshold, is group in order to form a calorimeter tower consisting of one or more HCAL cells geometrically corresponding to ECAL crystals. In the barrel region of the calorimeters, the unweighted sum of one single HCAL cell and 5×5 ECAL crystals form a projective calorimeter tower. The association between HCAL cells and ECAL crystals is more complex in the endcap regions. A four-momentum is associated with each tower deposit above a certain threshold, assuming zero mass, and taking the direction of the tower position as seen from the interaction point.

3.1.3 Reconstruction of Primary Vertices

The goal of primary-vertex reconstruction is to measure the position, and the associated uncertainty, of all proton-proton interaction vertices in each event, including the signal vertex and any vertices from pileup collisions, using the available reconstructed tracks. The reconstruction procedure consists of three steps. Initially the tracks are selected according to their consistency of being produced promptly in the primary interaction region. The consistency is measured imposing some requirements on the transverse impact parameter, the number of strip and pixel hits associated with a track and finally a cut on the normalized χ^2 of the fitted trajectory. Subsequently, the tracks that appear to be produced from the same interaction vertex are clustered on the basis of their z-coordinates at their point of closest approach to the centre of the beam spot. This clustering allows for the reconstruction of primary proton-proton interactions in the same LHC bunch crossing and it is performed by the Deterministic Annealing (DA) algorithm [101]. Finally those vertex candidates containing at least two tracks are then fitted to compute the best estimate of vertex parameters, including its x, y and z position and the covariance matrix.

Results of the primary-vertex resolution in x and z positions are reported in Fig. 3.3 for the minimum-bias sample[4] and the jet-enriched data samples obtained by requiring each event to have a reconstructed jet with transverse energy greater then 20 GeV. The resolution in y coordinate is identical to x and it is not reported. For the minimum-bias events, the resolution of a vertex built out of 50 tracks in the x direction

[4]The minimum bias events ware selected by a trigger signal in any of the beam scintillator counters (BSC) [102], coincident with a signal from either of the two Beam Pick-up Timing for Experiments (BPTX) [103] detectors indicating the presence of at least one proton bunch crossing the interaction point. The events are then selected offline by requiring BPTX signals from both beams passing the interaction point, at least one reconstructed charged particle trajectory in the pixel detector originating from within 0.2 cm of the beam position in the transverse direction, a coincidence of at least one HF calorimeter tower with more than 3 GeV total energy on each of the positive and negative sides of the HF and finally requiring at least one reconstructed primary vertex.

is less than 20 μm, and less than 25 μm in z direction. Considering the same number of tracks for the jet data samples the situation is slightly better, indeed the resolutions is less than 10 μm and 12 μm in x and z, but it worsens to 100 μm and 150 μm in the case there are about 5 tracks. The efficiency of the primary-vertex reconstruction is estimated to be close to 100% when more than two tracks are used to reconstruct the vertex whereas it decreases to 98% when only two tracks are used [98].

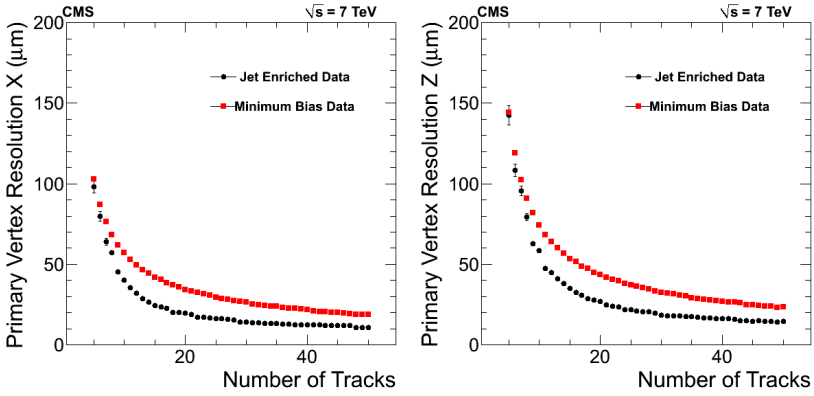

Figure 3.3: Primary Vertex resolution in x (*left*) and z (*right*) coordinates as a function of the number of primary tracks at the fitted vertex. The comparison is performed on data considering the minimum-bias sample and the jet data sample. The jets selected form the second sample have $E_T > 20$ GeV [98].

3.1.4 Track Reconstruction in the muon spectrometer

The first part of the muon reconstruction occurs at an individual chamber level, and in this sense it is labeled as local reconstruction. It is characterized by two steps: in the first part the hit position in each chamber are reconstructed and then this information is arranged in a more complex object called segment.

In the DT chambers the hit position is obtained combining the information of the electron drift velocity and the lapse of time needed by the electrons to reach the wire, assuming initially the same amount of time for each muon to reach the same layer of the DT chamber. However this assumption is not totally true due to both an intrinsic time-of-flight spread related to the muon momentum, and the possibility that the hit could have been produced by a muon from another collision. In order to take these effects into account each segment reconstruction in the DTs is performed as a fit assuming three free parameters: the muon crossing time, the muon crossing position, and finally the muon direction. The segments are reconstructed separately in each r-ϕ and r-z superlayer (SL), then the r-ϕ segments are refitted together in order to create

a single segment in the r-ϕ plane. Finally the segments reconstructed in the SLs are combined together to obtain a global segment in the chamber defined by the local x, and y position and the ϕ, and θ directions. The final segment has an angular resolution of 0.7 mrad in the ϕ direction and about 6 mrad in the θ direction.

Local hit reconstruction in the CSC detector measures the position of the traversing muon by combining information from the cathode strips and anode wires independently in each of the six layers of a chamber. Each CSC layer thus provides a measurement in two dimensions. The wires measure the radial coordinate r and the strips the azimuthal ϕ-coordinate. The charge distribution of a cluster of three neighbouring strips is fitted to the so-called Gatti function [104] to obtain a precise position measurement. Using the individual hits, segments are built as straight-line tracks inside the chamber. Their reconstruction starts by considering pairs of hits from the first and the last layers that are roughly compatible with a straight line passing through the beam spot. Compatible hits from the intermediate layers are added to the segment candidate, and a linear fit is performed. If the track segment is accepted, its hits are masked out and the procedure is repeated until all the hits in the chamber are used. The position resolution depends on the station: it ranges from 50 μm in the first one to about 250 μm in the fourth one; the direction resolution varies with the chamber, with an average of about 40-50 mrad in the ϕ direction and slightly worse in the θ direction.

The barrel and the endcap RPC systems are mainly used as trigger detectors. However they also contribute to the muon reconstruction by providing additional position and time information both in the barrel and in the endcap region. In particular in the RPCs the result of local reconstruction are points in the plane of the detector. The local hit position is obtained from the centroid of the clustered strips. Indeed, since the charges from a muon can be shared by more than one strip, adjacent strips are clustered together. In the barrel, where strips are rectangular, the centroid is simply the center of a rectangle. In the endcap, the computation is more complicated as the area covered by the clusters present a trapezoidal shape. The assumption here is that each group of strips is fired as a result of a single particle crossing and that the crossing can have taken place anywhere with flat probability over the area covered by the strips of the cluster [105].

After the local reconstruction in each muon sub-detector, the track reconstruction in the muon spectrometer continues using the Kalman filter technique. Reconstruction starts from the seed made up of groups of DT or CSC segments. At this point the seed is propagated up to the next compatible detecting layer and the track parameters are updated using the best measurement found on the layer. At this stage the Kalman filter is applied in the inside-out direction. Reconstructed hits from RPC chambers are also included. The propagation of the track parameters take into account both the material effects and the ionization energy losses in the return yoke. The best measurement is

searched looking at the compatibility based on a χ^2 method between the extrapolated track state and the measurement on the detection layer. As soon as the outermost compatible layer of the muon detector is reached the track propagation is stopped. The final filter is applied in the outside-in direction down to the innermost compatible layer of the muon detector. The final trajectory to be accepted has be composed by at least two measurements, one of which must be of the DT or CSC type. After the fake track suppression, the parameters are extrapolated to the point of closest approach to the beam line. In order to improve the momentum resolution a constraint represented by the passage of the muon through the nominal interaction point is imposed. The muons reconstructed using this approach are called *Standalone muon tracks*.

3.2 Particle-Flow Algorithm

The Particle-Flow Algorithm (PF) [106], aims to reconstruct and identify all the stable particles by combining all the informations obtained from all the CMS sub-detectors. The output is a list of particles with their energy, point of production and direction like the ones coming from simulations. The particles are then further used to identify jets, and reconstruct the missing transverse momentum. To fulfill this goal, the algorithm needs that some basic building blocks like the tracks reconstructed in the tracker and in the muon systems as well as the energy deposits in the calorimeter system are available. A key ingredient in this approach is the fine spatial granularity of the detectors layers. Thanks to the tracker immersed in a uniform 3.8 T and the iterative tracker algorithm, charged track particles can be reconstructed with large efficiency and small fake rate, while the granularity and tightness of the calorimeters permits the identification of photons and charged hadrons. Electrons are reconstructed combining the informations of the tracker and the calorimeter, instead the muons are obtained combining the two spectrometers. All these identified and measured particles are then clustered into jets. Finally the imbalance in the momentum on the transverse plane, the missing transverse momentum $\vec{E}_T^{miss}$, permits to identify the presence of invisible particle like neutrinos in the event.

3.2.1 PF building blocks

The building units, *particle-flow elements*, that are fed to PF are the tracks reconstructed neither in tracker or in the muon spectrometer and the clusters of energy identified in the calorimeters.

In general a particle is expected to produce more then one particle-flow element in the CMS detector. It is then mandatory to find a strategy to try to connect all the elements to each others avoiding any double counting. The connection is obtained by defining either a distance in the (η, ϕ) plane or a χ^2 restricted to the nearest neighbours in the (η, ϕ) plane. The output of this step of the algorithm are *PF blocks* of elements linked together. In each PF block, the identification and reconstruction sequence proceeds in a very precise order. The first particle to be reconstructed are muons and the corresponding PF elements (tracks and clusters) are removed from the list of PF elements. The electron identification and reconstruction follows with energetic and isolated photons. As for the muon case, the corresponding tracks and ECAL clusters are excluded from further consideration. The remaining elements in the block are then subject to a cross-identification of charged hadrons, neutral hadrons, and photons, arising from to jets [106].

3.2.2 Muon Reconstruction

At the LHC a lot of muon sources exist: light hadrons, heavy hadrons, charged and neutral boson. When combined all together these sources are able to produce muons that cover a very wide spectrum in transverse momentum from a handful of GeV up to hundreds of GeV. The low region of the spectrum ($p_T < 30$ GeV) is dominated by muons coming from light hadrons like pions and kaons, the charge and neutral bosons dominate the region above 40 GeV, whereas the heavy hadrons give a wide range muon p_T spectrum (Fig. 3.4) [107].

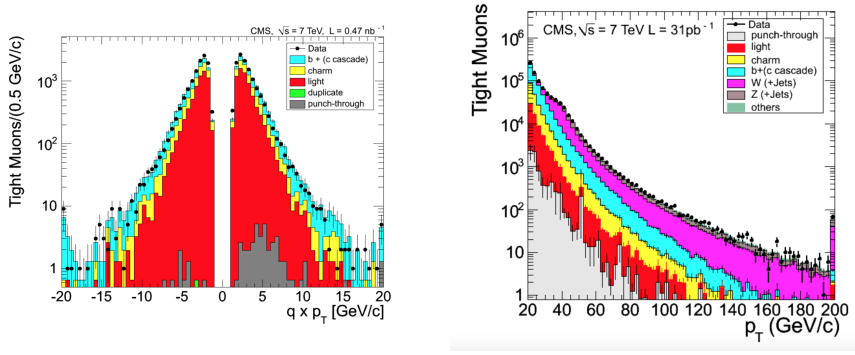

Figure 3.4: The distributions report the comparison between the simulations and the data collected by CMS in 2010 at the energy of 7 TeV. The result on the left shows the transverse momentum multiplied by its charge for global muons fired by the zero bias trigger. The p_T distribution is peaked at 0.5 GeV due to the minimum threshold required to reach the EndCap which is 0.5 GeV while in the Barrel is about 3-4 GeV. The result on the right reports the transverse momentum of the global muons with a p_T greater then 20 GeV. The events are selected by the single muon trigger applying a minimum p_T threshold of 15 GeV. In both the cases the muon tight requirements are applied [107].

The minimum transverse muon momentum value needed to reach the first muon detecting surface depends on the eta region. It is about 0.5 GeV in the EndCap region and about 4 GeV in the Barrel [107]. Given this situation, in order to improve the efficiency of the muon reconstruction over the entire spectrum, CMS adopts two different strategies to reconstruct muon tracks: an inside-out approach starting from the tracker track pointing to the muon spectrometer and an outside-in procedure starting from the latest muon layer down to the tracker detector. Tracks are first reconstructed independently in the inner tracker (*Tracker tracks*) and in the muon system (*Standalone muon tracks*). Based on these, two reconstruction approaches are used:

- *Global muon reconstruction*. The Global muon reconstruction targets all the muons with sufficient energy to cross more then one station in the muon spectrometer. All the muons reconstructed by the Global muon reconstruction are built "outside-in" propagating the Standalone muon track down to the Tracker. After the extrapolation down to the vertex of each Standalone muon track, a region of interest (ROI) in the η-ϕ plane is identified around the extrapolation of the muon track. The determination of the region of interest is based on the Standalone muon track parameters and their corresponding uncertainties of the extrapolated trajectory down to the Tracker, obtained with the assumption that the muon originates from the interaction point. Successively a matching onto a common surface between the muon track and all the tracker tracks inside the ROI is performed. The common surface is chosen in order to minimize the covariant error matrix of the propagated track parameters and reduce the number of matches per Standalone muon. It is possible to attempt to propagate the tracker track and Standalone muon track to any common plane, but it is natural to choose a plane that does not require the track to be propagated through a lot of a material. Some of the natural choices for a common surface are the tracker system outer boundary, the muon system inner boundary, the detector surface of the outermost tracker track hit, and the detector surface of the innermost muon track hit. Once the tracks are propagated to the common surface, a comparison of the track parameters is made using the position and the momentum of the tracks. Given the 5 parameters space $\vec{p_i} = \{q/|P|, x, y, p_x, p_y\}$, the discriminating variables are: the χ^2 of the $\vec{p_i}$ vector associated to the tracker tracks and the muon tracks falling in the same region of interest, the distance ($d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$), the χ^2 of the position vector ($\vec{d_i} = \{x, y\}$), and the ΔR of the involved tracks in the local frame of the ROI, and finally the ΔR of the $\vec{p_i}$ defined at the primary vertex. The tracks that satisfy precise requirements on the discriminating variables are selected and the global muon track is fitted combining hits from the tracker track and standalone-muon track, using the Kalman filter technique (KF).

- *Tracker muon reconstruction*. Tracker muons are built "inside-out" by propagating Tracker tracks to the muon system and looking for geometrical matches with DT or CSC segments. This reconstruction approach is designed to have high efficiency in the low momentum region ($p \leq 5$ GeV), where the muons have not enough energy to cross completely the muon spectrometer. In order to recover the largest possible fraction of these muons, the tracker tracks with $p_T > 0.5$ GeV and total momentum $p > 2.5$ GeV are considered as possible muon candidates and successively they are extrapolated up the muon spectrometer using the KF. The matching between the tracker track and the segment is performed in the local frame of the muon chamber requiring that either the dis-

tance between them in the x direction is less than 3 cm or the pull for x[5] is lower than 4 [107].

In the PF algorithm, the muon identification develops from the global and the tracker muon as described above, requiring further selection cuts on the isolation[6]. This selection is performed to improve the efficiency to identify muons in jets and keep the fake rate from misidentified charged hadrons low. Considering the chance that a charged energetic hadron reaches the muon spectrometer, a muon is defined as PF muon if it has a valid hit in the muon system and if the sum of the transverse energy of all its neighboring tracks and calorimeter cells inside a cone of radius 0.3 around the muon direction is less then 10% of its transverse momentum.

The combination of the double informations related to the tracker and the muon spectrometer allows to improve the resolution on the momentum as reported in the Fig. 3.5. In particular, the resolution is about 1% in the barrel and 3% in the endcaps for muons with momentum up to 100 GeV.

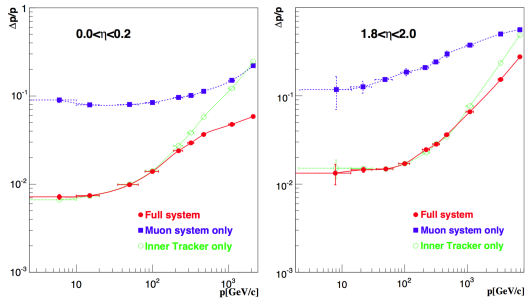

Figure 3.5: The resolution of the muon momentum as a function of its value using the muon spectrometer only, the tracker only, and finally their combination. The results are reported for two different pseudorapidity regions: the barrel ($|\eta| < 0.2$) on the left and the endcaps ($1.8 < |\eta| < 2.0$) on the right.

3.2.3 Electron and Isolated Photons Reconstruction

The PF electrons are reconstructed combining two informations: the tracker track and the energy deposits in the calorimeters [106, 108].

In principle, electron tracks can be reconstructed in the tracker using the Kalman Filter algorithm. This procedure can be sometimes compromised in the case electrons

[5]The pull is defined as the difference between the position of the matched segment and the position of the extrapolated track, divided by their combined uncertainties.

[6]The isolation quantifies the magnitude of the total transverse momentum of the particle in a cone with radius 0.3 or 0.5 around its direction. Much more details will be given in the following parts of this chapter.

experience large radiative losses in the tracker, leading in general to a reduced hit-collection efficiency and a poor estimation of its track parameters. For these reasons, a customized tracking procedure is designed for them. As for a generic track, also for the electrons the first step of the track reconstruction is the seeding which is then followed by the track building and the track fitting.

Two complementary algorithm for the seed estimation are used. The first procedure is called *ECAL-based seeding* and it starts from the SC energy and position to estimate the electron trajectory in the first layer of the tracker. This approach relying on the fact that the energy-weighted average position of the clusters is on the helix corresponding to the initial electron energy propagated through the magnetic field without emission of radiation. The electron seed is selected from all the reconstructed seeds. The SC are selected in order to limit the number of misidentified seeds requiring that both $E^{SC}_{T,ECAL} < 4$ GeV and $E^{ECAL}_{SC}/E^{HCAL}_{Tower} < 15\%$ the energy of the SC within a cone of radius 0.15 around the electron direction. This procedure works efficiently for isolated electrons with transverse energy greater then 4 GeV. For electrons in jets, however, the energy and position of the associated superclusters are often biased by the overlapping contributions from other particle deposits, leading to large inefficiency during the linking procedure. In addition, also the backward propagation from the SC to the inner tracker is likely to be compatible with many hits from other charged particles causing a substantial misreconstruction rate. Finally, for electrons with small transverse momentum, their tracks are bent by the magnetic field and its radiated energy is spread over a region larger than the limits imposed to the supercluster size. As a consequence, in these cases the position of the SC and the assignment of the tracker hits are biased.

For all these reasons, a *Tracker-based seed* approach is developed. To recover the seeding efficiency for low p_T or non isolated electrons, tracker seeds are formed by combining hits with the vertices obtained from tracks reconstructed in the pixel. For all those tracks that fulfills the matching energy (E)-momentum (p) criterion of $r_{th} <$ E/p < 3 (r_{th} is set to 0.65 for electrons with 2 < p_T < 6 GeV, and to 0.75 for electrons with p_T > 6 GeV), the algorithm starts with tracks reconstructed with the KF algorithm. In this case, the algorithm collects hits up to the ECAL detector and the track is matched to the closest PF cluster. For those tracks that have experienced a significant bremsstrahlung radiation and fail the matching energy-momentum condition a second selection is attempted. Initially the KF algorithm either stops collecting hits, or keeps collecting them but relaxing their χ^2 with the electron trajectory. Successively, the new tracker-seeded candidate is refitted using a Gaussian sum filter (GSF) [109].

The final electron seed is then selected combining ECAL-seed and Tracker-seed in a multivariate (MVA) analysis [110].

In the region where the transverse momentum is greater then 4 GeV, the introduction of the Tracker-based seeding almost doubles the electron efficiency obtained using only the ECAL-based approach. The situation is much better for lower values of

p_T. Indeed, the inefficiency of the ECAL-based seed is improved making the electron reconstruction possible for lower transverse momentum [106] (Fig. 3.6).

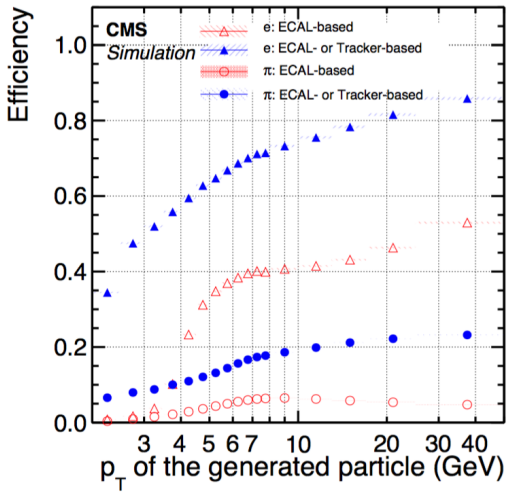

Figure 3.6: Electron seeding efficiency for electrons (triangles) and pions (circles) as a function of their transverse momentum. The results are obtained using a simulated events samples enriched in b quark jets with p_T between 80 GeV and 170 GeV, and with at least one semileptonic b hadron decay. Both the ECAL- based seeding efficiency and the Tracker-based seeding efficiency are displayed. [106].

The track-seeding is then followed by the track-building and the track-fitting. The track-building is similar to the standard tracking workflow except that the χ^2 threshold, used by the KF to decide the compatibility between an hit and the trajectory, is weakened from 30 to 2000 and the energy loss are modeled in order to take into account high radiative energy loss. The track fitting is performed using the Gaussian sum filter [108].

The last ingredient needed by the PF algorithm to identify the electron is its energy deposit in the ECAL detector. The clustering of the energy is performed separately in each subdetector in the barrel and in the endcaps. The cluster seeds are identified as cells with an energy larger then a given threshold and the energy of the closest cells. The topological-cluster is then grown from the seeds by aggregating cells with at least a corner in common with a cell already in the cluster and with an energy in excess of the cell threshold set to twice the noise level. The PF cluster corresponding to the electron at the ECAL surface is the one matched to the track at the exit of the tracker. The total ECAL energy is then corrected in order to recover all the energy dis-

sipated by photons during the bremsstrahlung emission. Following the electron track path, tangents are extrapolated towards the ECAL. If the extrapolation falls within the cluster, then the cluster is added to the electron PF cluster as bremsstrahlung photon.

The final electron candidates are constructed from the association of an electron track and a cluster in the ECAL. For ECAL-seeded a geometrical matching between the track and the SC in the η-ϕ plane is required, whereas for tracker-seeded electrons a global identification variable is defined using an MVA approach that combines informations on track observables, the electron PF cluster observables, and the association between the two. The electron momentum resolution estimated using electrons coming from the Z boson decay and with p_T about 45 GeV ranges from about 2% in the barrel region to about 4% in the endcaps for electrons with momentum close to 100 GeV. The region which benefits the most from the introduction of the tracking reconstruction step is the low momentum region. In the Barrel region the improvement is more then 50% for p_T < 10 GeV, whereas in the EndCap is about 10% [108] (Fig. 3.7).

The isolated photons are reconstructed using only the calorimeter information. Photon candidates are retained if they are isolated from other tracks and calorimeter clusters in the event, and if the ECAL SC energy and the ratio between the HCAL and ECAL energies are compatible with those expected from a photon shower. In particular the energy measured in the HCAL tower with a ΔR from the photon SC smaller than 0.15 must not exceed 10% of the supercluster energy. The latest studies about photon performances are obtained targeting $H \rightarrow \gamma\gamma$ process and are reported in [111]. In the barrel the energy resolution is about 1%, whereas in the endcaps, the resolution worsen up to about 2.5 in the case R_9[7] ≥ 0.94. The resolution for photons converted in e^+e^- pair is about 1.5% in the Barrel and about 3-3.5% in the EndCap. The resolution as a function of $\mid \eta \mid$ is reported in Fig. 3.7.

3.2.4 Neutral Particles and Charged Hadrons

Once muons, electrons, and isolated photons are identified the corresponding PF blocks are removed from the PF blocks collection. The remaining particles to be identified are hadrons from jet fragmentation and hadronization. These particles can be detected as charged hadrons, neutral hadrons, non isolated photons, or muons produced by early decays of charged hadrons.

The neutral hadrons and photons are reconstructed from all the clusters in ECAL and HCAL not linked to a track. Two different cases can be distinguished. If the

[7] The R_9 variable is defined as the energy sum of the 3×3 crystals centered on the most energetic crystal in the supercluster divided by the energy of the supercluster itself. This variable is used to discriminate between converted and unconverted photons, indeed the showers of photons that convert before reaching the calorimeter have wider transverse profiles and lower values of R_9 than those of unconverted photons.

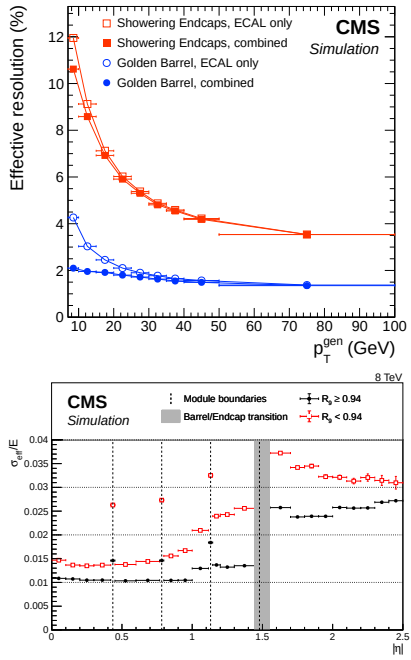

Figure 3.7: (*top*) Resolution in electron momentum after combining the E_{SC} and p estimates (solid symbols), compared to that of the SC energy only (open symbols), as a function of the electron p_T. The circles points represent all the electrons in the barrel not affected by bremsstrahlung, whereas the squares points show the electron resolution in the case of tight radiative losses. Electrons are generated with uniform distributions in η and p_T [108]. (*bottom*) Relative energy resolution, σ_{eff}/E, as a function of $|\eta|$, in simulated $H \rightarrow \gamma\gamma$ events, for photons with $R_9 \geq 0.94$ (solid circles) and photons with $R_9 < 0.94$ (open squares). The vertical dashed lines mark the module boundaries in the barrel, and the vertical grey band indicates the range of $|\eta|$, around the barrel/endcap transition, removed from the fiducial region [111].

calorimeter cluster falls within the tracker acceptance ($|\eta| < 2.5$), a photon is associated with the ECAL SC and a neutral hadron with the HCAL cluster[8]. If the calorimeter cluster is outside the tracker acceptance region, a photon is associated with the ECAL cluster if there is no presence in the PF block of an HCAL cluster. On the contrary if the HCAL cluster is present in the PF block, the combination of ECAL+HCAL cluster give rise to a hadron.

[8]This choice is supported by the fact that usually 25% of the energy of a charged hadron is carried by photons, whereas neutral hadrons leave only 3% of their energy in the electromagnetic calorimeter

In the case where the clusters can be linked to more tracks the particles are determined comparing the sum of the tracks momenta and the calorimetric energy as explained in the following.

The simplest case happens when the energy of the calorimeter cluster is compatible with the sum of the tracks momentum. In this case a charge hadron is reconstructed and its momentum is recomputed applying a χ^2 fit between the measurement in the tracker and in the calorimeter.

If the energy in the cluster is in excess of the sum of the track momenta by an amount larger than the expected calorimetric energy resolution of the cluster, the excess is interpreted as photons or neutral hadrons depending of the amount of the excess. Indeed, if the excess is smaller than the total ECAL energy and larger than 500 MeV, it is identified as a photon. Otherwise, the recalibrated ECAL energy still gives rise to a photon, and the remaining part of the excess, if larger than 1 GeV, is identified as a neutral hadron. A charged hadron is then reconstructed from each track and its momentum is taken from the track momentum.

The last case identifies all the situations where the energy in the cluster is significantly smaller compared to the sum of the tracks momentum by three standard deviations of the energy. In this case a relaxed search for muons with little deposit in the calorimeter is performed. As soon as a new muons is detected, its momentum is subtracted from the sum of the track momenta associated with the calorimeter cluster. If the sum of the track momenta is still larger than the cluster energy, the excess in momentum is assumed to arise from residual mis-reconstructed tracks which are removed until no such tracks remain in the PF block or until the momentum excess disappears, whichever come first.

3.2.5 Jets Reconstruction

Jets are reconstructed by clustering candidate particles obtained using the PF algorithm. In particular the results of the analysis are obtained using the *anti-K_T* algorithm (anti-kT) [112].

The process of collecting particles is guided by the distance between them defined as:

$$d_{ij} = min\left(k_{ti}^{2p}, k_{tj}^{2p}\right)\frac{\Delta_{ij}^2}{R^2} \tag{3.1}$$

and the distance of each particle *i* from the beam defined as:

$$d_{iB} = k_{ti}^{2p}, \tag{3.2}$$

where $\Delta_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$, k_{ti}, η_i and ϕ_i are the transverse momentum, the rapidity and the azimuth angle of the ith particle. The parameter R sets the

scale of the distances between particles and basically determines the maximum distance for particles to be added to jets. In the case of the anti-kT, the distance from the beam coincides with the square of the inverse of the transverse momentum of the ith parton. The distance between particles (Eq. 3.1) also depends on the values of the parameter p which is, for the particular case of the anti-kT, equal to -1.

The clustering proceeds by identifying the smallest of the distances among the d_{ij} and d_{iB}. If the smallest distance is identified by d_{ij} then the ith and jth particle are recombined in a single particle, in the other case ith particle is promoted to a jet and removed from the list of the particles. The distances are recalculated and the procedure repeated until no distance is left.

Considering the definition in the anti-kT, the distance d_{ij} between a hard and a soft particle is smaller then the distance between two soft particles, and the algorithm tends to cluster soft particles with hard ones rather than cluster soft particles with each other.

After this stage the jet momenta need to be calibrated. There are several factors that justify a need for calibrating the jet energies like the non linear response of the calorimeter, the detector resolution, the electronic noise, the pileup, the underlying events. Corrections can also be defined as a function of the flavor of the initial partons. CMS has developed a factorized multi-level jet energy corrections (JEC) to be applied in a precise fixed order [113]. The first jet energy correction is used to remove contamination from pileup, underling events and electronic noise (label as L1 correction). The contamination is event based and it is estimated considering the mean density energy released by these sponger processes, weighted by the area associated with each reconstructed jet [114,115]. The second and the third corrections are related to possible variations in jet response with pseudorapidity and transverse momentum (label as L2 and L3 corrections rispectively). Others corrections are released centrally by the experiment to further improve the jet energy resolution but they are not mandatory [113].

To further reduce the effect of pileup interactions, a dedicated pileup-jet identification algorithm [116] has been developed to identify pileup jets (Pileup Jet ID). The algorithm exploits both the vertex and the shape properties of a jet. Indeed, each pileup jet presents two features that allow its identification. The first characteristic is related to the jet shape, indeed pileup jets are more spread compared to the hard jets which are collimated. The second feature takes advantage from the fact that pileup charged particles don't point to the primary vertex. These characteristics allow the pileup identification both in and out the tracker acceptance. For each jet reconstructed by the detector, both vertex and shape informations are combined through a multivariate analysis and the output discriminator is used in order to discarded the jet in the case it is identified as a pileup jet.

Another technique commonly used in CMS and orthogonal to the pileup jet id is known as Charged Hadron Subtraction [117]. This procedure consists in removing from the collection of particles used to cluster the jets all the charged hadrons which can clearly be associated with a pileup vertex. Those ones without a clear association or associated with the high p_T vertex remain in the jet. This identification procedure is applied on the jet objects before the jet energy corrections. The remaining offset energy contamination in the hard jets has to be subsequently subtracted using the L1 corrections.

The Jet Energy Resolution is reported in Fig.3.8. The study is performed comparing the performance in different pileup scenari considering $\gamma + jets$, and $Z + jets$ sample as a function of the jet transverse momentum. Two different pseudorapidity region are reported. In both the regions the resolution is stable against pileup for jets with $p_T \geq 100$ GeV. In particular it is better than 10% above p_T = 100 GeV and 5% when $p_T > 1$ TeV. In the low momentum region ($p_T < 20$ GeV) the performances degrade up to 30% in the high pileup case (μ = 75). In both examples the CHS is applied.

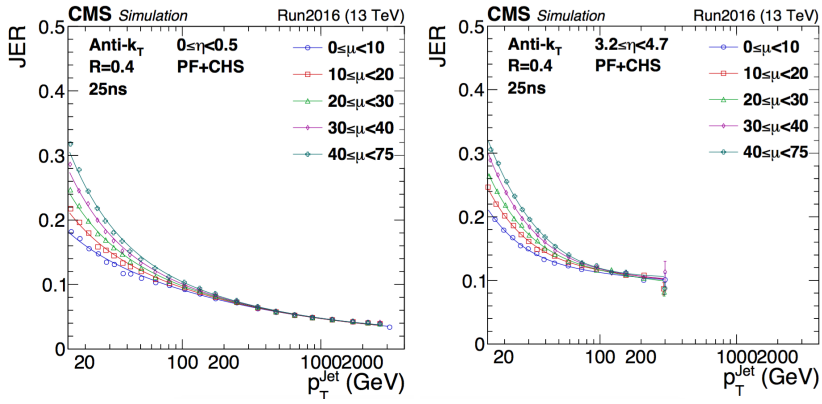

Figure 3.8: Jet energy resolution (JER) as a function of the transverse momentum measured in simulation ($\gamma + jets$, and $Z + jets$) after the pileup cleaning. The jets are reconstructed using the anti-kT clustering with R = 0.4. The JER is reported as a function of the transverse momentum and different pileup scenario identified by the mean number of interaction μ. For sake of illustration two pseudorapidity region are reported $0 <| \eta |< 0.5$ and $3.2 <| \eta |< 4.7$ [118].

Another important task which obliged the CMS collaboration to develop a specific tools for it, is the identification of jets originating from b quarks [119]. The tools is called *Combined Secondary Vertex v2* (CSVv2) algorithm. Using a multivariate technique, the CVSv2 algorithm takes advantage from different informations as

the secondary vertex, the impact parameter significance and finally the jet kinematics. The tool is able to discriminate between b quarks from c quarks, and b quarks from light quarks. Several working points for CVSv2 discriminator are available: loose, medium, and finally tight points. Each of them are defined as the values of the discriminator cut for which the rate for misidentifying a light jet as a b jet is 10%, 1%, and 0.1%, respectively, as measured for jets with p_T greater then 25 GeV in $t\bar{t}$ sample and p_T greater than 40 GeV in multijet sample [120].

3.2.6 Missing Energy Reconstruction

Although neurtrinos or new particles predicted by some exotic theoretical models do not leave any signal in the detector, their presence can be inferred indirectly from the momentum imbalance in the plane perpendicular to the beam direction. This asymmetry is quantified by the missing transverse momentum, $\vec{E}_T^{miss}$, and its magnitude, the missing transverse energy (E_T^{miss}). CMS presents several distinct and complementary algorithms to reconstruct the E_T^{miss} [121]. The one used in the analysis is the PF missing transverse momentum (PF $\vec{E}_T^{miss}$) defined as the negative vectorial sum over the transverse momenta of all PF particles reconstructed in the event:

$$\vec{E}_T^{miss} = -\sum_{i=1}^{N_{Pt}^{PF}} \vec{P}_{T,i} \tag{3.3}$$

A less commonly used algorithm is the Calo $\vec{E}_T^{miss}$ which uses the energies contained in calorimeter towers and their directions relative to the centre of the detector. The sum excludes energy deposits below noise thresholds but is corrected for the calorimeter deposits of muons, when they are present, by adding their momentum to the sum [122].

As for jets, also the $\vec{E}_T^{miss}$ and its magnitude needs to be calibrated to correct for the nonlinearity of the response of the calorimeter, inefficiencies in the tracker or calorimeters, thresholds in the calorimeter clusters and in the transverse momentum. The bias can be reduced correcting the transverse momentum of the jets involved in the missing transverse energy computation applying the jet energy corrections (JEC) on top of them.

$$\vec{E}_T^{miss,Type-I} = \vec{E}_T^{miss} - \sum_{jets}(\vec{P}_{T,jet}^{corr} - \vec{P}_{T,jet}) \tag{3.4}$$

The E_T^{miss} is also corrected for pileup. This correction removes all the charged hadrons originating from secondary vertex but also the neutral pileup particles. The estimation of the pileup neutral component is not easy. Under the assumption of no genuine missing energy in pileup events, the neutral and charged pileup contributions are assumed to be exactly equal [121]. Finally, a further correction has to be applied

to take into account ϕ asymmetries, which cause a variation in the average $\vec{E}_T^{miss}$ components along the x and y axis [121]. This asymmetry is caused by imperfect detector alignment, inefficiencies, a residual ϕ dependence of the calibration, and a shift between the centre of the detector and the beam line [123].

The resolution of the missing transverse energy is dominated by the hadronic activity in the event and it is measured using candle process like the Drell-Yan (DY) characterized by no physical $\vec{E}_T^{miss}$. The transverse momentum of the Z or γ boson is used as a reference. In principle this momentum should be totally balanced by the hadronic activity in the event. The parallel and orthogonal components to the p_T of the Z or γ boson are measured and used to quantify the resolution (Fig. 3.9). The distributions of these two components are presented in Fig. 3.10.

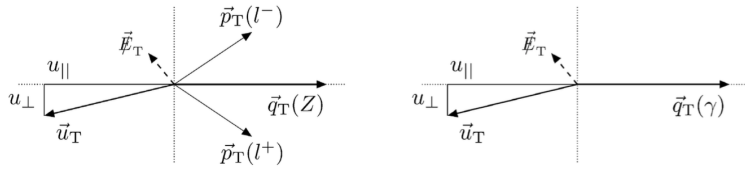

Figure 3.9: Illustration of the Drell-Yan event kinematics in the transverse plane for the Z boson (*left*) and the photon (*right*). The vector $\vec{u}_T$ denotes the vectorial sum of the transverse momentum of all particles reconstructed in the event except for the two leptons from the Z decay or the photons [121].

In parallel to PF E_T^{miss}, the CMS experiment has developed different alternative strategies to estimate $\vec{E}_T^{miss}$ in order to face up the degradation of the missing transverse energy resolution with the increasing of the pileup [121]. One of these techniques, developed for Run-II, is the pileup per particle identification (PUPPI) [125] and attempts to use local shape informations, event pileup properties and tracking informations to reduce the pileup dependence on various jet and E_T^{miss} variables. A weight for each particle is computed on an event-by-event basis based on the local shape for a pileup associated to charged particles to rescale their momenta depending on their pileup likeliness. Although the two different approaches have the same performance in presence of low pileup condition, the discrepancies between the two algorithms tends to increase with the increase of the vertices. In the case of 25 vertices the relative discrepancy is about 20-25%. In addition as presented in Fig. 3.11, the missing transverse energy estimated by PUPPI is less affected by the increase of the pileup since the difference of the resolution in both the recoil components between

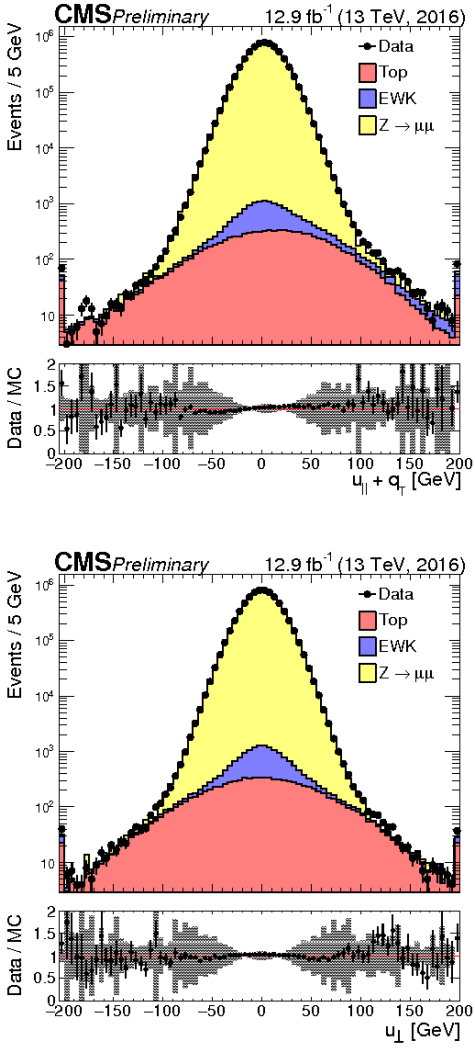

Figure 3.10: Distributions of u_T (*top*) and $u_{\parallel} + q_T$ (*bottom*) for PF E_T^{miss} for $Z \to \mu^+\mu^-$. The points in the lower panel of each plot show the data/MC ratio, including the statistical uncertainties of both data and simulation. The grey error band displays the systematic uncertainty of the simulation. The first (last) bin contains the underflow (overflow) content. The increased uncertainty in the u_T and $u_{\parallel} + q_T$ distributions around ± 70 GeV is due to the jet energy resolution uncertainty which is bigger in the region where the boson goes with an hard jet [124].

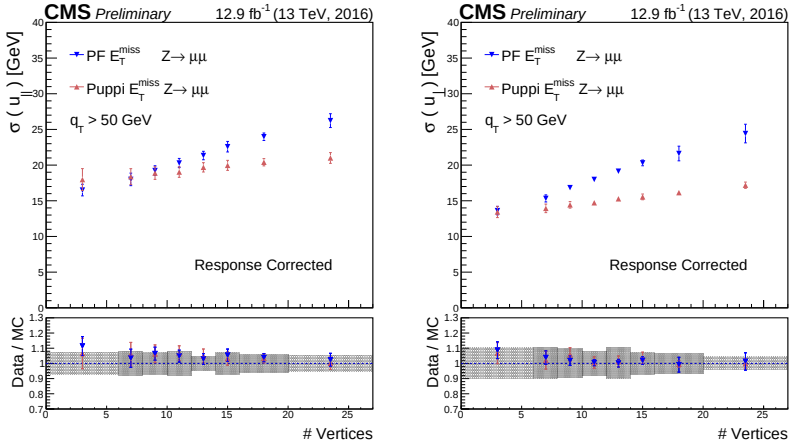

Figure 3.11: Resolution curves of the parallel (*left*) and perpendicular (*right*) recoil component versus the number of reconstructed vertices. The comparison between PF E_T^{miss} (blue triangle) and PUPPI E_T^{miss} (red triangle) is presented in the plots. In all the distributions the events in $Z \rightarrow \mu^+\mu^-$ channel are reported. The upper frame of each figure shows the resolution in data, instead the lower frame shows the ratio of data to simulation [124].

the « no pileup » case (3 vertices) and the « pileup » case (25 vertices) is of the order of GeV, whereas for the PF E_T^{miss} the discrepancy is about 20 GeV.

3.2.7 Lepton Isolation

Lepton isolation is one of the main variable used to select prompt muons and electrons produced during electroweak processes and for rejecting the leptons produced in jets by the decay on flight of hadrons. The isolation is quantified by estimating the magnitude of the total transverse momentum of the particle in a cone with radius 0.3 or 0.5 around its direction (Eq. 3.5),

$$I_{PF} = \frac{1}{p_T} \cdot \left(\sum_{h^\pm} p_T^{h^\pm} + \sum_{\gamma} p_T^{\gamma} + \sum_{h^0} p_T^{h^0} \right) \tag{3.5}$$

where the sums run over charged hadrons ($h^\pm$), photons (γ), and neutral hadrons (h^0), and p_T is the magnitude of the transverse momentum carried by each particle in the sums. As can be understood, this variable is strongly affected by the pileup activity during the collision. To mitigate the impact of pileup on the discrimination based on isolation, the expected pileup contributions are subtracted from the neutral hadrons and photons components. Only charged hadrons associated with the primary

vertex are used in the computation of the isolation. Three different strategies are adopted to estimate the mean contribution of the pileup depending if a muon or an electron/photon is considered.

In the first case the *delta-beta* correction is applied. This correction is achieved by subtracting from the neutral contribution a fraction of the sum of the transverse momenta of the charged particles not originating from the primary vertex. The isolation for the muon case can then be rewritten as:

$$I_{PF} = \frac{1}{p_{\mathrm{T}}(\mu)} \left[\sum_{h^{\pm}} p_T^{h^{\pm}} + \max(0, \sum_{h^0} p_T^{h^0} + \sum_{\gamma} p_T^{\gamma} - \Delta\beta \cdot \sum_{h^{\pm},pileup} p_T^{h^{\pm}}) \right] \quad (3.6)$$

The factor $\Delta\beta$ is estimated to be equal to 0.5 which corresponds approximately to the ratio of the neutral particles to charged hadron in inelastic proton-proton collision [126].

When dealing with electron candidates, the neutral flux is corrected by using the average energy density due to pileup and underlying event (ρ), and an effective area (A_{eff}) which is an ad-hoc coefficient chosen precisely to achieve an isolation efficiency independent from pileup. Figuere 3.12 reports the energy density of the electron produced by a Z boson as a function of the number of vertices generated during a collision between two bunches.

$$I_{PF} = \frac{1}{p_{\mathrm{T}}} \left[\sum_{h^{\pm}} p_T^{h^{\pm}} + \max(0, \sum_{h^0} p_T^{h^0} + \sum_{\gamma} p_T^{\gamma} - A_{eff} \cdot \rho) \right] \quad (3.7)$$

In the case of photons, it can happen that some deposits of its energy are missed during the energy clusterization around its seed. This energy leaks are instead used during the energy clusterization of hadrons or secondary photons that are used later to estimate the γ isolation. This effect is p_T dependent, and it becomes significant at high transverse momentum, whereas it is dwarfed by the noise contribution at low momentum. To handle this problem the requirements on the photon isolation are parametrized with respect to its transverse momentum and they are applied separately to each contribution ($h^{\pm}$, γ, and h^0) used to estimate the isolation (Tab. 4.8).

ρ-corrected PF charged hadron	α
ρ-corrected PF neutral hadron	$\beta + \theta \cdot p_T + \delta \cdot p_T^2$
ρ-corrected PF photon	$\omega + \mu \cdot p_T$

Table 3.1: Generic transverse momentum parametrization to estimate the photon isolation thresholds as a function of the photon p_T.

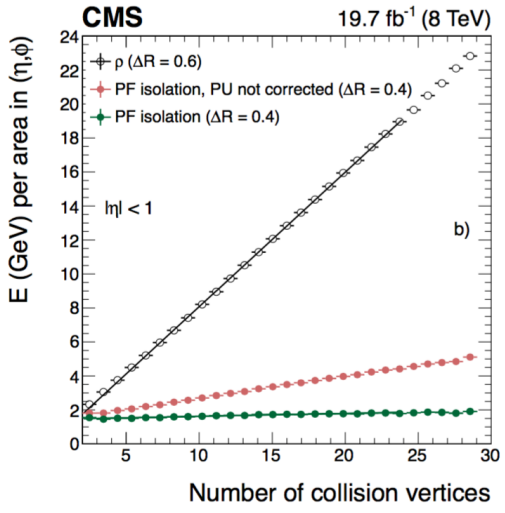

Figure 3.12: Average energy density as a function of the number of reconstructed p-p collision vertices, for electron candidates with $p_T > 20$ GeV and $|\eta| < 1$ in $Z \to e^+e^-$ events. The figure reports the energy density ρ (open dots) and the energy density on the PF isolation in the electron before (pink dots) and after (green dots) the correction for pileup [108].

3.3 Simulation

In the previous paragraphs and chapters it was described how it is possible to reconstruct the data produced by particle detectors in order to identify and measure the particles produced in the collisions. All this procedure would be totally useless if there wasn't anything to compare with. The goal of the simulation is precisely to facilitate the comparison of the observed data with those that would be expected from a model that is already known (background).

In CMS the process of obtaining simulated data is achieved through three different steps:

- *Generation*. The production of a samples (background or signal) starts with the generation of events that describe the physical processes happening in the LHC collisions. A process can be divided in two main parts. The first part is the hard scattering, which identifies all the phenomena that occur at a very high energy regime and can be described by a perturbative approach. The second part can be described by the non perturbative processes like the showering and the hadronization of the partons in the final event.

- *Simulation*. The particles produced in the physical process are propagated in this step through the experimental setup. In this step a detailed description of the magnetic field and material of the CMS apparatus, must be provided to simulate the propagation of the particles to the detecting layers, including their interactions that can give rise to secondary particles.

- *Digitization*. This step provide a simulation of the detector response. This step also includes the effect of the pileup, and the emulation of both the L1 and the HLT trigger.

After these three steps, simulated data can be processed through the reconstruction software exactly like the real data produced by the CMS detector. In other words, during the reco-step all the algorithms described previously in the chapter are run to reconstruct and identify the particles that were simulated during the previous three steps

The data produced by the reconstruction step is not yet the one used during the final analysis. A skimming process is applied on top of it to obtain an high-level data tier object to serve the needs of the mainstream physics analyses while keeping a small event size (30-50 kb/event). The final sample is denominated "MiniAOD" [127] and its main content is:

- High level physics objects (leptons, photons, jets, E_T^{miss}), with detailed information in order to allow e.g. retuning of identification criteria. Some preselection requirements are applied on the objects. If the objects fail them they are

either not stored or stored only with a limited set of informations. The selection is quite basic and it consists of a cut on the transverse momentum whose value is particle dependent. Two jets collections with different R value (0.4 and 0.8) are stored in the sample. Some high level corrections are applied, like the sets of the Jet Energy Corrections which are mandatory for the jets or the Type-I corrections for the missing transverse energy.

- Basic information for all PF candidates is included in Mini-AOD: four-momentum, charge, impact parameters, and particle type (electron, photon, etc.). Quality flags related to the association with the interaction vertex and tracking hits in the detector are also provided. To reduce as much as possible the occupied space, all the information is stored with a limited precision of about 0.1%. Finally, the physics objects and PF candidates are fully cross referenced, so each object is linked to the PF candidates from which it was reconstructed. The presence of PF candidates is useful in the case an analysis needs to work with objects that are not provided centrally and allow a better performance of the analysis itself.

- MC Truth information: a subset of the particles produced in the hard scattering process, jet flavor information, and final state leptons and photons are saved. Jets with p_T > 8 GeV are also stored, and so are the other MC summary information (e.g event weights, LHE header, PDF, PU information). In addition, all the stable particles associated with the final state are also saved, to allow reclustering of jets with different algorithms and substructure studies.

- Trigger information: the trigger bits associated with all the trigger paths, and all the trigger objects that have contributed to firing at least one path within the trigger are stored in the file. In addition, all the objects reconstructed at L1, the L1 global trigger output, and the prescale values of all the triggers are also stored in the MiniAOD.

The production of simulated samples is usually implemented centrally by the collaboration during some very precise periods of the data taking. All the samples produced during the same campaign share the same configurations, the same tuning and pileup conditions, the same beam energy, and the same software. For data, only the reconstruction and the skimming steps are needed. Indeed the Generation step is superseded by real physics collisions, the simulation and the digitization step is substituted by the real CMS detector.

Chapter 4

Search for a spin-zero high mass resonance with the $ZZ \to 2l2\nu$ channel using 2016 data in CMS

A search for a generic scalar resonance decaying in two Z bosons is described in this chapter. The analysis has been performed using data collected by the CMS experiment during 2016 for an integrated luminosity of 35.9 fb^{-1}. All the aspects of the analysis, starting from the development of the analysis framework and up to the extraction of the final exclusion limits, have been carried out during the last three years of the doctoral work in collaboration with the "Université Libre de Bruxelles", the "University of Delhi", and "Beihang University". All the results presented in this chapter have been used to perform a combination analysis with two other decay channel of the $H \to ZZ$ family: $ZZ \to 4l$ and $ZZ \to 2l2q$ [128].

4.1 Introduction

After the observation of a new boson with mass close to 125 GeV [1, 2] by the ATLAS and the CMS experiments at the LHC many questions seek experimental answers. As presented in the first chapter of this thesis, direct searches for an additional scalar boson can play an important role in this process, and are well motivated in the context of many models beyond the Standard Model (SM) like the 2HDM, EWS, and MSSM[1]. CMS has excluded a heavy SM-like Higgs boson in the mass range between 200 GeV and 1000 GeV using the full Run-I dataset [71]. With an increased center-of-mass energy of 13 TeV, it is possible to extend this search to higher masses, and probe for new physics up to 3 TeV.

The most direct way to discover a new particle is to observe a resonance at a specific mass (m). The mass reconstruction can be achieved taking advantage from the relativistic relation between the energy (E), the tri-momentum ($\vec{p}$), and the mass of the particle:

$$m^2 = E^2 - p^2 = E^2 - (p_x^2 + p_y^2 + p_z^2) = E^2 - (\vec{p}_T^{\,2} + p_z^2) \tag{4.1}$$

Looking at the simple case of a resonance decaying in two particles (two-body decay) and following the notation reported in the Fig. 4.1, the previous mass relation can be rewritten as

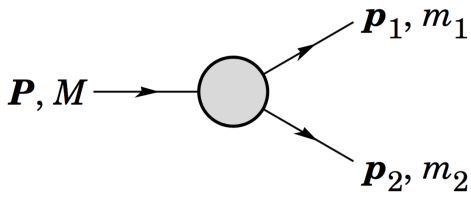

Figure 4.1: Definitions of the variables for the two-body decays. P, p_1 and p_2 identifies the tri-momentum of the particles involved in the process, whereas M, m_1, and m_2 their mass.

$$\begin{aligned} m^2 &= E^2 - p^2 = (E_1 + E_2)^2 - (p_1 + p_2)^2 \\ &= (E_1 + E_2)^2 - ((\vec{p}_{1,T} + \vec{p}_{2,T})^2 + (p_{1,z} + p_{2,z})^2) \end{aligned} \tag{4.2}$$

[1] A search for a scalar boson in the context of EWS and 2HDM (Type-I and Type-II) was performed in the $H \to ZZ \to 2l2\nu$ channel using the data collected in 2015 by the CMS detector. Their results are reported in this article [47] but they will not be described in this manuscript.

When all decay products of the particle can be properly reconstructed and identified the mass is well defined. This is the case of the Drell-Yen process when the Z boson decays in a pair of electrons or a pair of muons. However, it can also happen that the final state is composed by particles which don't interact with the detector like the neutrinos, as the case of the W boson when it decays leptonically or the case of the Z boson when it decays in two neutrinos. A similar situation occurs in the $ZZ \to 2l2\nu$ channel where the event can't be totally reconstructed due to a lack of information originating from the presence of neutrinos.

Despite this problem, a partial identification can be achieved exploiting the conservation of the transverse component of quadri-momentum. Indeed, since both the colliding particles gee along the longitudinal direction and the momentum has to be conserved, the total final momentum on the transverse plane has to be zero. The estimation of the total transverse momentum of the final invisible particles must balance the total transverse momentum of the visible particles. The imbalance of the transverse momentum is quantified by the missing transverse momentum ($\vec{E}_T^{miss}$) as already explained in Chapter 3.

In this context, even if the direct measurement of the mass can't be achieved anymore, its value can be constrained with the transverse mass variable (M_T) [129]. Defining the transverse energy E_T as $\sqrt{M^2 + \vec{p}_T^2}$, the transverse mass is defined as

$$M_T^2 \equiv (E_T(1) + E_T(2))^2 - (\vec{p}_T(1) + \vec{p}_T(2))^2 \tag{4.3}$$

The decay channel used in this analysis search is characterized by the presence of two Z bosons decaying respectively in two leptons and two neutrinos. For this reason the property of the missing transverse energy and the transverse mass are exploited in order to search for hints of possible new heavy scalar particles.

More precisely, the following analysis targets a generic heavy scalar resonance (indicated as X for the entire chapter) with mass between 200 GeV up to 3 TeV and width between 5 GeV up to 100 GeV. The search is performed in the $ZZ \to 2l2\nu$ channel characterized by the presence of large missing transverse energy (E_T^{miss}). The selection of the signal region is based on a pure identification of events with a $Z \to l^+l^-$ ($l = \mu, e$) candidates in different event categories characterized by different numbers of accompanying jets. The optimization of the search in both flavor channels requires a good understanding of the background processes as well as precise predictions, and simulations, of the assumed heavy scalar signal. Where feasible, the backgrounds are estimated using data-driven methods as in the case of Drell-Yan and "non-resonant" backgrounds (i.e. di-lepton events not produced from a Z candidate). Alternatively, simulation predictions are used as in the case of the irreducible background (ZZ, WW, ZVV). The signal is modeled taking into account its interference with the continuum background ($pp \to ZZ$) and the SM Higgs.

Finally, the statistical analysis is performed using the transverse mass shape distribution.

4.2 Datasets and Simulated Samples

To test the hypothesis of a heavy scalar boson three ingredients are needed: the real data measured by the experiment, the background simulations, and finally the signal simulations.

4.2.1 Data

The Data used in the analysis are the ones collected and certified by CMS during 2016 for an integrated luminosity of 35.9 fb^{-1}. The search is performed using the double and single lepton trigger datasets. The Double lepton dataset (`DoubleElectron`, `DoubleMu`) is composed by all the events which present at least two leptons with a transverse momentum greater than a precise threshold depending on the trigger path used. In the particular case of the $ZZ \rightarrow 2l2\nu$, the final state must have two leptons with same flavor. In addition, to recover some inefficiency in the double lepton trigger dataset, the single muon trigger dataset (`SingleMu`) and the single electron trigger dataset (`SingleElectron`) are used. Finally, the photon (`Photon`) and the electron-muon (`MuEG`) trigger datasets are used to study some precise region of the phase space in order to estimate some background contribution directly from data (data-driven approach). The trigger paths used in this analysis will be properly explained in the section 1.3.2 of this chapter, whereas the Table 4.1 summarizes all the datasets used in the analysis.

Dataset	$\int \mathcal{L}(fb^{-1})$	Run range
/%%/Run2016B-23Sep2016-v3	5.933	273150-275376
/%%/Run2016C-23Sep2016-v1	2.646	275656-276283
/%%/Run2016D-23Sep2016-v1	4.353	276315-276811
/%%/Run2016E-23Sep2016-v1	4.117	276831-277420
/%%/Run2016F-23Sep2016-v1	3.186	277932-278808
/%%/Run2016G-23Sep2016-v1	7.721	278820-280385
/%%/Run2016H-PromptReco-v2	8.636	281207-284035
/%%/Run2016H-PromptReco-v3	0.221	284036-284068
Total	**35.9**	

Table 4.1: MiniAod datasets used for $ZZ \to 2l2\nu$ analysis. The symbol %% stands for `DoubleElectron`, `DoubleMu`, `SingleElectron`, or `SingleMu`. `MuEG` and `Photon` datasets are used as control regions. The integrated luminosity and the run-ranges are shown for each data period.

4.2.2 Background

The final state used in the analysis can be mimicked by other SM processes with a similar topology. The main background is $pp \to ZZ \to 2l2\nu$. Its cross section is 2-4 order of magnitudes greater than the signal cross-section, which is 1-3 order of magnitudes lower than the tri-boson production processes, depending on the resonance mass. Hints of new physics are expected to emerge in regions populated by leptons with high transverse mass where the contribution of the background is expected to be low. In this phase space region event yield predictions can be decisive in the process of a discovery. For this reason the most precise theoretical tools as in the case of the Monte Carlo event generators are preferred. The background processes are produced with several Monte Carlo (MC) event generators at the NLO precision and all of them are normalized according to the cross section officially provided by the CMS experiment [130]. The $W \to l\nu$ sample is generated with MadGraph5_aMC@NLO [131] and its cross section is calculated centrally by the CMS experiment at NNLO QCD expansion precision. MadGraph5_aMC@NLO is used also to simulate the Drell-Yan process $Z/\gamma^* + jets$. For this process, two samples are considered: the first one requiring that the final lepton invariant mass is between 10 GeV and 50 GeV at generator level, whereas the second one requiring a mass greater than 50 GeV at generator level. The corresponding cross section is normalized at NLO in the first case and at NNLO in the second case. Top-pair production, $t\bar{t} \to 2l2\nu$, is produced with Powheg [132–135]

and normalized to its NNLO cross section. The $t\bar{t}V$ (V = W, Z) are generated with MadGraph5_aMC@NLO. The top quark produced in association with the W boson is simulated using Powheg normalized to the NNLO cross section. The single top quark production in the s-channel is modeled availing of MadGraph5_aMC@NLO, whereas the t-channel using Powheg. The di-bosons processes are produced using both MadGraph5_aMC@NLO and Powheg. WW ($WW \rightarrow 2l2\nu$, $WW \rightarrow l\nu 2q$) channels are simulated by Powheg but normalized to the NNLO cross section. The $ZZ \rightarrow 2l2\nu$ and $ZZ \rightarrow 2l2q$ are generated at NLO precision using Powheg and MadGraph5_aMC@NLO respectively. $WZ \rightarrow 3l\nu$ is generated with Powheg and $WZ \rightarrow 2l2q$ with MadGraph5_aMC@NLO normalized to the NNLO cross section. Finally the tri-boson channels are all simulated by MadGraph5_aMC@NLO and normalized to the MC generator cross-section. The shower is handled by Phythia8 [136] using the FxFx matching scheme [137, 138]. During 2016, the PDF used centrally were the NNPDF3.0 [139]. The full CMS detector simulation is based on GEANT4 package [140]. For the 2016 MC production all the samples are simulated assuming 25 ns bunch spacing. The entire list of background processes used in $2l2\nu$ analysis associated with their Monte Carlo generators, is reported in the table 4.2.

Process	Dataset	Generator	σ (pb) [QCD precision]
$W \rightarrow \ell\nu$	/WJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/S16	MadGraph5_aMC@NLO	61526.7 [NNLO]
$Z \rightarrow \ell\ell$	/DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/S16	MadGraph5_aMC@NLO	18610 [NLO]
	/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/S16	MadGraph5_aMC@NLO	5765.4 [NNLO]
$t\bar{t}$+X	/TTTo2L2Nu_13TeV-powheg/RunIISpring16MiniAODv2-PUSpring16	Powheg	87.31 [NNLO]
	/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/S16	MadGraph5_aMC@NLO	0.2043 [NLO]
	/TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8/S16	MadGraph5_aMC@NLO	0.2529 [NLO]
Single top	/ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1/S16	Powheg	35.85 [NNLO]
	/ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1/S16	Powheg	35.85 [NNLO]
	/ST_t-channel_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1/S16	Powheg	70.69 [NLO]
	/ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1/S16	MadGraph5_aMC@NLO	3.362 [NLO]
Dibosons	/WZTo3LNu_TuneCUETP8M1_13TeV-powheg-pythia8/S16	Powheg	4.4297 [NLO]
	/WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8/S16	MadGraph5_aMC@NLO	5.595 [NNLO]
	/WWTo2L2Nu_13TeV-powheg/RunIISummer16MiniAODv2-S16	Powheg	12.178 [NNLO]
	/WWToLNuQQ_13TeV-powheg/RunIISummer16MiniAODv2-S16	Powheg	49.997 [NNLO]
	/ZZTo2L2Nu_13TeV_powheg_pythia8/S16	Powheg	0.564 [NLO]
	/ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8/S16	MadGraph5_aMC@NLO	3.22 [NLO]
Tribosons	/ZZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/S16	MadGraph5_aMC@NLO	0.01398 [NLO]
	/WZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/S16	MadGraph5_aMC@NLO	0.05565 [NLO]
	/WWZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/S16	MadGraph5_aMC@NLO	0.16510 [NLO]

Table 4.2: A list of the SM dilepton processes used in the $ZZ \rightarrow 2l2\nu$ is reported. For each of them both the MC generator and the expected cross section is quoted. In particular for each cross section, it is explicitly underlined the order of the QCD perturbation order expansion at which they are estimated. S16 stands for RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6(_ext*)-v1.

4.2.3 Signal Samples and Signal Modeling

The target of the analysis search is a scalar particle with a mass between 200 GeV and 3 TeV that could be produced in either the ggF or the VBF processes, assuming a generic width. For both ggF and VBF the heavy resonance is generated initially by Powheg [64, 141]. Thirteen samples are produced for each process: nine mass points are simulated between 200 GeV and 1 TeV using steps of 100 GeV, and four mass points are generated between 1 TeV and 3 TeV using steps of 500 GeV. In order to test different width hypothesis, the signal samples are reweighted using the MELA package [76–79]. Three width values are tested: 5 GeV, 10 GeV, and 100 GeV. Finally, the decay in $ZZ \to 2l2\nu$ is handled by the JHU Generator (JHUGen) [76–79]. A detailed description of the reweighting procedure is given in the next lines.

MELA is a LO matrix-element tool based on JHUGen and MCFM generator [74, 75]. The reweigthing of the events generated by Powheg has two main purposes:

- test new width hypothesis for each mass point assuming SM couplings between particles.

- take into account the interference between signal, continuum and the SM higgs boson (h_1) both in ggF and in VBF.

The reweigthing procedure is event-by-event based and it starts by initializing MELA with the mass and width values that has to be tested. In particular, MELA needs to know the process that has to be modeled. To set the final limits on the production cross section of a generic scalar resonance, three different distinct processes are generated for each mass point: the heavy scalar resonance X, the background defined as $pp \to ZZ \to 2l2\nu + pp \to h_1 \to 2l2\nu$ and finally the combination of the signal and the background ($pp \to ZZ \to 2l2\nu + pp \to h_1 \to 2l2\nu + pp \to X \to 2l2\nu$). For each reconstructed signal event, MELA computes a weight based on the kinematic of the hard process generated by Powheg that could give rise to the considered event. In particular, MELA needs to know both the four-momentum and the flavor of each initial and final particle involved in the hard process. In the VBF case, the information of the additional jets in the final state is required. The kinematic of each event is used to compute the square of the corresponding matrix element both for the SM case and for each of the three processes previously mentioned at LO and based on the MCFM generator. The square of each matrix element identifies the probability that the event is generated by that specific physical process. The two numbers obtained considering the SM case and the three processes previously mentioned are then finally combined together in a single object defined by their ratio in order to remove from one side the Powheg SM weight carried by each generated event and on the other side rescale the

same event for the new weight. The final number is then multiplied to the event. In addition, in order to mimic the NLO kinematic availing of a LO matrix element, the real extra radiation is reabsorbed by adding its quadri-momentum to the one associated to its closest parton in the ΔR^2 plane.

There are four sources of interference in the $X \to ZZ \to 2l2\nu$: the interference with the $ZZ \to 2l2\nu$ continuum, the interference with the SM Higgs (h_1), the interference between the light Higgs and the continuum, and finally the interference with $WW \to 2l2\nu$. If the third contribution can be neglected [72] thanks to the requirement on the invariant mass of the final leptons $l^{\pm}$ ($|M_{l^+l^-} - M_Z| < 15$ GeV), the other three interferences cannot because they affect significantly both the shape and the normalization of the resonance. The distributions reported in figure 4.2-4.4 describe in detail how the interference affects both the shape and the normalization for two different mass points and two different values of the width. Following the notation reported in the equations 4.4

$$
\begin{aligned}
X =& \mid pp \to X \to ZZ \to 2l2\nu \mid^2 .\\
Cont + h_1 + Interf_{Cont-h_1} =& \mid pp \to ZZ \to 2l2\nu + \\
&+ pp \to h_1 \to ZZ \to 2l2\nu \mid^2 .\\
Cont + h_1 + X + Interf_{All} =& \mid pp \to ZZ \to 2l2\nu + \\
&+ pp \to h_1 \to ZZ \to 2l2\nu + \\
&+ pp \to X \to ZZ \to 2l2\nu \mid^2 .\\
Cont + X + Interf_{Cont-X} =& \mid pp \to ZZ \to 2l2\nu + \\
&+ pp \to X \to ZZ \to 2l2\nu \mid^2 .\\
Interf_{All} =& \mid pp \to ZZ \to 2l2\nu + \\
&+ pp \to h_1 \to ZZ \to 2l2\nu + \\
&+ pp \to X \to ZZ \to 2l2\nu \mid^2 + \\
&- \mid pp \to ZZ \to 2l2\nu \mid^2 + \\
&- \mid pp \to h_1 \to ZZ \to 2l2\nu \mid^2 + \\
&- \mid pp \to X \to ZZ \to 2l2\nu \mid^2 .\\
Interf_{Cont-X} =& \mid pp \to ZZ \to 2l2\nu + \\
&+ pp \to X \to ZZ \to 2l2\nu \mid^2 + \\
&- \mid pp \to ZZ \to 2l2\nu \mid^2 + \\
&- \mid pp \to X \to ZZ \to 2l2\nu \mid^2 .
\end{aligned}
\tag{4.4}
$$

the different distributions show that the interference between the continuum and the heavy scalar resonance increases the tail on the left of the peak whereas it re-

[2] $\Delta R = \sqrt{(\eta_1 - \eta_2)^2 + (\phi_1 - \phi_2)^2}$, where 1 and 2 refer to the considered particles.

duces the right part; the introduction of the light higgs inside the total amplitude has a big effect because overturns the previous interference contribution, leading to a total interference that is destructive on the left of the peak and constructive on the right.

The interference is also affected by both the mass of the resonance X and its width. Indeed, the interference occurs in the region of the phase space covered by at least two different processes. If both the continuum and the light Higgs have a known shape, the mass spectrum of the new scalar particle X changes according to its mass and its width values. From the distributions reported in Fig. 4.2 to Fig. 4.4 one can infer that the effect of the interference spreads more and more around the peak as the width increases, while its effects are totally limited close to the region around the peak for lower values of the width.

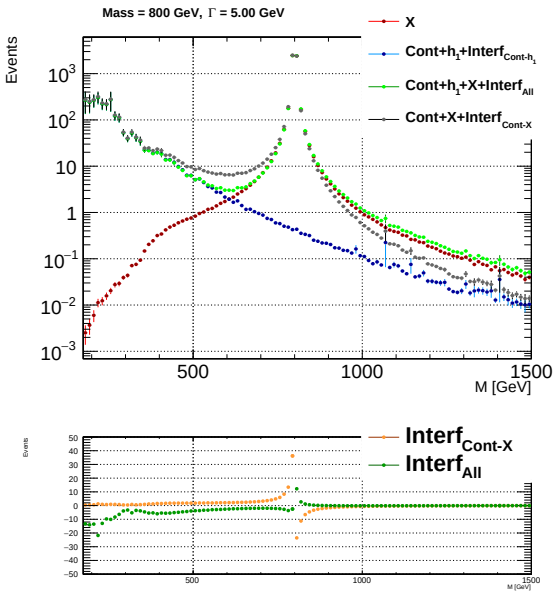

Figure 4.2: Mass spectrum of $ZZ \to 2l2\nu$ final state considering a mass of 800 GeV and width hypothesis of 5 GeV in gluon fusion mechanism. The red dots represent the heavy scalar boson (X), the blue shape ($\mid ZZ \to 2l2\nu + h_1 \to ZZ \to 2l2\nu \mid^2$), the light green dots represent the full amplitude considering also the contribution of the heavy boson, and finally the grey distribution ($\mid ZZ \to 2l2\nu + X \to ZZ \to 2l2\nu \mid^2$). The interference shapes are reported in the bottom plot in the case the SM higgs is neglected (orange distribution) or not (green distribution).

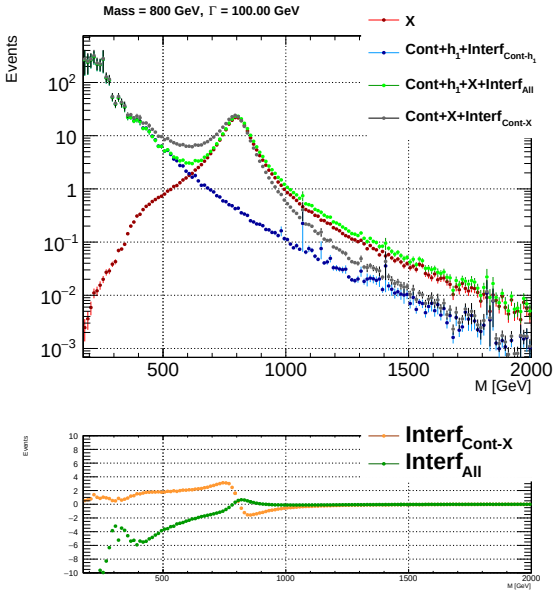

Figure 4.3: Mass spectrum of $ZZ \to 2l2\nu$ final state considering a mass of 800 GeV and width hypothesis of 100 GeV in gluon fusion mechanism. The red dots represent the heavy scalar boson (X), the blue shape ($\mid ZZ \to 2l2\nu + h_1 \to ZZ \to 2l2\nu \mid^2$), the light green dots represent the full amplitude considering also the contribution of the heavy boson, and finally the grey distribution ($\mid ZZ \to 2l2\nu + X \to ZZ \to 2l2\nu \mid^2$). The interference shapes are reported in the bottom plot in the case the SM higgs is neglected (orange distribution) or not (green distribution).

The last step concerning the signal modeling is related to the estimation of the signal cross section. Actually, for each background sample the cross section is obtained availing the theoretical SM predictions and it is used as an input parameter to rescale each sample to the correct luminosity[3]. This procedure can't be applied to the signal sample because no specific theoretical model is used as a benchmark. To face this problem, the cross section is estimated on top of the MELA reweigthing procedure for each mass and width value. In this situation the Powheg signal cross section is meaningless and then for simplicity it is set to 1 pb for each mass point. As a consequence all the different shapes that can be generated by MELA loose their

[3]The usual strategy to generate a sample is to fix the number of produced events in order to cover as much a possible the entire kinematical phase space (N_{prod}). It means therefore that each sample has its own luminosity ($\mathscr{L}_{tot} = N_{prod}/\sigma$) which doesn't correspond to the one measured by the experiment. For this reason each sample has to be normalized to recover the number of events that will be measured considering the experimental conditions of the data taking.

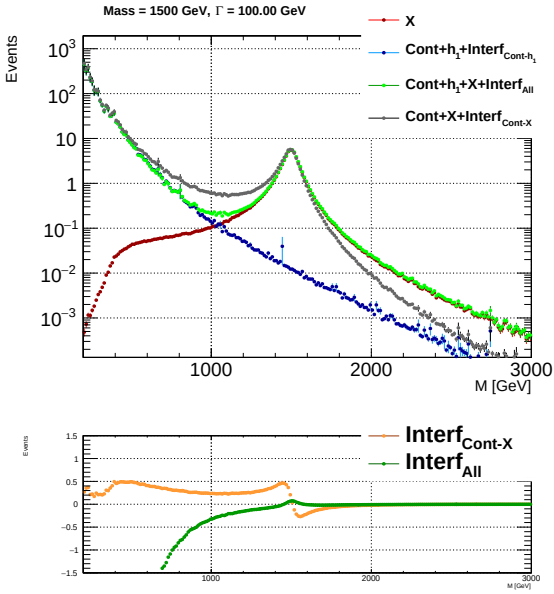

Figure 4.4: Mass spectrum of $ZZ \to 2l2\nu$ final state considering a mass of 1500 GeV and width hypothesis of 100 GeV in gluon fusion mechanism. The red dots represent the heavy scalar boson (X), the blue shape ($| ZZ \to 2l2\nu + h_1 \to ZZ \to 2l2\nu |^2$), the light green dots represent the full amplitude considering also the contribution of the heavy boson, and finally the grey distribution ($| ZZ \to 2l2\nu + X \to ZZ \to 2l2\nu |^2$). The interference shapes are reported in the bottom plot in the case the SM higgs is neglected (orange distribution) or not (green distribution).

right normalization which is restored by using the continuum background as it will be explained in the following lines.

In the case the MELA package is used to simulate the $pp \to ZZ$ process (MELA continuum), its predictions need to be compatible with the same physical process simulated by another MC generator like MCFM. The Fig. 4.5 presents a comparison between the MCFM mass spectrum and the mass shape predicted by MELA starting from the Powheg sample normalized to 1 pb. It is evident that the MELA shape is afflicted by a global normalization problem which is recovered scaling the shape to the reference distribution by applying a global scale factor ("continuum-SF"). The continuum-SF is defined as the ratio of the integral of the mass spectrum estimated in the two cases. The goodness of the rescaling procedure is reported in the same Figure 4.5. Although, the absolute normalization of the continuum background is not saved during the procedure, the relative normalization between MELA continuum and the other distributions obtained starting from the same Powheg sample is conserved.

Taking advantage from this relation, the continuum-SF can be used further to recover their right absolute normalization of any process obtained by MELA.

The final signal cross section is estimated swamping the relation $N_{events} = \sigma \cdot \mathscr{L}$, where the number of events N_{events} is computed as the integral of the mass spectrum for each mass and width hypothesis and $\mathscr{L}$ is the luminosity recorded by the CMS experiment during the 2016.

In the case of the VBF channel, where the proper signal sample was not centrally produced in $2l2\nu$ final state, the rescaling is performed following the same procedure just described but starting from the $4l$ final state at LO and taking into account the different branching ratio between the different channels decay $4l$ and $2l2\nu$ final state.

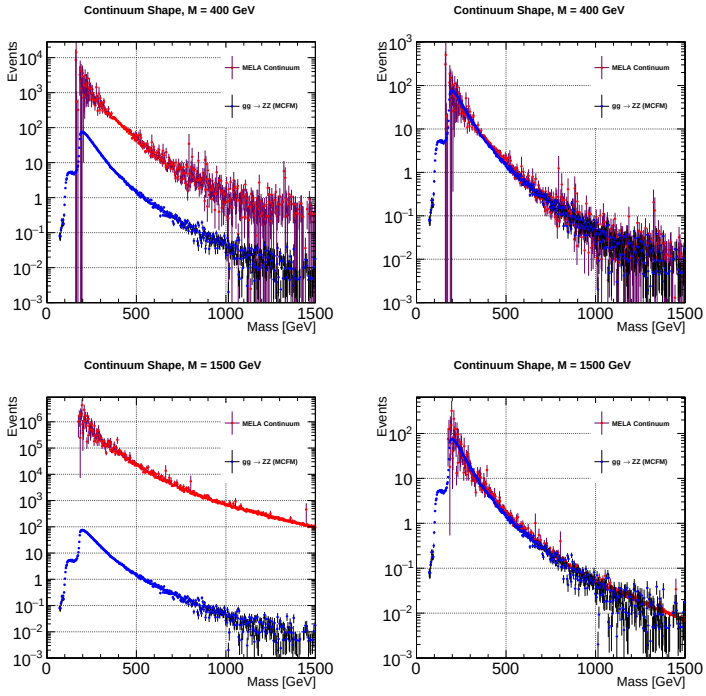

Figure 4.5: MELA Continuum mass spectrum for $gg \to ZZ$ process. The comparison is performed considering two mass values (400 GeV (*top*) and 1500 GeV (*bottom*)) before (*left*) and after (*right*) the continuum reweighting procedure.

This procedure was validated comparing the cross-section predicted using this recipe and assuming a heavy scalar resonance SM-like with the values computed by the "Higgs Cross Section Working Group" for a heavy scalar resonance SM-like and reported in the following article [60]. The comparison has shown a good agreement

between the two values. The final cross sections for each mass point and each width hypothesis are reported in the Table 4.3.

M_X (GeV)	Γ_X(GeV)	σ_{ggF} (pb)	σ_{VBF} (pb)
200	100	$7.16943 \cdot 10^{-4}$	$4.01088 \cdot 10^{-6}$
	10	$3.10052 \cdot 10^{-3}$	$1.41713 \cdot 10^{-5}$
	5	$5.73029 \cdot 10^{-3}$	$2.53557 \cdot 10^{-5}$
300	100	$9.60252 \cdot 10^{-3}$	$1.43903 \cdot 10^{-3}$
	10	$6.59359 \cdot 10^{-2}$	$1.09335 \cdot 10^{-2}$
	5	$1.27286 \cdot 10^{-1}$	$2.14829 \cdot 10^{-2}$
400	100	$1.76147 \cdot 10^{-2}$	$1.64836 \cdot 10^{-3}$
	10	$2.08345 \cdot 10^{-1}$	$1.50163 \cdot 10^{-2}$
	5	$4.22868 \cdot 10^{-1}$	$2.98804 \cdot 10^{-2}$
500	100	$1.72144 \cdot 10^{-2}$	$1.56461 \cdot 10^{-3}$
	10	$2.06409 \cdot 10^{-1}$	$1.52646 \cdot 10^{-2}$
	5	$4.18509 \cdot 10^{-1}$	$3.04568 \cdot 10^{-2}$
600	100	$1.21598 \cdot 10^{-2}$	$2.67087 \cdot 10^{-3}$
	10	$1.39248 \cdot 10^{-1}$	$2.69125 \cdot 10^{-2}$
	5	$2.81284 \cdot 10^{-1}$	$5.38395 \cdot 10^{-2}$
700	100	$9.97052 \cdot 10^{-3}$	$2.61042 \cdot 10^{-3}$
	10	$1.08503 \cdot 10^{-2}$	$2.68473 \cdot 10^{-2}$
	5	$2.17461 \cdot 10^{-2}$	$5.34991 \cdot 10^{-2}$
800	100	$7.23416 \cdot 10^{-3}$	$2.53221 \cdot 10^{-3}$
	10	$7.74346 \cdot 10^{-2}$	$2.64465 \cdot 10^{-2}$
	5	$1.55051 \cdot 10^{-1}$	$5.29065 \cdot 10^{-2}$
900	100	$6.98813 \cdot 10^{-3}$	$2.4129 \cdot 10^{-3}$
	10	$7.31695 \cdot 10^{-2}$	$2.52191 \cdot 10^{-2}$
	5	$1.4615 \cdot 10^{-1}$	$5.05514 \cdot 10^{-2}$
1000	100	$5.4009 \cdot 10^{-3}$	$3.26896 \cdot 10^{-3}$
	10	$5.57424 \cdot 10^{-2}$	$3.45785 \cdot 10^{-2}$
	5	$1.12064 \cdot 10^{-1}$	$6.97972 \cdot 10^{-2}$
1500	100	$1.93455 \cdot 10^{-3}$	$2.77825 \cdot 10^{-3}$
	10	$1.96021 \cdot 10^{-2}$	$2.91523 \cdot 10^{-2}$
	5	$3.9561 \cdot 10^{-2}$	$5.85943 \cdot 10^{-2}$
2000	100	$5.56375 \cdot 10^{-4}$	$1.49636 \cdot 10^{-3}$
	10	$5.40282 \cdot 10^{-3}$	$1.54154 \cdot 10^{-2}$
	5	$1.07575 \cdot 10^{-2}$	$3.11816 \cdot 10^{-2}$
2500	100	$1.89043 \cdot 10^{-4}$	$1.31605 \cdot 10^{-3}$
	10	$1.77042 \cdot 10^{-3}$	$1.3566 \cdot 10^{-2}$
	5	$3.51486 \cdot 10^{-3}$	$2.70974 \cdot 10^{-2}$
3000	100	$6.72087 \cdot 10^{-4}$	$7.61465 \cdot 10^{-4}$
	10	$5.99182 \cdot 10^{-4}$	$7.58377 \cdot 10^{-3}$
	5	$1.19645 \cdot 10^{-3}$	$1.5133 \cdot 10^{-2}$

Table 4.3: Cross section for each mass and width hypothesis of the signal hypothesis tested in the analysis. Both the Gluon Fusion and the Vector Boson Fusion are reported.

4.3 Selection and Efficiency

The second step of the analysis is represented by the selection of the $ZZ \to 2l2\nu$ candidate events and it can be summarized by the flow chart depicted in Figure 4.6. After a skimming obtained requiring that the events passed the kinematical requirements associated to at least one of the trigger path previously mentioned, the analysis proceeds by selecting a boosted Z boson candidate from a pair of electrons or muons, the events are then divided in different jet bins category. The events that are accidentally selected because they present a missing energy produced by a misreconstruction of the jets, or have a couple of same flavor leptons in the final state produced by a non-resonance boson, are discarded using both the transverse mass and the missing energy shape. Events where a Z boson candidate is potentially recoiling against neutrinos, which yield genuine E_T^{miss}, are selected and tagged as $ZZ \to 2l2\nu$ candidate events. The final limits are set using the transverse mass shape.

The next sub-chapters will be used to describe in more details each single step used in the selection workflow of the $ZZ \to 2l2\nu$ analysis.

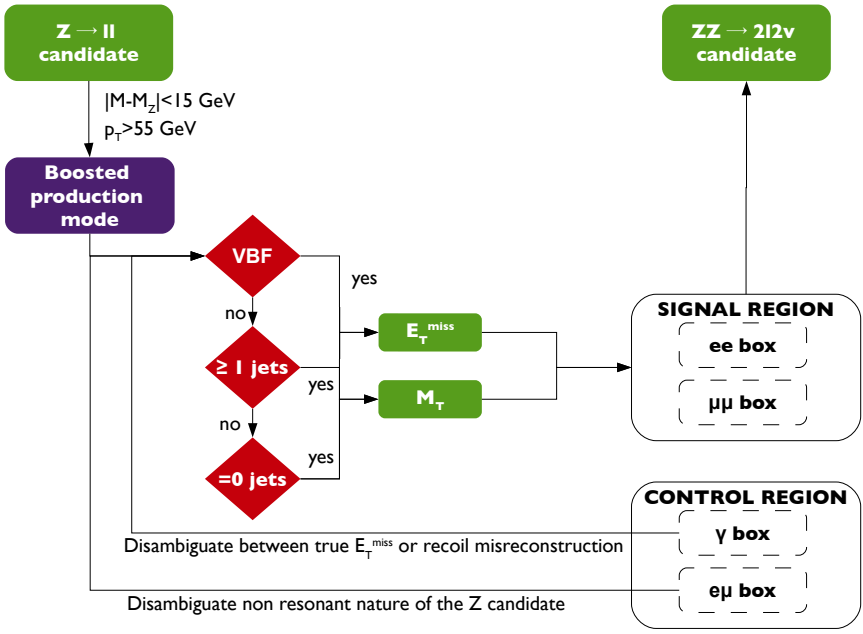

Figure 4.6: Schematic representation of the analysis.

The next lines will precede the description of the pre-selection cuts, and they aim to introduce all the ingredients used to identify the final state of this analysis.

4.3.1 Trigger

The first barrage met by the events is the trigger selection and actually only the events passing some precise trigger paths are used. The trigger paths used for the search of the signal are both the double lepton triggers (`DoubleElectron`, `DoubleMu`) and the single lepton trigger (`SingleElectron`, `SingleMu`). All these trigger paths are used in order to reach an efficiency close to 100% as the double lepton trigger inefficiencies are cured by the single lepton trigger. Double lepton triggers (`DoubleElectron`, `DoubleMu`) select two same flavor leptons (electrons or muons) requiring a threshold on their transverse momentum. In particular for the muon case, the leading lepton must have a $p_T > 17$ GeV and a $p_T > 8$ GeV for the sub-leading muon. In the electron trigger path either the leading electron must have a transverse momentum greater than 23 GeV and the sub-leading a p_T greater than 12 GeV or both the two electrons must have a transverse momentum greater than 33 GeV. The Single lepton trigger (`SingleElectron`, `SingleMu`) requires a single lepton with transverse momentum greater than 27 GeV in the electron case, and 22 GeV or 24 GeV in the muon case. The previous triggers are applied both on simulation (MC) and data. Even if an efficiency close to 100% is measured for both of them a slightly discrepancy between the two is observed. For this reason scale factors defined as the ratio between the efficiency in data and in MC are applied on each simulated event. The trigger efficiencies were computed by the analysis group and estimated as a function of both the transverse momentum and the pseudorapidity of the leptons. For the electron trigger the efficiency is estimated using the tag and probe method [142] whereas for muon trigger the efficiency is computed using the so called "Reference Efficiency" method [143]. The electron trigger efficiency is reported in the Figure 4.7-4.8.

Finally the photon triggers (*SinglePhoton*) and the electron-muon trigger (*MuEG*) are used to fire the events that populate precise phase space regions (control region) used to estimate some background contributions directly from data (data-driven approach). In this particular case no efficiency corrections are estimated. A detailed list of the complete trigger paths is available in the Table 4.4.

Before the identification of the different objects, the event collections are cleaned from all the events produced by beam-gas interaction, beam-halo and calorimeter noises which can contribute to produce fake missing transverse energy. These filters are called E_T^{miss} filters and a detailed description can be found in the following reference pages [144–146].

Primary dataset	Trigger path	Type
DoubleMu	Mu17_TrkIsoVVL_Mu8_TrkIsoVVL Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ	signal
SingleMu	IsoMu24 IsoTkMu24 IsoMu22 IsoTkMu22	signal
DoubleElectron	Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ DoubleEle33_CaloIdL	signal
SingleElectron	Ele27_WPTight_Gsf	signal
MuEG	Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ	control region
SinglePhoton	Photon22_R9Id90_HE10_IsoM Photon22_R9Id90_HE10_Iso40_EBOnly_VBF Photon30_R9Id90_HE10_IsoM Photon36_R9Id90_HE10_IsoM Photon36_R9Id90_HE10_Iso40_EBOnly_VBF Photon50_R9Id90_HE10_IsoM Photon50_R9Id90_HE10_Iso40_EBOnly_VBF Photon75_R9Id90_HE10_IsoM Photon75_R9Id90_HE10_Iso40_EBOnly_VBF Photon90_R9Id90_HE10_IsoM Photon90_R9Id90_HE10_Iso40_EBOnly_VBF Photon120_R9Id90_HE10_IsoM Photon120_R9Id90_HE10_Iso40_EBOnly_VBF Photon165_R9Id90_HE10_IsoM Photon250_NoHE Photon300_NoHE	control region

Table 4.4: Trigger paths used for the analysis and their provenance with respect to the primary datasets. The last column shows the purpose with which the trigger is used: signal - for the search of the signal; control region - to build specific control regions for data-driven background estimation.

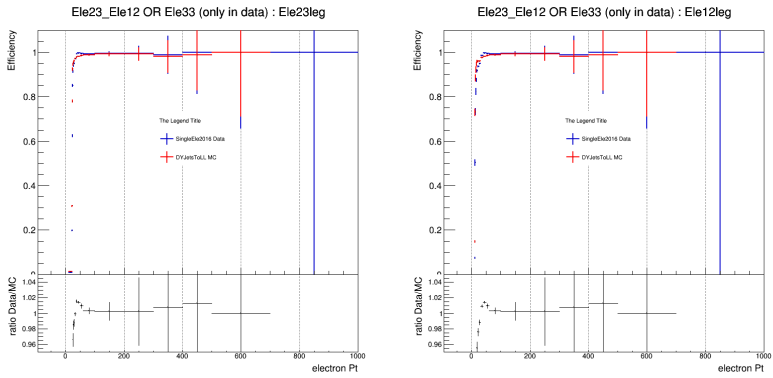

Figure 4.7: Di-electron trigger efficiency measured in data and MC as function of electron p_T. The *left* (*right*) plot correspond to the leading (sub-leading) electron.

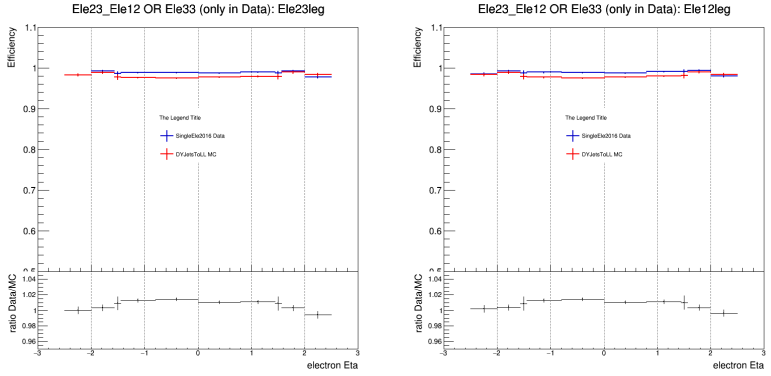

Figure 4.8: Di-electron trigger efficiency measured in data and MC as function of the electron pseudorapidity with $p_T > 25$ GeV. The *left* (*right*) plot corresponds to the leading (sub-leading) electron.

4.3.2 Physics Objects

In this section a detailed description of the reconstructed physics objects used in the analysis is presented. The first objects to be described will be the leptons used to reconstruct the visible Z boson.

Electron Identification

The electrons that will be used to reconstruct the Z boson have to satisfy requirements applied in the offline part of the event selection. Both electrons must have

a $p_T > 25$ GeV and be reconstructed in the tracker acceptance ($|\eta| < 2.5$). The electrons that fall inside the thin uninstrumented transition region between the ECAL barrel and ECAL endcap ($1.4442 < |\eta| < 1.5660$) are not considered in the analysis because of the large uncertainties associated with their identification and energy measurement.

The electrons are required to fulfill the *Tight* identification working point defined as set of cuts for which the identification rate for prompt electrons is about 70% [147]. The identification of an electron is achieved by a set of requirements on the property of its shower and its track. A description of the different identification requirements and the corresponding tight working point is listed below [147]

- the width of the ECAL cluster along the η direction computed in the 5×5 crystal block centered on the highest energy crystal of the seed cluster has to be lower then 0.00998 (0.0292) in the barrel (endcap) [$\sigma_{i\eta i\eta}$].

- the difference in η between the supercluster associated with the electron and its electron track at the interaction point ($\Delta\eta_{in}$) has to be lower then 0.00308 (0.00605) in the barrel (endcap) [$\Delta\eta_{in}$].

- the difference in the $\phi-$plane between the supercluster associated with the electron and its electron track at the interaction point has to be lower then 0.0816 (0.0394) in the barrel (endcap) [$\Delta\phi_{in}$].

- the ratio of the hadronic deposit over the electromagnetic shower has to be lower then 0.0414 (0.0641) in the barrel (endcap) [H/E].

- the absolute value of the difference between the reciprocal of the energy in the calorimeter and the reciprocal of the momentum in the tracker has to be lower then 0.0129 GeV^{-1} both in the barrel and in the endcap to reduce potential electrons from conversion in the calorimeter [$|1/E - 1/p|$].

- to reduce the contamination of fake electrons coming from photons conversion two requirements are imposed on the reconstructed electron

  - the electron is not associated with a conversion vertex[4].

[4]Conversion vertices are identified by the presence of a pairs electron tracks which are nearly collinear at their production vertex [148].

– the maximum number of missing tracker hits is reduced to 1.

To reduce the probability of selecting electrons originating within jets or faked by jets a *Tight* isolation selection is applied. The isolation is computed from the Particle-Flow candidates within a cone of R = 0.3 built around the lepton direction. In the Barrel the electron are considered if their isolation is lower then 0.0588, whereas in the EndCap is it is lower then 0.0571 [147].

The complete set of requirements used in the electron ID are summarized in the table 4.5.

Requirement	Barrel		Endcap	
	Loose	Tight	Loose	Tight
$\sigma_{i\eta i\eta}$	< 0.011	< 0.00998	< 0.0314	< 0.0292
$\Delta\eta_{in}$	< 0.00477	< 0.00308	< 0.00868	< 0.00605
$\Delta\phi_{in}$	< 0.222	< 0.0816	< 0.213	< 0.0394
H/E	< 0.298	< 0.0414	< 0.101	< 0.0641
$\lvert 1/E - 1/p \rvert$	< 0.241	< 0.0129	< 0.14	< 0.0129
Missing hits	= 1	= 1	= 1	= 1
pass conversion veto	yes	yes	yes	yes
PF iso	< 0.0994	< 0.0588	< 0.107	< 0.0571

Table 4.5: Electron identification requirements for the loose and tight selection both in the Barrel and Endcap region.

For every step of the lepton selection, i.e. lepton identification, and isolation, there is a probability to mis-identify a lepton and a probability that a genuine electron is not identified. As mentioned above, the efficiencies of the various steps of the selection have been computed from data using a tag and probe technique. In addition the efficiency in the simulations needs to be corrected in order to match the values measured on the real detector. In order to do so, scale factors defined as respect to the pseudorapidity and transverse momentum of the electron and provided by the CMS group in charge of the electron and gamma reconstruction and identification are used [149].

Muon Identification

Like for the electrons both muons must have a transverse momentum greater than 25 GeV and be reconstructed inside the muon detector acceptance ($|\eta| < 2.4$). To have the best prediction of muon track parameters also in a region with $p_T > 200$ GeV, the *Tracker High-pT* identification is applied. This identification consists of a set of requirements both on the tracker track and on the vertex as reported in the table 4.6.

Requirement	Soft	Loose	Tight	Tracker High-pT
Global		or tracker	true	
PF		true	true	
isTracker				true
track				Nonnull
muonBestTrack (ptError/pt)				<0.3
Muon arbitration	TMOneStationTight			
χ^2/dof (global)			<10	
χ^2/dof (inner)				
Valid pixel hits			>0	> 0
Pixel layers with measurement	>0			
Tracker layers with measurement	>5		>5	>5
Valid muon hits			>0	
Matched muon stations			>1	>1
$d_0(vtx)$ (cm)	0.3		0.2	<0.2
$d_z(vtx)$ (cm)	20		0.5	<0.5
PF iso		< 0.20	< 0.15	
track rel iso		< 0.10	< 0.10	

Table 4.6: Muon identification requirements.

To reduce the probability to select misidentified muons or genuine muons produced in jets a double isolation cut is applied. The isolation is defined using two different recipes both recommended by the collaboration: PF-based isolation (Chap.3) and Tracker-based isolation:

$$I_{\rm rel} = \frac{1}{p_{\rm T}(\mu)} \left[\sum p_T^{Trk(PV)} \right] \tag{4.5}$$

For the PF-based isolation the flux of particles is computed within a cone of R = 0.4, whereas the Tracker-based isolation is defined as the sum of the magnitude of the transverse momentum of all the tracks associated with the primary vertex inside a cone of R = 0.3. In both cases a cleaning procedure is performed: for the PF-based isolation all possible photons irradiated by the muon candidate are excluded, while the Tracker-based isolation is cleaned from all those extra muons that fall inside the isolation cone and are identified as Tracker High-pT muons. Following the recommendation of the CMS group in charge of the muon reconstruction and identification, the tight working point is chosen for the PF-based isolation. For the Tracker-based isolation the loose selection was chosen in order to reduce as much as possible the misidentification of muons inside a jet with prompt muons without compromising the isolation efficiency of the signal sample in the high transverse momentum region.

As for electrons also for prompt muons, there is a chance to not identify a genuine muon and mis-identify a jet as a muon. The efficiencies of the isolation step was computed internally in the analysis group using a tag and probe technique applied on

data. These results was then used to correct the simulation in order to match the values measured on the real detector. For the muon identification efficiencies the corrections to applied to the simulations are borrowed from the following analysis [150] which is using the same muon identification selection of the $ZZ \to 2l2\nu$.

Jet Identification

The jets are one of the most important ingredient in the analysis, indeed they are used to categorize the events, to estimate the missing energy and indirectly the transverse mass. The jets used in the analysis are reconstructed by the Particle-Flow algorithm and further corrected using the scale factors described previously in Chapter 3. The jets with a transverse momentum greater than 15 GeV and a pseudorapidity lower than 4.7, have to pass other quality requirements related to the number of constituents inside the cone and their fraction of electromagnetic or hadronic energy. For the analysis a "PF-loose" identification point was chosen and the associated requirements are reported in the table 4.7.

Id	Variable	Comment
PF-loose	$n >1$ if $\|\eta\| < 2.7$	number of constituents
	$nhf < 0.99$ if $\|\eta\| < 2.7$	neutral hadron energy fraction
	$nef < 0.99$ if $\|\eta\| < 2.7$	neutral electromagnetic energy fraction
	$cef < 0.99$ if $\|\eta\| < 2.4$	charged electromagnetic energy fraction
	$chf > 0$ if $\|\eta\| < 2.4$	charged hadron energy fraction
	$nch > 0$ if $\|\eta\| < 2.4$	charged multiplicity
	$nef > 0.01$ if $2.7 < \|\eta\| <= 3.0$	neutral electromagnetic energy fraction
	$nhf < 0.98$ if $2.7 < \|\eta\| <= 3.0$	neutral hadron energy fraction
	$nnp > 2$ if $2.7 < \|\eta\| <= 3.0$	number of neutral particles
	$nef < 0.90$ if $\|\eta\| > 3.0$	neutral electromagnetic energy fraction
	$nnp > 10$ if $\|\eta\| > 3.0$	number of neutral particles

Table 4.7: Jet id requirements.

The selected jets are expected to be mostly coming from ISR gluons or light-quarks or from the hard process in the VBF process case. To reject events containing at least one b-jets a loose cut on the CVSv2 b-tag variable is used.

Missing Energy and Transverse Mass

The ingredients previously described are used together to define the two most important variables of the entire analysis: the missing energy ($\vec{E}_T^{miss}$) and the transverse

mass (M_T). The first variable is used to purify the signal region from the background contamination, while M_T is used to perform the final statistical analysis. The missing energy used in the analysis is the PF-$\vec{E}_T^{miss,Type-I}$ described in third chapter.

In the context of this analysis the definition of the Transverse Mass reported at the beginning of the Chapter in the equation 4.3 is recasted according to the $X \to ZZ \to 2l2\nu$ decay channel. Indeed the particle one and two in the equation 4.3 are superseded by the Z boson decay in electrons and muons, and by the Z boson decay in neutrinos. The M_T is then defined as

$$M_T^2 = \left[\sqrt{p_{T,ll}^2 + m_{ll}^2} + \sqrt{(E_T^{miss})^2 + m_Z^2}\right]^2 - \left[\vec{p}_{T,ll} + \vec{E}_T^{miss}\right]^2. \quad (4.6)$$

where $\vec{p}_{T,ll}$, $p_{T,ll}$ and m_{ll} identified the kinematics of the visible Z boson. $\vec{E}_T^{miss}$, and E_T^{miss}, and m_Z identify the kinematics of the invisible Z boson. m_Z is the Z pole mass. Both M_T and E_T^{miss} distributions are reported in Fig. 4.22 and Fig. 4.21. The results are obtained after the pre-selection requirements described in the next sub-chapter.

Photon Identification

Although not directly used in the search for the signal, photons are part of a control sample in a data-driven estimate of the instrumental background as it will be described in section 4.4.2. Photons are required to satisfy *tight* identification and *tight* isolation requirements as documented in [151]. The photon is identified if it fulfills a set of basic requirements on the shape of the ECAL supercluster associated with it (for the different variables meaning refers to the previous Electron Identification sub-section): $\sigma_{i\eta,i\eta} < 0.00994$ for the barrel region, and $\sigma_{i\eta,i\eta} < 0.03000$ for the endcap region. $H/E < 0.0269$ for the barrel region and $H/E < 0.0213$ for the endcap region.

Table 4.8 summarizes the photon isolation thresholds used for the tight selection.

Photon events are required to have a single photon selected in the barrel region and the jets in photon events are required to be reconstructed with at least Δ R > 0.4 from the selected photon. Under this condition the photons are required to have a p_T larger then 55 GeV.

4.3.3 Pre-Selection

One of the most important aspects in the analysis is the reduction of the background contamination. This goal is achieved thanks to the following pre-selection requirements:

barrel	
ρ-corrected PF charged hadron	0.202
ρ-corrected PF neutral hadron	0.264 + 0.0148*p_T + 0.000017*p_T^2
ρ-corrected PF photon	2.362 + 0.0047*p_T
endcap	
ρ-corrected PF charged hadron	0.034
ρ-corrected PF neutral hadron	0.586 + 0.0163*p_T+ 0.000014*p_T^2
ρ-corrected PF photon	2.617 + 0.0034*p_T

Table 4.8: Thresholds used for *tight* photon isolation both in the barrel and in the endcap region.

- exactly two leptons satisfying tight identification and isolation requirements with $p_T > 25$ GeV, with invariant mass $\mid M_{l^+l^-} - 91 \mid < 15$ GeV and total transverse momentum $p_{l^+l^-}$ greater than 55 GeV.
- no 3^{rd} lepton must be present in the event. An electron is identified as the third lepton if it satisfies a loose identification requirement with the Isolation lower then 0.15, and p_T > 10 GeV. The muon is identified as a third lepton if it satisfies a loose identification requirement, its isolation is lower then 0.2 and it has a transverse momentum greater then 3 GeV.
- no b-tagged jet has to be present in the event (b-veto cut). A jet is classified as a b-jet if it has a $p_T > 30$ GeV, $\mid \eta \mid < 2.5$, and satisfy the CSVv2 loose identification requirement (For more details about the CSVv2 algorithm refers to the third chapter of this thesis).
- to reduce possible contamination of fake E_T^{miss} due to a significant mis-measurement of one or more jet energies the events are discarded if the absolute value of the difference in the ϕ plane between the jets and the E_T^{miss} is lower then 0.5 ($\mid \Delta\phi(jet, E_T^{miss}) \mid$).
- events with E_T^{miss} lower than 125 GeV are discarded.
- when E_T^{miss} is greater than 125 GeV, all the events where the difference in the ϕ plane between the Z boson reconstructed with the two visible leptons selected in the analysis workflow and the E_T^{miss} is lower than 0.5 are discarded. Some more details about this cut will be presented in the background section ($\mid \Delta\phi(Z, E_T^{miss}) \mid$).

The flow of the events through the pre-selection filter is summarized in the plots 4.9. It is evident that the most important requirements are the ones on the p_T^Z and on

the E_T^{miss}. All the other variables used in the pre-selection are reported in the figure 4.10. All the backgrounds at this level are estimated using pure Monte Carlo and present a good agreement with data.

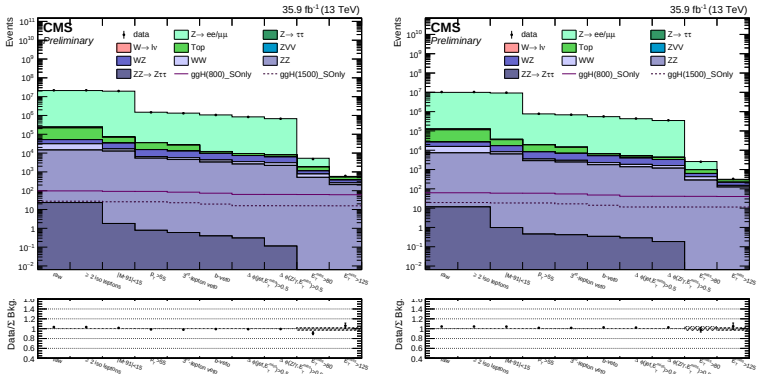

Figure 4.9: Number of events passing each pre-selection cuts in the di-muon (*left*) and di-electron (*right*) sample MC based. Reference signal distributions for m_H = 800 GeV and 1500 GeV are superimposed to both the distributions.

4.3.4 Jet-bin categorization

After the pre-selection cuts, the events are categorized in three different groups: the exclusive 0-jet category, the inclusive 1-jet category, and finally the VBF category. After ranking the jets by decreasing transverse momentum, the exclusive 0-jet category is selected if the event doesn't have any jet with p_T greater than 30 GeV, whereas the inclusive 1-jet category is chosen if the event present at least on jet with $p_T > 30$ GeV without requiring any further requirements. The VBF category is selected via some kinematical requirements on the two leading jets (j_1, j_2) in the event:

- at least two jet with p_T greater than 30 GeV
- the distance in the pseudorapidity plane between the two leading jets ($\Delta\eta_{j_1,j_2}$) has to be greater than 4.
- the invariant mass of the leading and the sub-leading jets has to be greater then 500 GeV
- the pseudorapidity of the Z boson (η_Z) reconstructed by the two visible leptons selected by the analysis workflow needs to satisfy the following relation: $\eta_{j_1} < \eta_Z < \eta_{j_2}$ where η_{j_1} and η_{j_2} are the pseudorapidity values associated the two leading jets in the event

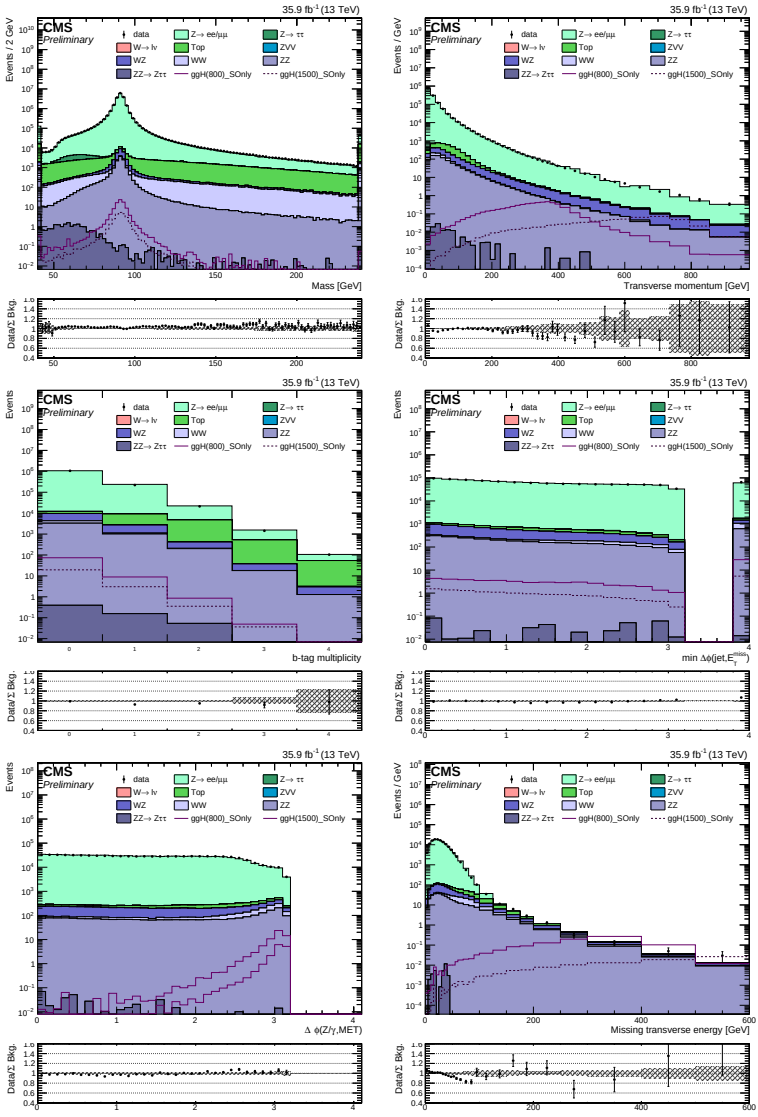

Figure 4.10: Distributions showing the variables used in the pre-selection in the inclusive category for the $\mu\mu$ channel. Reference signal distributions for m_H = 800 GeV and 1500 GeV are superimposed to all the distributions. The overflow bin in the $\mid \Delta\phi(jet, E_T^{miss}) \mid$ is coming from all the events that present no jets. All the shapes are obtained before the corresponding selection cut is applied. The invariant mass is obtained directly from the leptons that pass the identification and isolation selection, whereas the Z p_T spectrum is produced considering only the leptons that pass the Z mass window requirements. The b-jet multiplicity distribution is obtained considering the events which don't present a third lepton in the final state, whereas $\mid \Delta\phi(jet, E_T^{miss}) \mid$ and $\mid \Delta\phi(Z, E_T^{miss}) \mid$ after the b-tag veto and the cut on the minimum ϕ-angle between the jet and the missing transverse energy. Finally, the E_T^{miss} is obtained at the end after all the other selection requirements.

- no jet ($\bar{j}$) with $p_T > 30$ GeV and $\eta_{j_1} < \eta_{\bar{j}} < \eta_{j_2}$ has to be present in the event

4.4 Background Estimation

All the events that survive to the last cut of the pre-selection (E_T^{miss} > 125 GeV) can be sorted between the signal events and the background events. In particular, the background can be grouped in two main big family: irreducible and reducible background. The irreducible background consists of those processes that have the same final state as the signal searched for. Production of ZZ and ZWW belong to this category. The WZ and the other ZVV (V=Z,W) processes, in which the charged lepton from the W decay goes out of acceptance or remains undetected, are also considered as being part of this category. All the other processes, like Z+jets production, are instead called in the following "reducible background". Each background has its own estimation strategy but two core procedure can be highlighted already at this point: the irreducible background is estimated using simulations, whereas the reducible one using a data-driven approach. The next lines will be devoted to make clear all the peculiarities of each background estimation.

4.4.1 Irreducible Background

The irreducible background is composed by all the physical processes characterized by a final state that is the same as that of the signal. For this reason its contamination, in general, can't be reduced significantly by simply applying some kinematical cuts. In addition to the ZZ process, the WZ also belongs to this category of backgrounds in the case the charge boson decays hadronically and the jets are misreconstructed and produce fake E_T^{miss} or the W decay leptonically and the lepton either falls outside the acceptance region or it doesn't satisfy the third lepton veto requirement in the analysis. The tri-boson processes (ZVV) have a lot of possibilities to mimic the signal final state but due to their low cross section which is two order of magnitude lower then WW, and 2 to 5 times lower then ZZ, their contribution is small.

This source of background is totally predicted using pure NLO Monte Carlo corrected taking into account both QCD and electroweak (EWK) corrections (K-factors) applied at a generator level [152]. For $qq \rightarrow ZZ \rightarrow 2l2\nu$ two K-factors are considered: one to take into account the EWK NLO corrections and the other one to take into account the NNLO corrections in QCD. For the WZ process only the EWK corrections are considered.

For ZZ, the EWK corrections come in three distinct categories: emission of a real γ/Z/W boson, emission of a parton due to the presence of photons in the initial state (γ-induced process, for a schematic representation refers to the Figure 4.11), and finally virtual corrections with a γ, Z or W in the loop. However, only virtual corrections are important for the process [153, 154]. The corrections are implemented on a basis of the table given by the authors of Ref. [153, 154], and they are computed as a function of the Mandelstam variables, $\widehat{s}$ and $\widehat{t}$, in the center-of-mass frame and

the flavor of the initial quarks which initiated the hard process. The computation has shown that the corrections are negative and dependent on both the m_{ZZ} and $p_{T,Z}$, in particular the corrections increase in module with the increase of m_{ZZ} as presented in the fig.4.12.

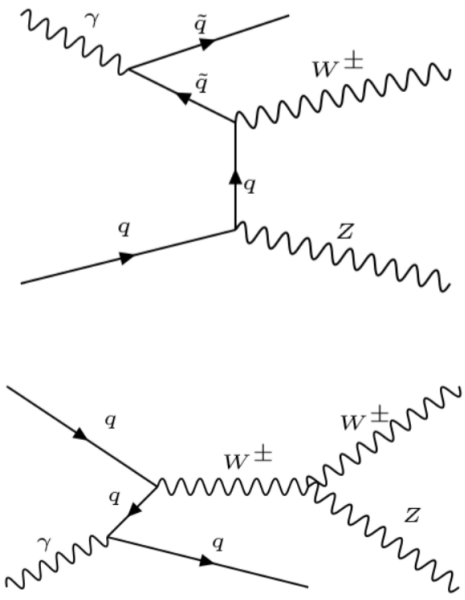

Figure 4.11: Examples of diagrams for the γ-induced processes in $pp \to WZ + X$ [152].

The electroweak corrections to WZ come in two different contributions: virtual and photon-induced. The virtual part gives a negative contribution and it is computed using the same approach for ZZ, whereas the γ-induced corrections are positive and of the same order of magnitude of the virtual part. They are estimated generating the events with MadGraph5_aMC@NLO and the LUXqed photon PDF [155] and cross-checked with the EW corrections calculated in MadGraph5_aMC@NLO framework [131, 156, 157]. The final K-factor is obtained multiplying the partial K-factors together. The final result is reported in the Fig. 4.13.

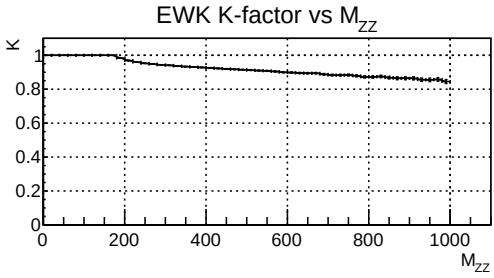

Figure 4.12: Evolution of the NLO electroweak corrections K-factor on $pp \to ZZ$ as a function of m_{ZZ} at a generator level. EW corrections are expected to be small at low energies, for Z-pair production a shift of about 4% is observed close to threshold, whereas the corresponding corrections are below 1% for lower energies. [153, 154]

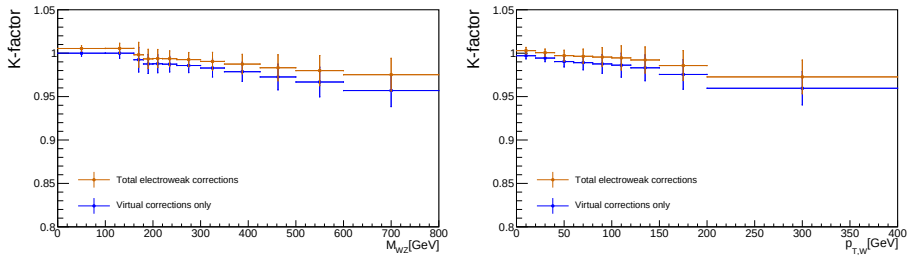

Figure 4.13: Evolution of the NLO electroweak corrections K-factor on $pp \to W^+Z$ as a function of m_{WZ} (left) and $p_{T,W}$ (right), using only virtual corrections (blue) or using both virtual and photon-induced corrections (brown) [152].

4.4.2 Reducible Background

Non-resonant Background

All the reducible background processes that present two charged leptons in the final state coming from a non-resonant particle, are grouped in this background (NRB) which represents the second contribution to the total yield. The non-resonant background includes top production (single top or top pairs), WW, WWW and W+jets in the case the jet is misidentified as a lepton. The two most important cuts which reduce considerably their contamination in the final yields are the b-veto and the Z mass window cut. The events that survive are estimated using a data-driven approach called α-method.

This approach estimates the contamination of NRB di-leptons events in the signal region by rescaling the NRB $e\mu$ events that fall in the same signal kinematical region by the α-scale factor as reported in the equation 4.7 where N_{ll}^{NRB} quantifies

the amount of non-resonant same flavor events in the signal region, and $N_{e\mu}^{in}$ identifies the number of NRB opposite flavor events which populate a kinematical region defined by two opposite flavor and the kinematical cuts of the signal region. The equation 4.8 defines the α-scale factor: N_{ll}^{out} ($N_{e\mu}^{out}$) is the number of NRB same-flavor (opposite-flavor) events that populate a given sideband region.

$$N_{ll}^{NRB} = \alpha_l \cdot N_{e\mu}^{in} \tag{4.7}$$

$$\alpha_l = \frac{N_{ll}^{out}}{N_{e\mu}^{out}} \tag{4.8}$$

The sideband region is characterized by the same requirements of the pre-selection except for the mass window and the veto on the presence of b jets in the event. Indeed the mass window in this case is 40 GeV $< M_{ll} <$ 70 GeV and 110 GeV $< M_{ll} <$ 200 GeV and at least a single b jet is required in the control region in order to reduce further the contamination of DY events. It should be noted that this method suffers in principle from the contamination of events coming from WZ, WWZ and the part of Z and ZZ events involving leptonically decaying taus in the three control samples, i.e $N_{e\mu}^{in}$, N_{ll}^{out} and $N_{e\mu}^{out}$. It has however been proved that their contribution, as reported in figure 4.14, is totally negligible with respect to that of the events including top quarks and doesn't introduce a bias nor increase significantly the systematic uncertainty of the method.

The α value is measured as function of different E_T^{miss} thresholds and its dependency is reported in Figure 4.15. Eighteen different E_T^{miss} values are tested from 50 GeV up to 135 GeV with step of 5 GeV. The relation is fairly flat, even if it tends to be larger if the E_T^{miss} threshold is lower because of a bigger contamination of Drell-Yan events in the same sideband region.

The expected composition of the sidebands region and the prediction obtained from it ("Predicted" column) is presented in Table 4.9. The composition is shown incrementally for the two main processes expected to dominate the sidebands (Top and WW) and when all the simulated processes are considered. The rightmost column is the ratio between estimated yields using the data-driven method described above and the yields expected from MC. The relatively good agreement (usually within 5%) represents a successful closure test of the method.

Further closure tests of this method were performed by comparing the inclusive (40 GeV $< m_{ll} <$ 70 GeV and 110 GeV $< m_{ll} <$ 200 GeV) and the high-mass sidebands (110 GeV $< m_{ll} <$ 200 GeV) in both the ee and $\mu\mu$ flavor channels. Using the same values of the E_T^{miss} thresholds adopted to study the α-value dependence

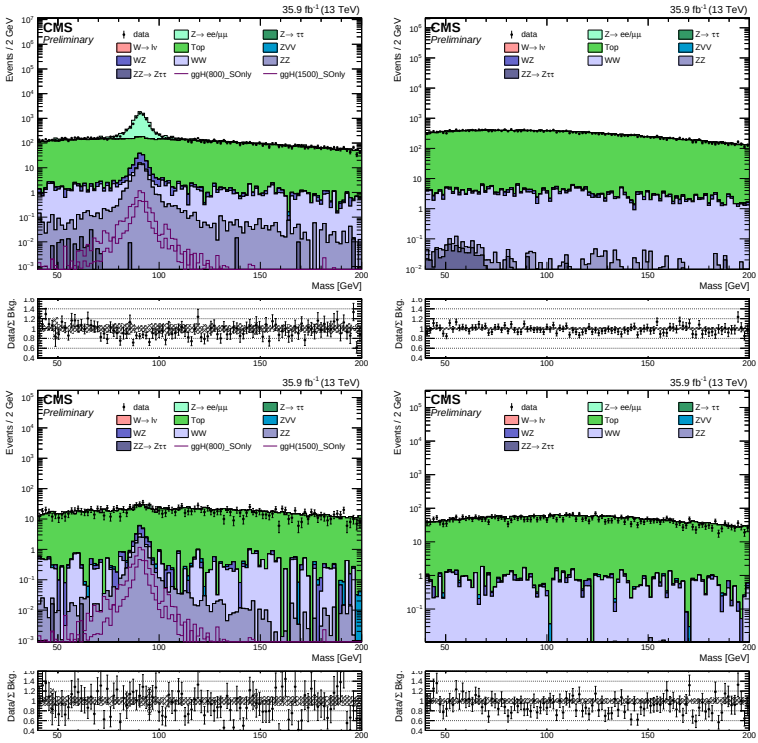

Figure 4.14: Number of events populating the invariant mass spectrum in the di-electron (*left*) and muon-electron (*right*) channel in the final state. Two different requirements on the E_T^{miss} are applied: E_T^{miss} greater than 50 GeV (*top*) and greater than 125 GeV (*bottom*).

from the E_T^{miss}, the number of events predicted by the α-method is compared to the expectations from MC simulations. The bias is defined as:

$$bias = \left(\frac{N_{predicted}}{N_{expected}} - 1 \right). \tag{4.9}$$

The results of the closure test comparison are presented in the Figure 4.16 and they show a good agreement in the performance between the side band region and the up sideband region.

The table 4.10 summarizes the final values of α used in the analysis as computed from data, and, for reference, as obtained from simulation reporting both the sideband region and the up sideband region. Thanks to the flat dependence of the α-value respect to the E_T^{miss} threshold (Fig. 4.15), the final α-values can be chosen in order to reduce the total uncertainties on the method reported in Figure 4.17. For this reason,

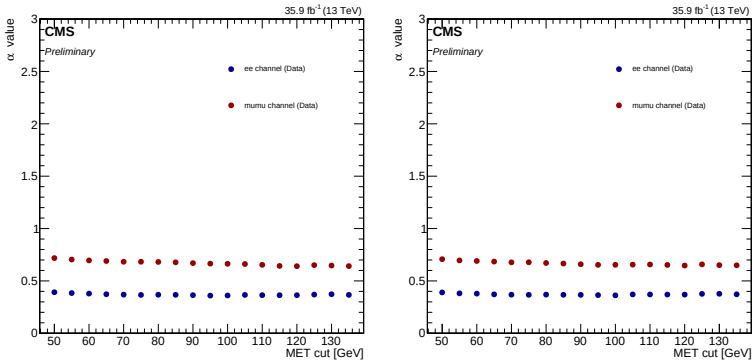

Figure 4.15: α value as a function of the different requirements on the missing transverse energy using Data samples in di-electron (*red*) and di-muon (*blue*) channel. The dependance of α on the E_T^{miss} is flat. The value is computed using both the sideband region (40 GeV $< m_{ll} <$ 70 GeV and 110 GeV $< m_{ll} <$ 200 GeV) (*left*) and the up sideband region (110 GeV $< m_{ll} <$ 200 GeV) (*right*). In both the cases, at least one b jet in the event is required.

b-tagged sample, 40 GeV $< m_{\ell\ell} <$70 GeV and 110 GeV $< m_{\ell\ell} <$200 GeV							
ee	$N_{e\mu}^{in}$	$N_{e\mu}^{out}$	N_{ee}^{out}	α	$N_{ee}^{SR,est}$	$N_{ee}^{SR,exp}$	$f_{est/exp}$
Top	219.5 ± 5.2	14185.6 ± 40.7	5216.9 ± 24.4	0.368 ± 0.002	80.7 ± 2.0	76.7 ± 3.0	1.05 ± 0.05
Top+WW	292.5 ± 6.6	14333.2 ± 41.1	5270.7 ± 24.6	0.368 ± 0.002	107.5 ± 2.5	103.7 ± 3.8	1.04 ± 0.05
All MC	300.1 ± 7.4	14395.2 ± 44.0	5290.3 ± 26.2	0.368 ± 0.002	110.3 ± 2.8	103.7 ± 3.8	1.06 ± 0.05
$\mu\mu$	$N_{e\mu}^{in}$	$N_{e\mu}^{out}$	$N_{\mu\mu}^{out}$	α	$N_{\mu\mu}^{SR,est}$	$N_{\mu\mu}^{SR,exp}$	$f_{est/exp}$
Top	219.5 ± 5.2	14185.6 ± 40.7	9708.6 ± 34.2	0.684 ± 0.003	150.2 ± 3.6	153.9 ± 4.3	0.98 ± 0.04
Top+WW	292.5 ± 6.6	14333.2 ± 41.1	9806.8 ± 34.5	0.684 ± 0.003	200.1 ± 4.6	204.0 ± 5.5	0.98 ± 0.03
All MC	300.1 ± 7.4	14395.2 ± 44.0	9856.9 ± 37.3	0.685 ± 0.003	205.5 ± 5.1	204.0 ± 5.5	1.01 ± 0.04
b-tagged sample, 110 GeV $< m_{\ell\ell} <$200 GeV							
ee	$N_{e\mu}^{in}$	$N_{e\mu}^{out}$	N_{ee}^{out}	α	$N_{ee}^{SR,est}$	$N_{ee}^{SR,exp}$	$f_{est/exp}$
Top	219.5 ± 5.2	9479.3 ± 33.4	3552.1 ± 20.2	0.375 ± 0.003	82.2 ± 2.0	76.7 ± 3.0	1.07 ± 0.05
Top+WW	292.5 ± 6.6	9579.5 ± 33.7	3588.2 ± 20.4	0.375 ± 0.003	109.5 ± 2.6	103.7 ± 3.8	1.06 ± 0.05
All MC	300.1 ± 7.4	9596.7 ± 34.4	3598.8 ± 20.5	0.375 ± 0.003	112.5 ± 2.9	103.7 ± 3.8	1.09 ± 0.05
$\mu\mu$	$N_{e\mu}^{in}$	$N_{e\mu}^{out}$	$N_{\mu\mu}^{out}$	α	$N_{\mu\mu}^{SR,est}$	$N_{\mu\mu}^{SR,exp}$	$f_{est/exp}$
Top	219.5 ± 5.2	9479.3 ± 33.4	6354.7 ± 27.7	0.670 ± 0.004	147.1 ± 3.6	153.9 ± 4.3	0.96 ± 0.04
Top+WW	292.5 ± 6.6	9579.5 ± 33.7	6419.1 ± 28.0	0.670 ± 0.004	196.0 ± 4.5	204.0 ± 5.5	0.96 ± 0.03
All MC	300.1 ± 7.4	9596.7 ± 34.4	6431.2 ± 28.0	0.670 ± 0.004	201.1 ± 5.1	204.0 ± 5.5	0.99 ± 0.04

Table 4.9: Closure test using the sideband regions 40 GeV $< m_{ll} <$ 70 GeV and 110 GeV $< m_{ll} <$ 200 GeV (*top*) and up sideband region 110 GeV $< m_{ll} <$ 200 GeV (*bottom*) to compute α for E_T^{miss} >70 GeV. The results are obtained for 13 TeV. The "est" label stands for estimated and "exp" for expected.

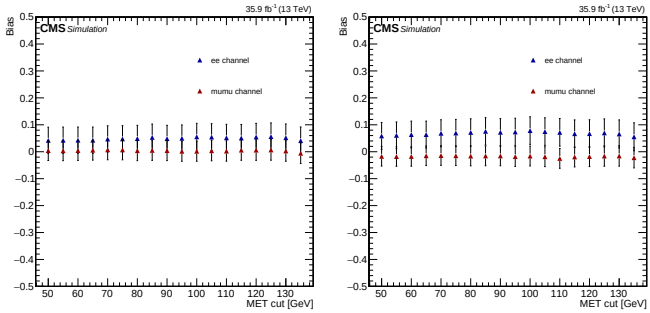

Figure 4.16: Closure test for different E_T^{miss} requirements. The distributions are reported for both the sideband region (*left*) and the up sideband region (*right*) requiring the presence of at least a b jet in the event.

the sideband region and a E_T^{miss} cut of 70 GeV are chosen. The corresponding α value are: 0.37 ± 0.01 for the electron channel and and 0.68 ± 0.01 for the muon channel. The Table 4.11, reports the actual event numbers in the control region and the final estimated events in data using the previous values of α. The right most column shows the ratio between the estimated NRB events and the MC predicted events. One could see that the estimated event numbers are consistent with MC prediction.

Channel		b-tag, all side band	b-tag, right side band
ee	Data	0.369 ± 0.006	0.369 ± 0.007
	MC	0.368 ± 0.002	0.375 ± 0.003
$\mu\mu$	Data	0.683 ± 0.009	0.677 ± 0.011
	MC	0.685 ± 0.003	0.670 ± 0.004

Table 4.10: α computation with the sideband region and the up sideband region. In both cases, at least one b jet is required in the event.

Channel	$N_{e\mu}^{in}$	$N_{e\mu}^{out}$	N_{ee}^{out}	α_{Data}	$N_{ee}^{SR,est}$	$N_{ee}^{SR,exp}$	$f_{est/exp}$
ee	298.0 ± 17.3	13970.0 ± 118.2	5156.0 ± 71.8	0.369 ± 0.006	110.0 ± 6.6	107.7 ± 4.2	1.02 ± 0.07
$\mu\mu$	298.0 ± 17.3	13970.0 ± 118.2	9543.0 ± 97.7	0.683 ± 0.009	203.6 ± 12.1	208.4 ± 5.5	0.98 ± 0.06

Table 4.11: Event numbers in the control region and final estimated Non-Resonant background events in data.

Although the results for different E_T^{miss} requirements are correlated, given they partially use the same subset of events, no bias larger than 13% is observed at the value of the missing transverse energy cut used in the analysis at 125 GeV. This value

is assumed to be the uncertainty in the background prediction of the method and it is estimated taking into account three different contributions: the uncertainties associated to the closure test, relative difference on the alpha value with different E_T^{miss} cut in data and finally the statistical error on α. The plot 4.17 gives an overview of the total uncertainties as a function of the missing transverse energy both in the $\mu\mu$ channel and in the ee channel.

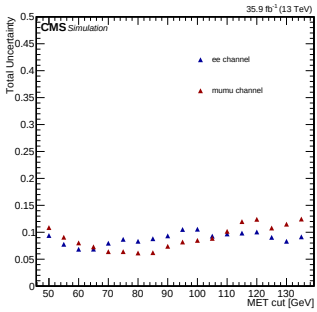

Figure 4.17: Total uncertainties as a function of the E_T^{miss}. The uncertainties is defined as the sum in quadrature of the following uncertainties: closure test, relative difference on alpha value with different E_T^{miss} requirements in data, and stats error on α.

$Z + jets$

The process $pp \to Z+j \to l^+l^-j$ $(l = e, \mu)$ doesn't present any genuine E_T^{miss}. However events from this process are present in any missing transverse energy distribution inside an experiment. The reason for this is the jet energy mis-measurement. This effect affects the low missing transverse energy region distribution and is also well known to be not properly modeled by the simulation. The background from $Z + jets$ events is estimated using a data-driven approach based on a control sample containing single, isolated photons. Single, isolated photons are generated through the same processes as Z bosons and are expected to preserve similar kinematics. The size of the $\gamma + jets$ sample is significantly high which reduces the statistical uncertainty associated with the estimate of the background yield obtained with this method. The procedure adopted is the following:

1. Select events fired by the single photon triggers reported in the table 4.4 and satisfying the pre-selection requirements as in the dilepton signal region with the exception of the Z mass window constrain. The events are then categorized according to the number of jets in the final state.

2. Reweight the photon p_T spectrum in order to match the p_T spectrum of electrons and muons produced by a resonant Z boson in each jet bin category. Both the leptons are selected from the double and single lepton datasets reported in the table 4.1 requiring tight identification and isolation constrains, and a $p_T > 25$ GeV. In order to select electrons and muons produced by a resonant Z boson the following requirement on their final invariant mass is applied: $\mid M_{l^+l^-} - M_{Zpole} \mid < 15$ GeV. All these requisites allow to select from data a sample of Drell-Yan characterized by a high-purity, indeed the contributions from other background processes used in the analysis is lower then 1% for the inclusive selection, and of the order of the percentage in the jet bin category[5]. The comparison of the dilepton and photon spectrum before the reweighting is reported in the Figure 4.18. For illustrative purposes the inclusive one jet bin category is reported. The exclusive 0 jet and the VBF categories present the same features of the inclusive one jet category and are not reported. The photon weights in the one jet category are reported in Figure 4.18. Also in this case only the inclusive 1 jet category is reported, but in the exclusive 0 jet category and in the VBF categories the weights are very similar to those computed in the 1 jet category.

3. Generation of the Z boson mass for the reweighted photon events in order to compute the transverse mass. The mass for each reweighted photon event is estimated by choosing randomly a value from the Z mass spectrum predicted by the MC Drell-Yen and reconstructed by two leptons (l^+, l^-) satisfying tight identification and isolation requirements, with a p_T greater than 25 GeV, and $\mid M_{l^+l^-} - M_{Zpole} \mid < 15$ GeV. This procedure is performed in each jet bin category.

4. After this re-weighting, the distributions of interest, i.e. E_T^{miss} and M_T are computed from the photon sample and used as models for $Z + jets$.

5. MC samples which present a genuine missing transverse energy are subtracted from reweighted photon data. As it was described above the missing transverse energy in $Z + jets$ is fake and it is caused by the jets energy mis-measurement. However the photon sample receives contributions also from the processes containing real E_T^{miss} in the final state. Such processes refer to $Z + \gamma \rightarrow \nu\nu\gamma$, $W + \gamma \rightarrow l\nu\gamma$, $W \rightarrow l\nu$, $Z + \gamma \rightarrow l^+l^-\gamma$, $Z \rightarrow \nu\nu$, $t + \gamma$, and finally $t\bar{t} + \gamma$. The contribution of these processes has to be taken into account in the estimation of the photon p_T spectrum. In order to do so, these processes predicted by MC and reweighted by the same weights used to rescale the photon transverse momentum, are subtracted from reweighted photon data.

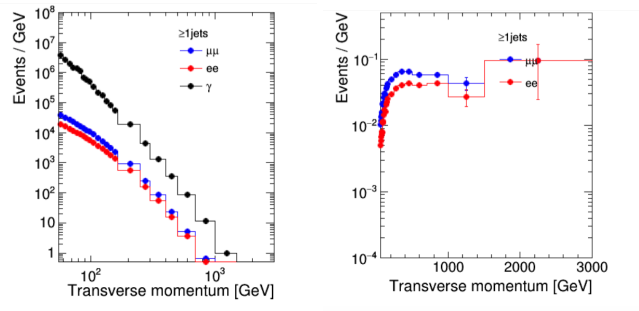

Figure 4.18: (*left*) p_T spectra of events passing the analysis in the ee, $\mu\mu$ and γ channels in the inclusive 1 jet category. (*right*) Photon weights used to reweight the photon control sample in the inclusive 1 jet category.

In order to evaluate the validity of the method, a closure test is performed by applying the steps described above to simulated photons and Drell-Yen samples. The resulting distributions are shown in Figure 4.19 for the E_T^{miss}, M_T and $|\ \Delta\phi(jet, E_T^{miss})\ |$ kinematic variables. In all the figures beside the comparison between the MC Z boson p_T spectrum and the MC $\gamma + jets$ reweighted p_T spectrum, the ratio between the two shapes is reported. The agreement between the two predictions is good except in high missing transverse energy regions which are affected by low statistic.

As mentioned during the description of the Pre-selection step, the Figure 4.20 shows that the prediction of the $|\Delta\phi(Z, E_T^{miss})|$ distribution appears to mismodel the data in the region below 0.5 when the missing energy is required to be greater than 125 GeV. This discrepancy is more evident in the di-muon channel. For this reason the events presenting a value lower then 0.5 are discarded.

The goodness of the data-driven approach in the estimation of the Drell-Yan process can be appreciated in the Figures 4.21-4.22 where a comparison between the NLO MC predictions and the data-driven method is presented in both the E_T^{miss} shape and the transverse mass distribution. In both the cases, the reweighted $\gamma + jets$ sample is able to cure the visible discrepancy between data and background which is present in the low region of both the E_T^{miss} and the M_T when the $Z + jets$ background is fully estimated by the simulations. The discrepancy between data and simulations which is present in the E_T^{miss} distribution is less evident in the M_T variable thanks to the mitigation effects due to the good agreement between data and MC in both the Z boson p_T and mass spectrum.

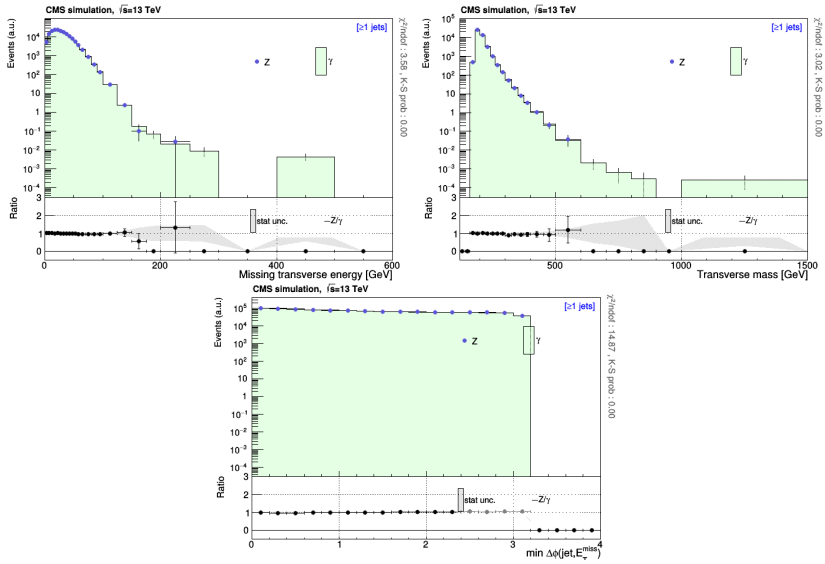

Figure 4.19: Closure test of the γ+jets method for the E_T^{miss} (*top-left*), E_T (*top-right*), and $\mid \Delta\phi(jet, E_T^{miss}) \mid$ (*bottom*) variable using MC samples. The gray distributions correspond to the DY sample and the points to the γ sample after the reweighting procedure described in the text. As an illustrative example only the inclusive 1 jet category is considered. The plots on the bottom show the bias estimated for the prediction obtained from the γ+jets sample.

[5]The selection requirements applied on data is applied in all the CMS analysis measuring Drell-Yan properties or cross-section [158]

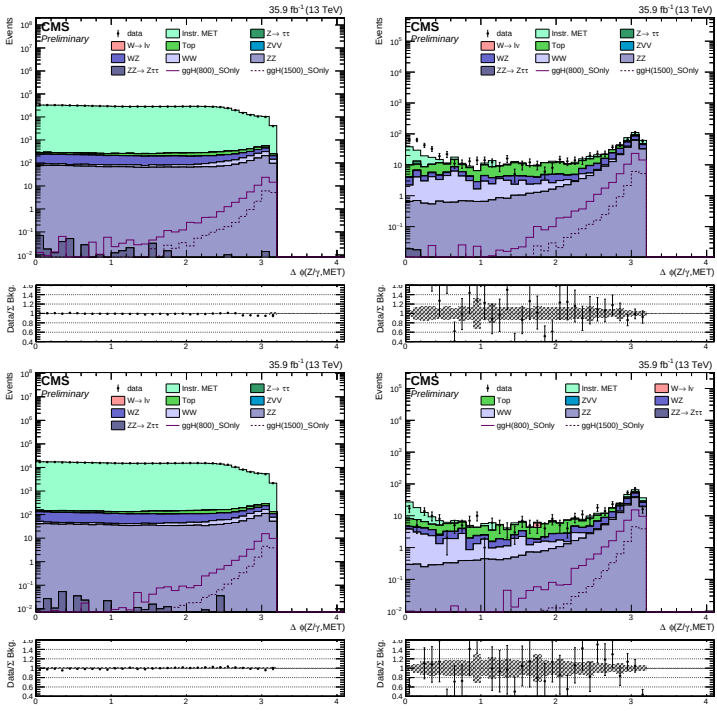

Figure 4.20: $\mid \Delta\phi(Z, E_T^{miss}) \mid$ variable in di-muon channel (*top*) and di-electron channel (*bottom*) for $E_T^{miss} > 0$ GeV (*left*) and greater then 125 (*right*). In both the cases the events used in both the distributions have passed all the pre-selection requirements up to the $\mid \Delta\phi(jet, E_T^{miss}) \mid$ cut.

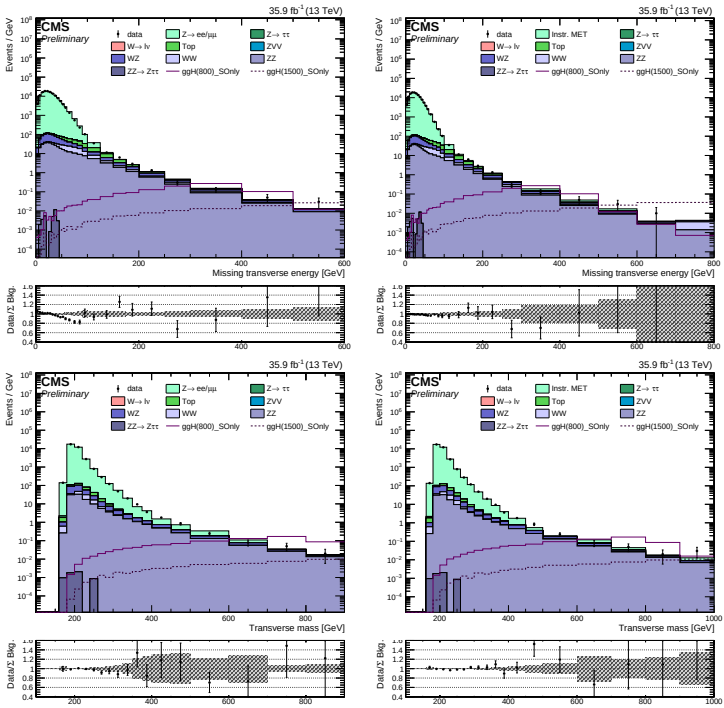

Figure 4.21: Comparison of the background prediction using MC samples (*right*) and the data-driven approach (*left*) applied for the Drell-Yan process in the $\mu\mu$ channel looking the E_T^{miss} (*top*) and the M_T (*bottom*) variable. Both the distributions are obtained using events that have passed all the pre-selection requirements except for the missing transverse energy cut.

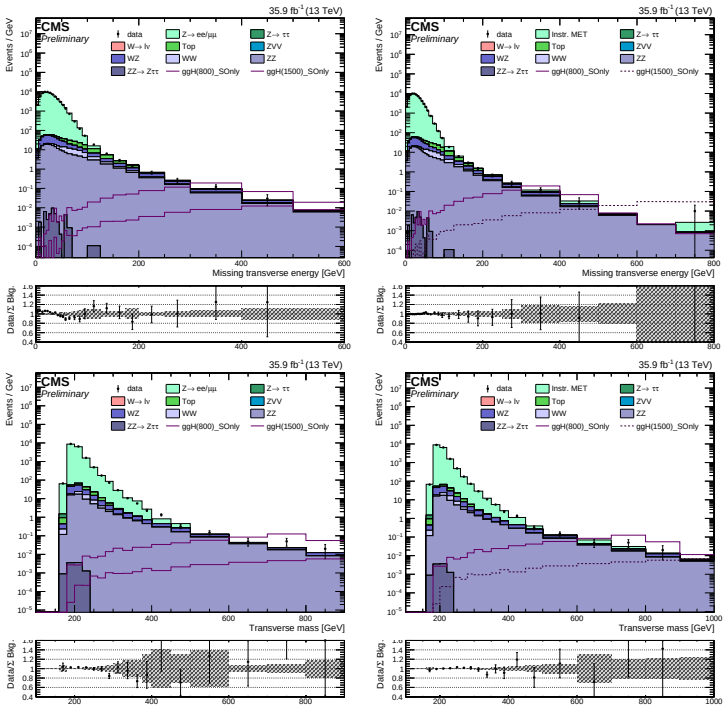

Figure 4.22: Comparison of the background prediction using MC samples (*right*) and the data-driven approach (*left*) applied for the Drell-Yan process in the ee channel looking the E_T^{miss} (*top*) and the M_T (*bottom*) variable. Both the distributions are obtained using events that have passed all the pre-selection requirements except for the missing transverse energy cut.

4.5 Systematic Uncertainties

The last important ingredient in the analysis of data is the estimation of the uncertainties associated with the final results. The sources of the uncertainty are categorized in two big groups: instrumental uncertainties and theoretical uncertainties. The first one group is related to the detector performance, the calibration techniques and the reconstruction. The second group is related to the modeling of the considered physics processes. This chapter is totally devoted to the description of these topics.

4.5.1 Statistical Uncertainty of the Simulated Sample

The limited number of simulated events induces a statistical uncertainty that contributes to the uncertainty estimated from such samples. This contribution is computed for each background of the analysis and each signal hypothesis directly of the M_T shape, considering both the flavor and jet bin categories. The impact[6] of this systematic is computed independently bin-by-bin varying up and down the nominal bin yield by the square root of its content directly on the M_T shape.

4.5.2 Theoretical Uncertainties

The theoretical uncertainties are estimated only for the MC simulations (signal samples, $ZZ \to 2l2\nu$, $WZ \to 3l\nu$) and include the uncertainties related to the PDFs and the α_S value used in the MC production, the uncertainties related to the missing higher order contributions in the estimation of the cross-section of the samples and finally the uncertainty associated with the QCD and EWK corrections used to improve the simulation predictions.

Systematic uncertainty in $gg \to ZZ$ K-factor

A normalization uncertainty of 10% is applied to the assumed cross section for the $gg \to ZZ$ continuum to account for possible differences between the K-factor computed for $gg \to ZZ$ and the exact $gg \to ZZ$ result. This error is derived from renormalization and factorization scale variations [128].

PDF and α_S uncertainties

All the background and the signal samples are produced centrally using the NNPDF3.0 PDFs. Following the recommendations of the PDF4LHC group [159], the impact of the PDF and the α_S uncertainty is initially estimated for each event and

[6] The impact of a systematic is defined as the variation of the nominal M_T shape in the case the ingredient concerned is modified by the corresponding uncertainty

successively propagated to the M_T shape. It's impact is estimated to be between 1% and 4% according to the flavor and the jet bin category.

Scale uncertainties

This uncertainty in the renormalization (μ_R) and factorization (μ_F) scales quantifies the impact on the the M_T shape of the QCD missing higher order contributions to the cross section of the MC samples. It is estimated for both the background samples ($ZZ \to 2l2\nu$, $WZ \to 3l\nu$, and ZVV background) and the signal samples. It is accounted for by a normalization uncertainty that varies according to the jet bin category. The uncertainty is computed varying the normalization and the factorization scale up and down, or both in the same direction, by a factor 2 or 1/2 and considering the combination which returns the cross section with the biggest shift from the nominal cross section value. This previous procedure is applied for the inclusive 1 jet category and the VBF category. Scale uncertainties on exclusive jet categories as the 0-jet category in the analysis, are under-estimated using this method because of the compensation of two opposite effects: collinear radiations and higher-order QCD effects. To account for this, the recipe proposed in [160] is used and the uncertainty in the 0 jet category is given by:

$$\Delta\sigma^{0j} = \sqrt{\left(\Delta\sigma^{\geq 0j}\right)^2 + \left(\Delta\sigma^{\geq 1j}\right)^2}.$$

In the particular case of the signal where its cross-section is the physical quantity measured by the analysis, the normalization uncertainties computed with the scale variation have to take into account only the jet bin migration between the different categories. In order to do so, the new cross-section with the modified μ_R and μ_F in each jet bin category is normalized by the ratio between the total nominal cross-section and the total new one obtained by changing both the factorization and the normalization scale.

The theoretical uncertainties for $ZZ \to 2l2\nu$ and $WZ \to 3l\nu$ at 13 TeV are shown in Table 4.12.

Electroweak uncertainties

The uncertainty associated to the NLO electroweak corrections on the background samples is estimated for each event availing the corresponding generated information. This uncertainty error is then propagated to the M_T shape through the event in order to generate a "shape up" and "shape down" (Fig. 4.23). The uncertainty associated to each event are estimated according to the event recoil [154], defined as $\rho \equiv \frac{\left|\sum_{i=1}^{4} \vec{p}_T^{\,i}\right|}{\sum_{i=1}^{4}\left|\vec{p}_T^{\,i}\right|}$, with i running on the leptons coming from each vector bosons particle. The relative uncertainty in the corrected cross section is then estimated as follows:

Table 4.12: Scale (μ_R and μ_F) uncertainties (in %) for ZZ, WZ and ggH production cross sections at 13 TeV. For the signal case three mass points are considered: 800 GeV, 1 TeV, and finally 2 TeV.

Process	Uncertainty per category (%)		
	=0 jets	$\geq$1 jets	VBF
ZZ	6.3	5.4	40
WZ	9.5	5.1	40
ggH($M_X = 800$ GeV)	3.2	3.1	6.0
ggH($M_X = 1$ TeV)	3.4	2.8	5.3
ggH($M_X = 2$ TeV)	1.7	1.2	0.1

$$\delta = \begin{cases} \left|\left(1 - K_{QCD}^{NLO}\right)\left(1 - K_{EWK}^{NLO}\right)\right|, & \text{if } \rho < 0.3 \\ \left|1 - K_{EWK}^{NLO}\right|, & \text{if } \rho \geq 0.3 \end{cases}$$

where the K_{QCD}^{NLO}-factor is taken from [161]. The choice of uncertainty for $\rho < 0.3$ is motivated by the fact that uncertainties are coming mainly from missing diagrams in $\alpha\alpha_S$. The two contributions (NLO QCD and NLO EWK) are fully correlated and then they are linearly combined. The case $\rho \geq 0.3$ has a maximal error associated to it, i.e. 100% of the electroweak correction. This case represents about 25% of the events. The Figure 4.23 shows the distribution of the nominal transverse mass after applying the EWK corrections with $\pm$ 1 σ uncertainty following the recipe previously described.

In WZ, the uncertainty associated to the photon-induced process is related to the photon PDF error. However, this contribution is very small (less than 1 % in the whole spectrum) and it can be neglected.

4.5.3 Instrumental Uncertainties

Luminosity

The uncertainty in the total luminosity delivered by CMS in 2016 data is estimated to be 2.5% [162]. It affects the signal and the MC background yields which need to be normalized to the measured luminosity.

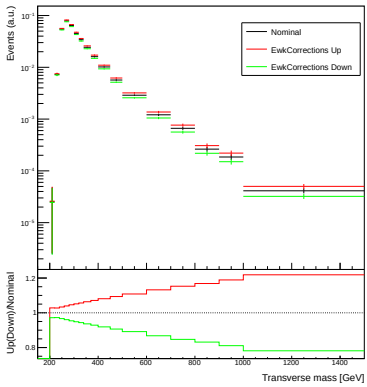

Figure 4.23: Distribution of the transverse mass for the nominal, upper and lower values of the electroweak corrections for $pp \to ZZ$.

Lepton Trigger, Lepton Identification and Lepton Isolation

The systematic uncertainty arising from the lepton selection is accounted for by a normalization uncertainty. The total uncertainty varies between 6% and 8% , depending on the lepton flavor, and it is obtained by combining in quadrature the uncertainty associated to the trigger scale factors, the identification scale factors and finally the isolation scale factors.

Lepton Momentum Scale

The impact of the leptons momentum scale correction uncertainties is estimated by shifting up and down the nominal energy of each lepton in the event by the uncertainty associated to the scale corrections following the recommendations of the CMS group in charge of the muon and electron identification [149, 163]. The shift is then propagated to the transverse mass. The uncertainty in the lepton energy scale is 0.01% for muons and 0.3% for electrons.

Jet Energy Scale, Jet Energy Resolution and Unclustered MET

The uncertainties associated to the simulated jet energy scale corrections are dependent both on the jet p_T and the jet η and cover the range between 8% for low jet momentum region and 1.5% for very energetic jets (1 TeV) [118]. The effects of this uncertainties on the M_T shape is estimated for each event shifting the correction factor up and down by 1σ respect to the nominal value. This variation is then finally propagated to the M_T shape and its impact is measured to be between 1% up to 40% , depending on the jet bin category.

The effects of the unclustered energy[7] on the transverse mass are considered and they are estimated shifting the transverse energy deposit according to its energy and η position following the prescription of the CMS group in charge of the missing transverse energy reconstruction [165]. The associated uncertainties on the M_T do not exceed the 10%

b Jet Veto

The scale factors (SF) associated to b-tagging and mistagging efficiency in simulated events are provided as a function of the transverse momentum and the pseudorapidity of the b-jet. To estimate the impact of the SF uncertainty, the nominal SF value is varied up and down according to its uncertainty. The uncertainties are p_T-dependent and vary from less then 1% up to 3% for the lower p_T region ($p_T < 30$ GeV). The variations are performed separately for heavy flavor jets and for light jets and then propagated to the M_T shape. The final effect of the b-tagging uncertainty is estimated to be between 2% and 6%.

Non-Resonant Background

Three kinds of systematic uncertainties are applied to the Non-Resonant Background. The systematic uncertainty in the background estimate stems from the uncertainties associated to the closure test, the relative difference on the alpha value with different E_T^{miss} cut in data and finally the statistical error on α. A total uncertainty of 13% is assigned to the method in both the flavor channel (Fig. 4.17) and it is accounted as a normalization uncertainty.

$Z + jets$ Background

Three different shape systematic uncertainties are considered for the $Z + jets$ background. The first uncertainty is related to the statistical error associated to both the γ data sample and the MC samples used to predict the genuine E_T^{miss} contribution. This uncertainty is estimated directly on the M_T shape combining in quadrature the statistical variation bin-by-bin of each of them.

A second contribution comes from the degree of agreement between the reweigthed $\gamma + jets$ prediction on the Drell-Yan. The uncertainty is computed in each jet bin category looking to the relative difference on the estimated events in the two different samples. A value of 10% is associated to it.

The last systematic takes into account the precision of the cross-section computation of the $Z + \gamma$, $W + \gamma$ and $W + jets$ processes used to estimate the genuine

[7] The unclustered energy refers to the energy associated to jets with a $p_T < 10$ GeV plus the contribution from the PF objects which were not clustered in any jet [164].

MET contribution. Both the QCD and EWK corrections are considered for the exception of $W + jets$ where no EWK corrections are applied. The numbers reported in the literature [166–169] are used on the M_T shape in order to determine a global up and a global down shape variation. The biggest uncertainty is related to the QCD corrections on the $Z + \gamma$ sample and it amounts to 25%.

Finally the uncertainty associated to the closure test and the uncertainty related to the missing higher order corrections to the cross section are combined in quadrature bin-by-bin to estimate a global final shape up and a global final shape down variation.

Theoretical Uncertainty [%]	
NNLO ($gg \to ZZ$) K factor	10
PDF + α_S	1-4
Renorm./Factor. Scales	0.1-40
EW Corrections (*)	7-40
Experimental Uncertainty [%]	
Integrated Luminosity	2.5
Lepton Trigger, Identification and Isolation	6-8
Lepton Momentum Scale (*)	0.01-0.3
Jet Energy Scale, Jet Energy Resolution (*)	1-40
Unclustered E_T^{miss} (*)	10
b Tagging (*)	2-6
Data Driven Background [%]	
Z + jets (*)	20-50
Non Resonant Background	13

Table 4.13: Summary of the different sources of uncertainties used in this analysis given in percent. The numbers shown as ranges represent the uncertainties in different final states or categories. Most uncertainties affect the normalizations of the background estimations or simulated event yields, and those that affect the shape of kinematic distributions as well are labeled with (*).

4.6 Results

The search for a scalar resonance X in the $ZZ \to 2l2\nu$ channel is performed in the mass range between 200 GeV up to 3 TeV considering three different width values 5, 10 and 100 GeV.

The final yields after the entire selection is reported in the table 4.14. Each number is accompanied by the statistical uncertainty. For the combination of the different jet bins category the most important background is the irreducible ZZ process, followed by the non-resonant contribution, WZ and Drell-Yen are comparable whereas the ZVV is totally negligible. The zero jet category is totally dominated by ZZ, instead the 1-jet category is contaminated by the non-resonant background because of the presence of the top quark in the process. The VBF category is characterized by lower total yields and is dominated by the reducible backgrounds. The final M_T distributions used to set the limits are presented in Fig. 4.25 and Fig. 4.24 showing two mass points in gluon fusion.

As can be seen from both the Yield table (Tab 4.14) and the transverse mass shape (Fig. 4.25 and Fig. 4.24) no significance excess is measured over the SM background. For this reason the analysis has proceeded setting an upper limits on the $pp \to X \to ZZ$ cross section as a function of m_X and its width (Γ_X) exploiting the transverse mass shape (Fig. 4.25 - 4.24). The modified frequentist method called as CL_s described in [170] is used and the test statistic chosen is the profile likelihood modified for upper limits [171]. The limits are set in both the ggF and in the VBF separately. In the likelihood the number of events observed in each bin of the M_T distributions is compared with the predictions from background and signal. The different observations are combined by means of a product of Poisson probabilities. The predictions are subject to the multiple uncertainties described in the previous section and each associated with a so-called nuisance parameter that represents the source of the uncertainty. The effect of the uncertainty in each nuisance parameter is modeled by providing two modified distributions of M_T: one corresponding to an upward variation by one standard deviation of the central value of the nuisance parameter and a second corresponding to a downward variation of the nuisance parameter by the same value. Systematic uncertainties in signal and background normalizations are modeled with log-normal distributions whereas the shape uncertainties are treated by means of morphing procedures [172]. The final exclusion cross section at the 95% confidence level (CL_s) is reported in the Figure 4.26.

In order to avoid the usage of negative bin yield associated to the interference between the signal and the background and properly handle rescaling of the interference contribution by the $\sqrt{\mu}$, the parametrization of the expected event yield (ν) used

channel	Inc.	$\mu\mu = 0jets$	$\mu\mu \geq 1jets$	$\mu\mu vbf$	$ee = 0jets$	$ee \geq 1jets$	$eevbf$
ZZ	$\mathbf{332.8 \pm 1.0}$	113.2 ± 0.6	94.9 ± 0.5	1.27 ± 0.06	67.4 ± 0.4	55.3 ± 0.4	0.74 ± 0.04
WZ	$\mathbf{176.5 \pm 4.0}$	39.1 ± 1.9	69.6 ± 2.5	1.5 ± 0.3	22.0 ± 1.4	43.5 ± 2.0	0.86 ± 0.24
ZVV	$\mathbf{6.8 \pm 0.4}$	0.45 ± 0.10	3.8 ± 0.3	0.06 ± 0.03	0.22 ± 0.05	2.1 ± 0.2	0.04 ± 0.03
Instr. MET	$\mathbf{123.0 \pm 18.3}$	38.0 ± 9.93	38.0 ± 11.9	3.44 ± 1.93	21.2 ± 5.62	20.7 ± 7.59	1.73 ± 1.13
Top/W/WW	$\mathbf{313.5 \pm 13.3}$	18.4 ± 3.5	181.7 ± 11.2	3.4 ± 0.7	10.0 ± 1.9	98.2 ± 6.0	1.8 ± 0.4
total	$\mathbf{952.7 \pm 22.99}$	$\mathbf{209.2 \pm 10.72}$	$\mathbf{388.0 \pm 16.54}$	$\mathbf{9.7 \pm 2.076}$	$\mathbf{120.8 \pm 6.108}$	$\mathbf{219.8 \pm 9.88}$	$\mathbf{5.2 \pm 1.22}$
data	**938**	**194**	**405**	**11**	**115**	**209**	**4**
ggF(800)	$\mathbf{(212.2 \pm 6.9) \times 10^1}$	561.9 ± 36.0	697.0 ± 40.5	21.3 ± 3.8	349.6 ± 24.3	478.2 ± 34.5	14.2 ± 2.2
VBF(800)	$\mathbf{36.7 \pm 0.2}$	2.40 ± 0.04	11.71 ± 0.09	7.84 ± 0.07	1.64 ± 0.03	7.85 ± 0.07	5.26 ± 0.06
ggF(1000)	$\mathbf{(168.2 \pm 3.7) \times 10^1}$	408.6 ± 17.8	581.3 ± 21.8	23.5 ± 4.3	254.6 ± 14.2	399.0 ± 17.9	14.5 ± 2.8
VBF(1000)	$\mathbf{49.6 \pm 0.2}$	3.58 ± 0.07	15.3 ± 0.1	10.4 ± 0.1	2.36 ± 0.05	10.7 ± 0.1	7.25 ± 0.09
ggF(2500)	$\mathbf{46.0 \pm 1.7}$	7.6 ± 0.7	16.9 ± 0.9	0.98 ± 0.22	6.0 ± 0.6	13.9 ± 0.9	0.65 ± 0.15
VBF(2500)	$\mathbf{22.1 \pm 0.1}$	2.02 ± 0.04	6.66 ± 0.07	3.90 ± 0.06	1.57 ± 0.04	5.09 ± 0.07	2.85 ± 0.05

Table 4.14: Expected number of signal and background events for an integrated luminosity of 35.9 fb^{-1} at 13 TeV after applying the full selection requirements compared to the number of observed events in data. Uncertainties for ZZ/WZ/Z+Jets and non-resonant backgrounds as well as for signal include the statistical uncertainties.

to define the likelihood model [171] associated to a shape analysis is customized in the following way [173]:

$$\begin{aligned} \nu_i = (\mu - \sqrt{\mu}) \cdot S_i(\vec{\theta}) + \\ + \sqrt{\mu} \cdot (S_i(\vec{\theta}) + B_i(\vec{\theta}) + I_i^{S-B}(\vec{\theta})) \\ + (1 - \sqrt{\mu}) \cdot B_i(\vec{\theta}) + \tilde{B}_i(\vec{\theta}) \end{aligned} \tag{4.10}$$

where the multiplicative parameter μ, called *signal strenght* quantifies the strenght of the observed signal. S_i, B_i and I_i^{S-B} are the expected yields in the ith bin of the M_T shape respectively for the signal hypothesis (S_i), the background hypothesis (B_i) defined as the combination of the continuum sample and the light higgs, and finally the interference between the signal and the background (I_i^{S-B}). $\tilde{B}_i$ quantifies the yield in the ith bin predicted by the background described in the section 1.4.

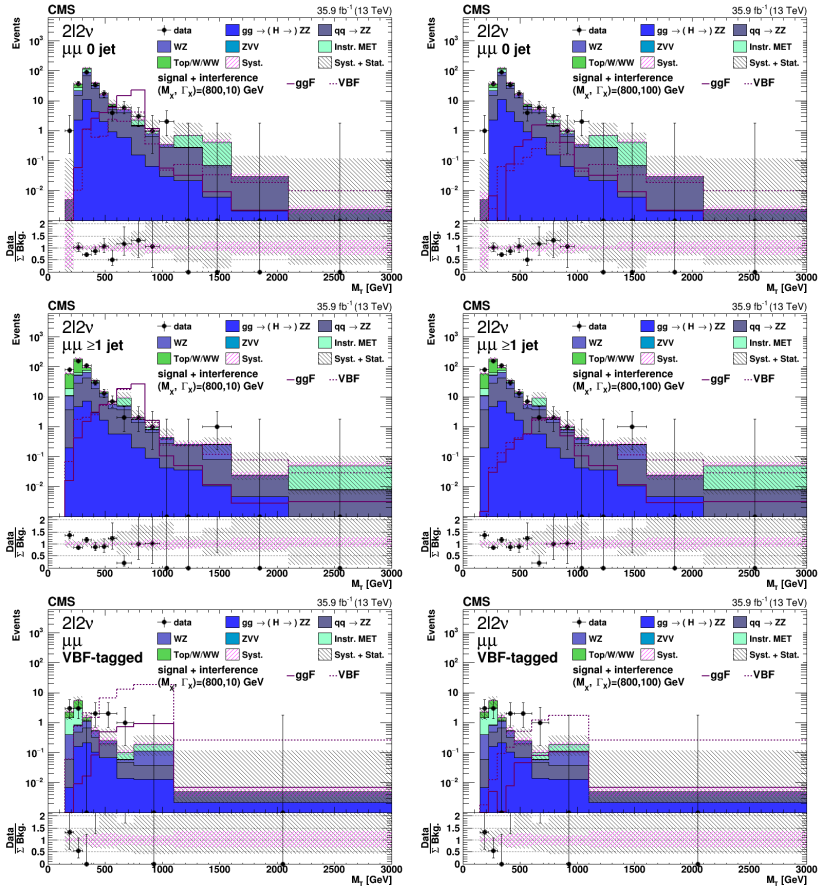

Figure 4.24: Transverse mass distributions in the different jet bins category after the final E_T^{miss} cut of 125 GeV in $\mu\mu$ channel. The signal distribution corresponds to a mass point of 800 GeV. Two width hypothesis were reported: 10 GeV (*left*) and 100 GeV (*right*) both in ggF and VBF. In each distribution both systematic and statistical errors are represented. The distribution on the bottom of the transverse mass reports the discrepancies between the data and the total background.

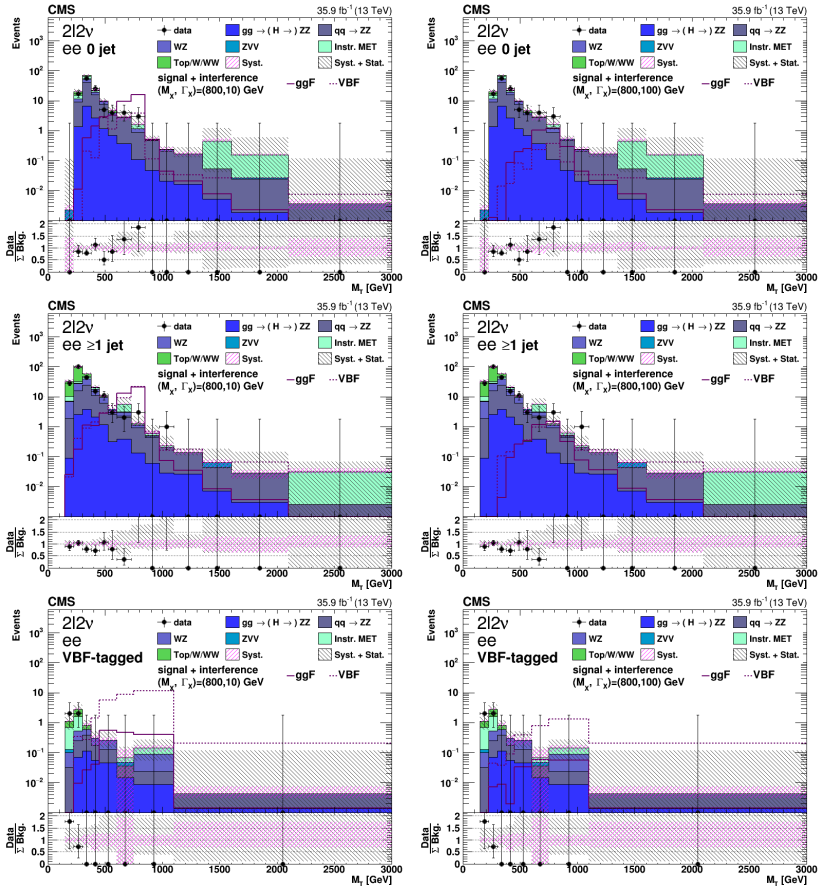

Figure 4.25: Transverse mass distributions in the different jet bins category after the final E_T^{miss} cut of 125 GeV in *ee* channel. The signal distribution corresponds to a mass point of 800 GeV. Two width hypothesis are reported: 10 GeV (*left*) and 100 GeV (*right*) both in ggF and VBF. In each distribution both systematic and statistical errors are represented. The distribution on the bottom of the transverse mass reports the discrepancies between the data and the total background.

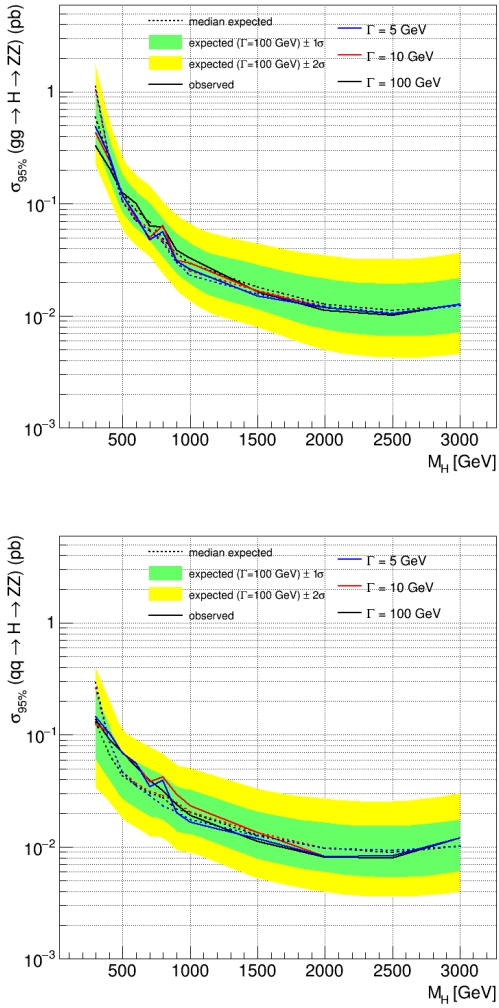

Figure 4.26: Upper limits at 95% CL set on the ggF (*left*) and VBF (*right*) cross section of a generic scalar boson as function of its mass for various values of width. Note that these limits are set on the $X \to ZZ$ total cross-section. In the x and y axis the heavy scalar resonance is indicated by H.

Chapter 5

The Dynamic Truncation Algorithm for Muon Reconstruction

Preserving good performance in the reconstruction of high-p_T ($p_T > 300$ GeV) muons is crucial especially for all the analysis targeting a high-mass resonance with muons in its final state. These muons suffer from catastrophic energy losses when traversing dense matter, which makes their reconstruction challenging. CMS has introduced different strategies to face this problem. One of them is the Dynamic Truncation Algorithm (DYT) and its development is one of the main topics of the present thesis. The first section of this chapter describes an overview of the radiation-matter interaction followed by a description of the CMS strategy to deal with this problem. The second part of the chapter will be devoted to describe the DYT algorithm and its performances.

5.1 High-Energy Muon Reconstruction

The standard muon reconstruction described in the third Chapter is well suited for all the muons with a momentum lower then 200 GeV. For very energetic muons, however, there is a high probability to produce electromagnetic showers in the calorimeters or in the iron of the CMS magnet return joke. This effect causes a significant degradation of the reconstruction because on one side the shower can leak in the muon detectors and compromise the local reconstruction and on the other the amount of energy lost in the shower process is considerable. The presence of many analysis looking for very massive or boosted objects decaying into muons has pushed the collaboration to develop new approaches which deal efficiently with this situation. Before examining the different approaches in depth, a brief summary of the interaction radiation-matter is provided.

5.1.1 Material Effects

There are five main processes characterizing the interaction of high-energy muons with matter: ionization, multiple scattering, electron-positron pair production, inelastic interaction with nuclei and bremsstrahlung. While the first two are present at all muon energies, the last three, which are called radiative processes, intervene only beyond 200 GeV, as shown in Figure 5.1 [174].

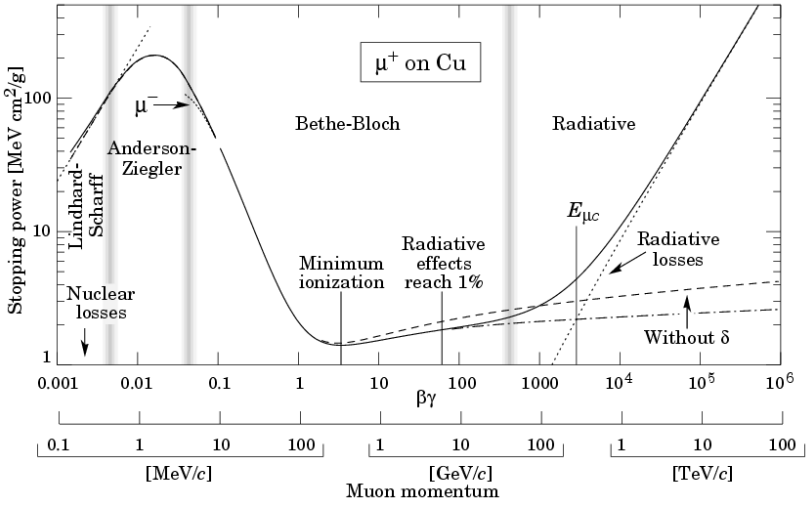

Figure 5.1: Stopping power for muon in copper as a function of its momentum [175].

These radiative processes are characterized by small cross sections, hard spectra, large energy fluctuations, and the associated generation of electromagnetic and hadronic showers. At these energies the treatment of energy loss as a uniform and continuos process is for the purpose of track reconstruction inadequate. It is convenient to write the average rate of muon energy loss as,

$$-dE/dx = a(E) + b(E)E. \tag{5.1}$$

In the equation above, $a(E)$ is the ionization energy loss given by the Bethe-Bloch equation, and $b(E)$ is the sum of electron-positron pair production, bremsstrahlung, and photonuclear contributions [175]. To the approximation that these slowly varying functions are constant, the mean range x_0 of a muon with initial energy E_0 is given by [175],

$$x_0 \approx (1/b)\ln(1 + E_0/E_c) \tag{5.2}$$

The *critical energy*, E_c, is defined as the energy at which radiative and ionization losses are equal and from Eq. 5.1 it is given by $E_c = a/b$. In the particular case of iron, which is relevant for the CMS detector, the critical energy occurs at $E_c \approx 300$ GeV. Below the critical energy ionization losses dominate, whereas above it radiative effects dominate. Fig. 5.2 shows the contributions to $b(E)$ for iron. Since $a(E) \approx 0.002$ GeV g^{-1} cm^2, $b(E)E$ dominates the energy loss above several hundred of GeV, where $b(E)$ is nearly constant. The rates of energy loss for muons in hydrogen, uranium, and iron are shown in Fig. 5.3.

The radiative cross sections can be expressed as functions of the fractional energy loss ν. With this choice the bremsstrahlung cross section goes roughly as $1/\nu$ over most of the range, while for the pair production case the distribution goes as ν^{-3} to ν^{-2} [175]. The probability for a muon to produce electrons with different energies through iron is shown in the Fig. 5.4. As reported in the figure the electron production cross-section rise with the muon momentum especially for the pair production case. This process gives the largest contribution to the electron production for electrons with low energy, whereas the Bremsstrahlung is the dominant contribution for very high energetic electrons. Indeed, considering electrons with energy close to 1 GeV, if one muon of 1 TeV traverses a meter of iron it will produce on average about 0.09 such electrons via ionization, about 0.01 via bremsstrahlung and finally about 0.5 via pair production. Instead, considering high energetic electrons with an average value of about 100 GeV, the biggest contribution comes from the bremsstrahlung with about $3{\cdot}10^{-3}$ electrons whereas the amount of electrons produced by both the pair production and the ionization are a factor 10 smaller.

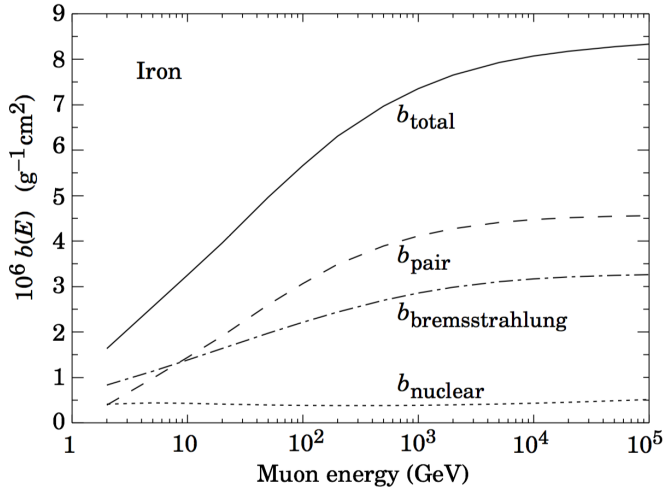

Figure 5.2: Contributions to the fractional energy loss by muons in iron due to electron-positron pair production, bremsstrahlung, and photonuclear interactions as a function of the muon energy [175].

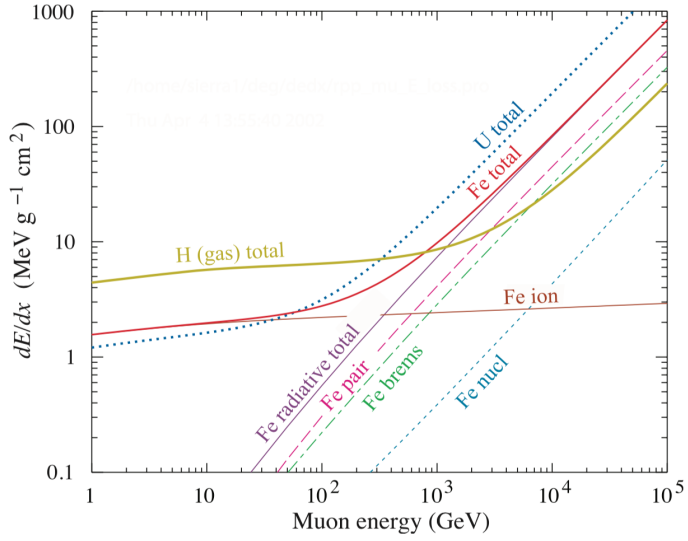

Figure 5.3: The average energy loss of a muon in hydrogen, iron, and uranium as a function of muon energy [175].

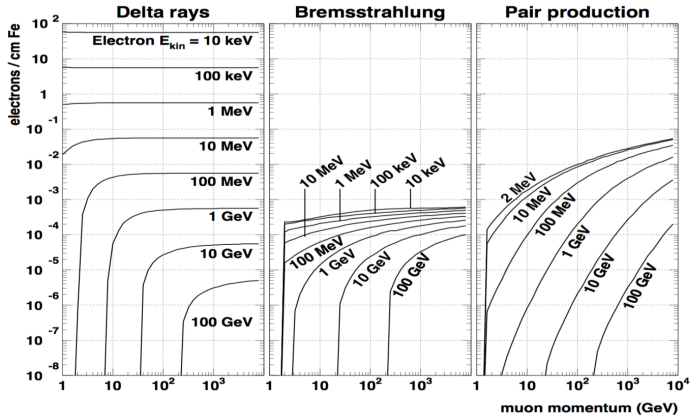

Figure 5.4: Cross sections to produce an electron with the same energy reported in the distribution by muons traversing matter. The cross section is shown as number of electrons produced per 1 cm of iron, as a function of muon transverse momentum [176].

Figure 5.5 shows the distribution of the final momentum of 1 TeV muons after having traversed three meters of iron. The most probable loss is 8 GeV whereas the effects of the radiative processes are visible in the left tail.

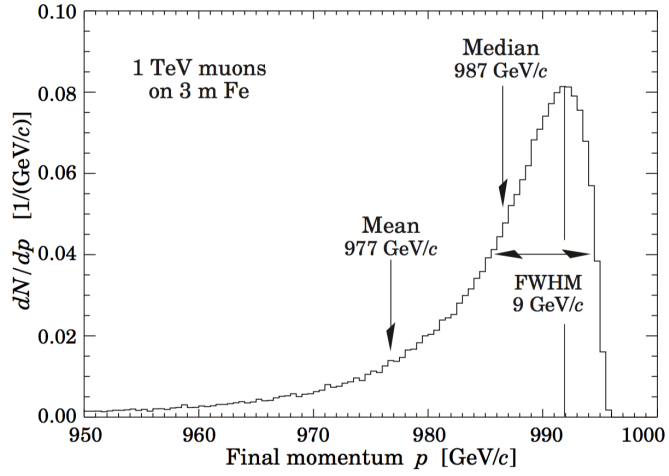

Figure 5.5: Distribution of the final momentum of 1 TeV muons after having traversed three meters of iron. The most probable loss is 8 GeV whereas the effects of the radiative processes are visible in the left tail [175].

Focusing now on the case of the CMS detector, all the muons produced at the collision point need to cross different sub-detectors before reaching the muon spectrometer. Indeed, as described in the second chapter of this thesis, the muon chambers are preceded by the tracker, the calorimeters and finally the magnet. If the tracker has a very little impact on the energy loss of the muons due to the thickness of the silicon detector, the rest of the material which separates the tracker from the solenoid can lead to the production of showers due to radiative losses. Indeed, the electromagnetic calorimeter is composed by crystals of $PbWO_4$ 23 cm long characterized by a radiation length (X_0) of 0.89 cm corresponding to a total radiation length of 25.8 cm. In addition, the muon critical energy in lead tungstate is measured to be about 160 GeV [177]. The HCAL is made of brass interposed between two layers of stainless steel. Its total absorber thickness varies from 5.8 λ_I ($\lambda_I = 16.42$ cm) when η = 0 up to 10 λ_I in the endcap, which can be translated in 63.8 radiation length ($X_0 = 1.49$ cm) in the barrel and about 110 X_0 in the endcap. The iron joke which encapsulates the muon chambers is made of iron characterized by a radiation length of 1.8 cm ($E_C \sim 300$ GeV). In the barrel, the innermost iron layer starts at R = 4590 mm and the outermost (the third one) ends at R = 7000 mm. The thickness of the first layer is 295 mm, whereas the thickness of the second and third layers is 630 mm [178]. Both endcap yokes are composed of three disks with a diameter of 13.9 m. The two innermost disks are 600 mm thick and the outermost 300 mm. The inner disk is separated from the adjacent barrel ring by 600 mm, and all disks are equally spaced by 600 mm to accommodate the endcap muon chambers. Attached to the innermost disk, and facing the interaction point, there is a solid iron piece 5.4 m in diameter and 700 mm thick called the nose. A global view of the material thickness in radiation length of the CMS detector is reported in Fig. 5.6.

The impact of the material in front of the muon spectrometer can be tracked studying the frequency at which the muon station is used in the reconstruction of the muon track performed by the Dynamic Truncation algorithm (A more detailed description of the algorithm will presented in the next sections). The results are reported in Fig. 5.7 for both the barrel and the endcap case as a function of the nominal muon momentum. In both the barrel and the endcap the frequency at which the first station is not used increase with the muon momentum. In particular it is skipped 25-30% of the times for muon at 400 GeV whereas at 3 TeV the first station is skipped 35-40% of the times. No differences emerge between the barrel and the endcap because the CMS material thickness in radiation lengths is similar in the first station of the muon chambers as reported in Fig. 5.6.

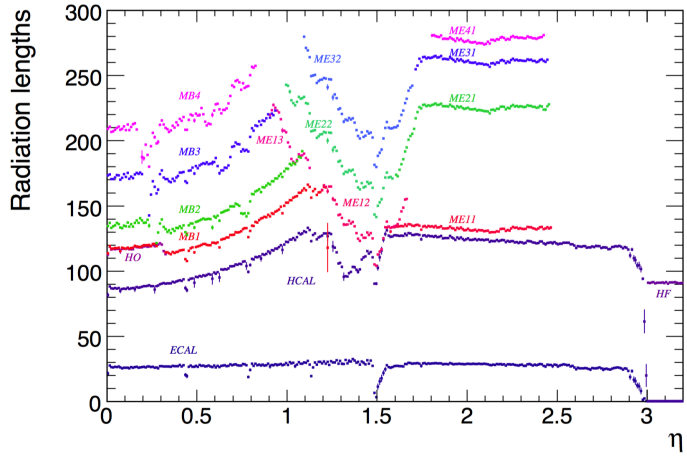

Figure 5.6: Material thickness in radiation lengths after the ECAL, HCAL, and at the depth of each muon station as a function of pseudorapidity. [95].

5.1.2 High-p_T Muon refitters

As extensively explained in the previous section, muons with energy greater than 300 GeV and crossing layers of iron have a significant chance to loose a big amount of their energy. These large energy losses can degrade the performance of the muon reconstruction for the two following main reasons:

- the muon can loose a large fraction of its energy and its momentum changes accordingly. Contrary to standard ionization and multiple scattering, radiative losses cannot be modeled as continuous processes and therefore are completely unaccounted in the Kalman Filter algorithm used to fit tracks. As a consequence the Kalman Filter estimate of the momentum of a muon track that has suffered a severe energy loss will be highly biased.
- the shower produced by the products of a radiative process can traverse the muon detectors. Segment reconstruction will be rendered highly complicated by the large number of hits recorded in the detectors and therefore also an incorrect trajectory parameters estimation will result.

Starting from the considerations above two possible strategies can be developed in order to identify a large radiative energy loss: either looking for an incompatibility between the extrapolated track states and the segments in the muon chamber or looking for high occupancy in the muon chambers.

Three specialized and complementary algorithms, generically called *Tev-refitter* algorithms, have been developed to face this problem. All the Tev-refitters have as

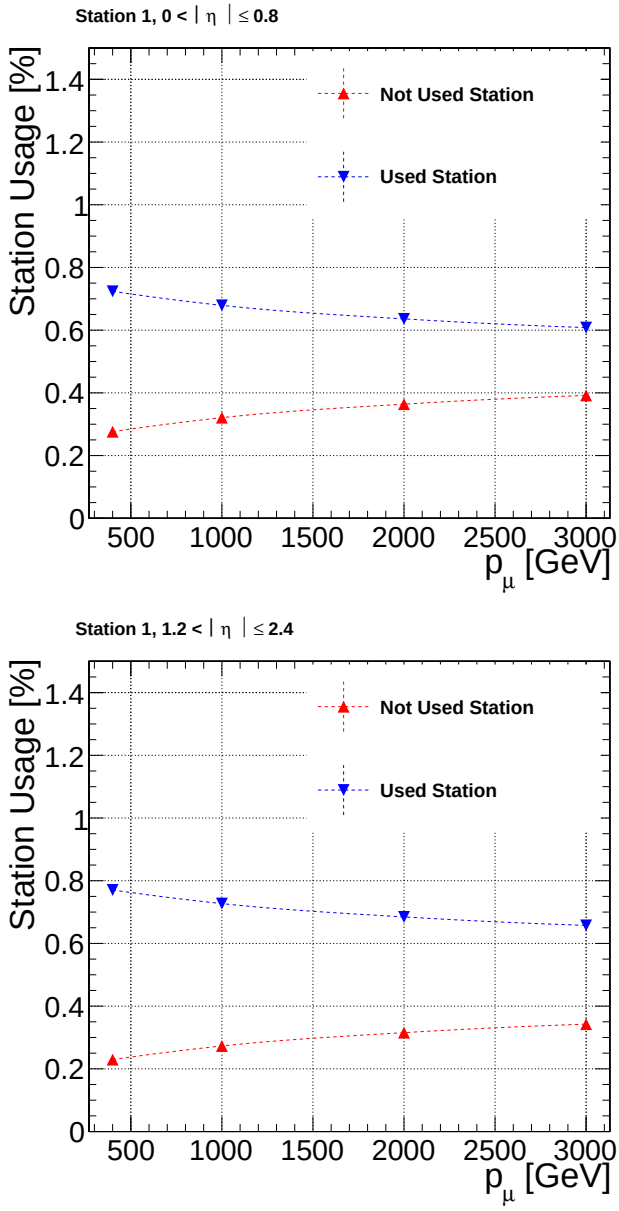

Figure 5.7: Distribution reporting the frequency at which the first station is used (Blue distribution) and skipped (Red distribution) as a function of the nominal muon momentum in both the barrel (*top*) and in the endcap (*bottom*) region of the muon spectrometer.

input the collection of hits associated to each Global Muon track and they output a new collection which can contain all, a subset of the initial hits or also additional hits. As soon as the new collection is available, a new fit with a new estimate of the momentum will be performed. The Tev-refitters are:

- *Tracker-Plus-First-Muon-Station (TPFMS) fit*. Most of the gain obtained by combining the inner tracker and the stand-alone muon comes from the first hit used in the muon system [85]. TPFMS takes advantage of this fact by performing a refit using only the tracker hits and the hits coming from the first muon station which is not shielded by any layer of iron.

- *Picky fit*. This refitter analyzes the level of occupancy inside each chamber to identify a possible catastrophic process. For each hit of the global muon, the multiplicity of hits in the associated chamber is checked. The scan is performed looking at the isolation of the hit inside a cone centered on it: if the number of hits inside the cone is greater than a threshold value, the chamber is tagged as contaminated. The second step of the algorithm consists of a loop over the hits of the contaminated chamber to check the compatibility of each hit with the extrapolated trajectory based on a χ^2 requirement. All the hits that have passed the last compatibility requirements are then finally used in the refitting procedure.

- *Dynamic Truncation (DYT) algorithm*. The DYT has been developed in the context of this thesis and uses a different and partly complementary approach compared to Picky. In most of the cases the shower produced by a radiative energy loss is fully contained within the iron without producing any high occupancy in the following chambers. The DYT quantifies the compatibility between the extrapolated track state and the reconstructed segments inside the chambers: if the smallest discrepancy is higher than a threshold the segments of that chamber are not taken into account in the reconstruction. The next section is fully dedicated to the Dynamic Truncation, and the reader will find more details there.

To further improve the performance of the reconstruction of high-energy muons the output of the previously described Tev-refitters is given as an input to the so-called tune-P algorithm. On a track-by-track basis TuneP chooses the best available track among the ones produced by the previous refitters and the traditional algorithms (Global Muon and Tracker Muon) described in the third chapter of this thesis using the tail probability of the χ^2 of the track fit. During Run-I, TuneP was using only Picky, TPFMS and the Tracker muon algorithms; since the beginning of Run-II the first version of the DYT has been integrated in the official CMS standard muon reconstruction framework and TuneP has included it accordingly.

The selection of the best available track in the Run-II version of TuneP, is still on a track-by-track basis using both the normalized χ^2 of the track fit and the relative p_T error measurement $\delta p_T/p_T$ (p_T^{err}). The algorithm starts its search for the best track considering initially the Picky hypothesis and comparing its p_T^{err} with the value estimated for the corresponding track but re-fitted by the DYT algorithm. The track with the lowest p_T^{err} value is kept and compared with the Tracker muon using the normalized χ^2 variable. The track with the lower χ^2 value is kept and finally compared with the TPFMS re-fitter algorithm. The final best track is chosen after the last comparison according to the lower normalized χ^2 value. At the end, if the final candidate track or the tracker-only track have p_T lower than 200 GeV, the tracker-only track is selected.

5.2 The Dynamic Truncation Algorithm

5.2.1 Basic principle

As explained in the previous section, the probability to have catastrophic energy losses for particles passing through iron and with momentum greater then 300 GeV is not negligible. Energy loss for radiative processes feature hard spectra and large fluctuations, making impossible its description as a uniform and continuous process in the Kalman Filter (KF) used by the CMS muon reconstruction. For this reason, after a large radiative energy loss the remaining reconstructed hits of a muon track will introduce a bias in the track fit given that no account is taken for them in the propagation step.

The Dynamic Truncation Algorithm (DYT) is based on the idea of stopping the Kalman Filter once a large energy loss is identified, avoiding that the remaining hits would bias the momentum measurement. Radiative processes are likely to happen within the calorimeters and the iron yoke, and, though they usually feature electromagnetic showers, the latter could not necessarily appear in the muon chambers because it could well be fully contained in the calorimeters or iron. Moreover no clear correlation exists between shower size and energy loss. For these reasons, a strategy for the identification of large energy losses based on their associated showers would be sub-optimal. Instead, the DYT strategy is based on the compatibility between the extrapolated track state from the outermost tracker layer, and the reconstructed segment in the muon chambers, as shown in the Figure 5.8. Because no attempt is made in the extrapolation to model the effect of radiative losses, compatibility is expected whenever such losses have not occurred and the considered reconstructed segment is the correct one. Within the CMSSW software, the compatibility test is performed between the so-called TrajectoryStateOnSurface (TSOS), which represents the track state on a detector layer, and a 4D reconstructed segment (DTRecSegment4D and CSCSegment). Both of them consist of a vector ($v = (x, y, dx/dz, dy/dz)$) containing the information of the local position (x, y) and the local direction $(dx/dz, dy/dz)$ of the segment estimated in the reference frame of the chamber on the plane with $z = 0$[1]. The degree of compatibility is measured in the local frame of the muon chamber by the operator defined as:

$$E = vMv^T \tag{5.3}$$

[1]The CSC local reference frame is defined in the following way: the local z-axis is parallel to the beam, whereas the y-axis is parallel to the r-axis direction and finally the x-axis is perpendicular to the y-z plane. The local reference frame of the DT chamber is defined by the z-axis pointing to the center of the CMS detector, the y-axis parallel to the beam line and finally the x-axis perpendicular to the y-z plane along the ϕ direction.

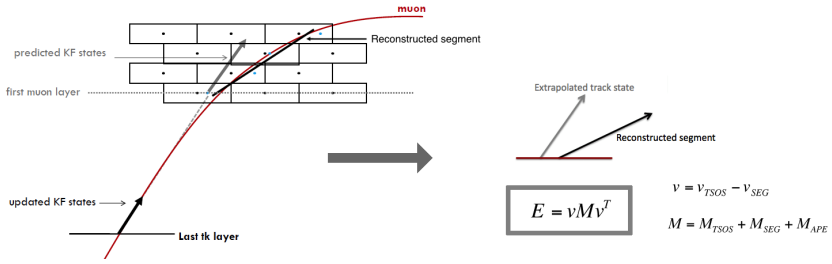

Figure 5.8: Representation of the comparison performed by the DYT between the predicted TrajectoryStateOnSurface (TSOS) and the reconstructed segments.

where v is the difference between the TSOS and 4D segments, and M is the sum of their covariance matrices and the alignment position error matrix. In the absence of correlations this estimator takes the classical form of a χ^2 variable associated with a model that predicts null measurements. In its full form correlations between measurements are taken into account.

The TSOS and the 4D reconstructed segments have the same structure with the difference that the former is a 5D vector, with the additional component given by q/p. For this reason, in order to compute the estimator E, the q/p component is removed from the TSOS and, consequently, the associated covariance matrix is reduced to a 4×4 matrix.

The way the DYT works is summarized in the flowchart 5.9. The TSOS on the last tracker layer is propagated, together with its uncertainties, to the first muon station and the resulting track state is compared to all the reconstructed segments in the chamber. If the minimum value of E, among all segments, is below a given configurable threshold (Thr), the most compatible segment is selected, the TSOS is updated[2] and propagated to next station. If the minimum value of E is above the threshold the DYT skips this measurement and checks whether such a large incompatibility is also found in the next station: if the incompatibility is confirmed then the collection of track measurements is stopped before the first incompatible muon station, otherwise just the incompatible measurement is skipped.

[2]In Kalman filter formalism the *update* step consists of computing the weighted average of the extrapolated track state, which carries informations from the measurements on the previous layers, and the measurement found on the detecting layer where the track state has been extrapolated to. The uncertainties play the role of weights in this matching.

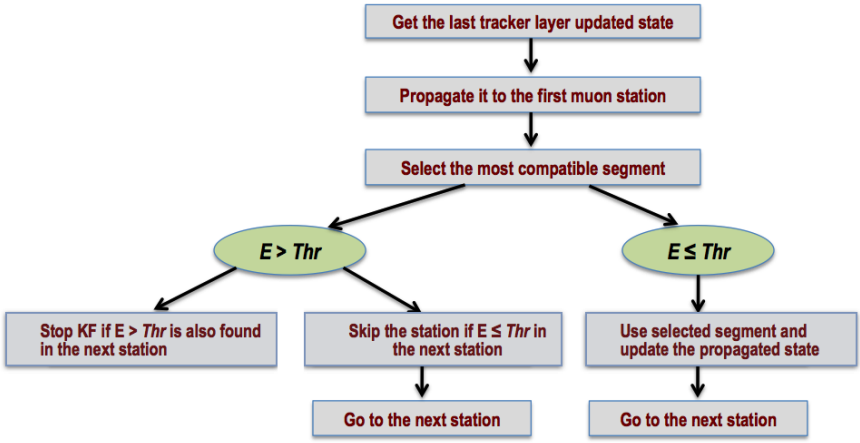

Figure 5.9: Flowchart representing schematically the procedure used by the DYT algorithm.

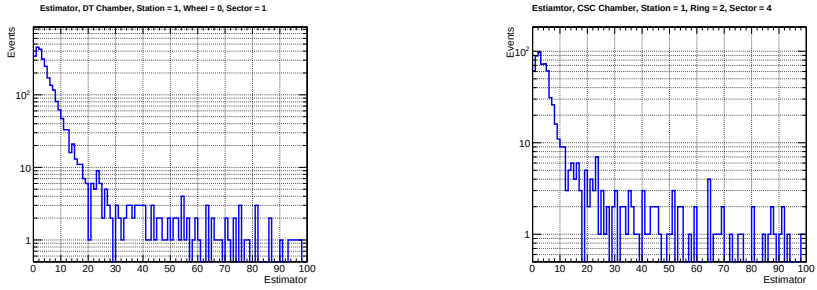

Figure 5.10: Estimator distribution in the barrel (*left*) and in the endcap (*right*). In the case no correlation is present in the error matrix, and in absence of biases, the estimator is expected to follow to a chi-square distribution.

5.2.2 First version of the DYT and its performance

The requirement of having accurate measurements for very energetic muons with a low probability of significantly overestimating its momentum, comes from all those analysis searching for boosted objects or very heavy resonances decaying into muons and characterized by low background. Any muon algorithm used in such a search has to reduce to a minimum the chance that reconstructed high-energy muons are the result of only a mis-reconstruction. For this specific reasons the performance of the DYT is optimized via its capability to reduce the tails of the measured momentum distribution. The value of the threshold on the estimator E used in the algorithm is

thus defined as the one that minimizes the number of events that populate the tails of the muon momentum spectrum as better defined in the following.

The first version of the DYT, which is the one still in use in the official CMS software, presents only one fixed threshold value applied to each muon regardless of its pseudorapidity and momentum. The optimization procedure was performed entirely using Double muons (back-to-back) MC samples generated using a flat probability distribution in η, in ϕ and in p range (*Muon Particle Gun Generator*) [179]. The muons were generated in the full CMS acceptance region ($\eta \in [-2.5, 2.5]$ and $\phi \in [-3.14, 3.14]$) with momentum range defined by the interval 1 TeV $\pm$ 0.01%. The threshold value that yields the minimum number of events outside the region [0,2]$\cdot p_{sim}$ in the momentum spectrum was chosen.

The reconstructed muon momentum distribution is reported in Fig. 5.11[3]. The performance of the DYT is checked considering its ability to reduce both the number of events that fall outside the region defined by [0,2]$\cdot p_{sim}$ and the spread in muon momentum around the peak of the distribution, which is quantified by the RMS value[4]. The mean value and the standard deviation of a Gaussian fit performed around the peak of the distribution is also looked at to evaluate the reconstruction performance.

In all the three distributions, the first aspect that emerges is the DYT ability compared to Picky[5] to reduce the events in the overflow region, despite a more sizable non-Gaussian tail in the region right below the overflow threshold. The region which benefits the most is the barrel region with an improvement of 62% compared to Picky, whereas in the endcap region the improvements settled to about 24%. In addition, it has to be underlined that including the events in the overflow region doesn't affect considerably the RMS value. The deterioration in the barrel amount to about 6% and it is related to the presence of the tail on the right of the momentum distribution peak. In both the Picky and DYT momentum distribution no bias is visible in the mean of the Gaussian fit performed on the core of the distribution, and no difference is observed in the sigma of the same fit.

[3]The meaning of the variables reported in each statistical box that goes with the distribution is here briefly described: *Entries* identifies the number of events used to produce the distribution, *Mean* reports the average of the histogram entries, *RMS* reports the standard deviation of the histogram entries, the *Overflow* quantifies the number of outliers that fall outside the upper boundary of the distribution, $\chi^2/ndof$ report the normalized χ^2 value obtained from the Gaussian fit, *Constant* is the amplitude of the fitted Gaussian function, *Mean* is the mean of the Gaussian function, and finally σ is the standard deviation of the Gaussian distribution.

[4]The RMS is estimated considering only the entries used to fill the histogram in the x-axis range and it is defined as $\sqrt{1/N \cdot \sum(x_i - x_{mean})^2}$, where the index i runs over all the entries. In the specific case of the muon momentum the range of the x-axis is defined as [0,2]$\cdot p_{sim}$.

[5]Among the Tev-refitters, Picky is the one which gives the best reconstruction performances for high energetic muons and for this reason, it will be used as benchmark in the evaluation of the DYT performances.

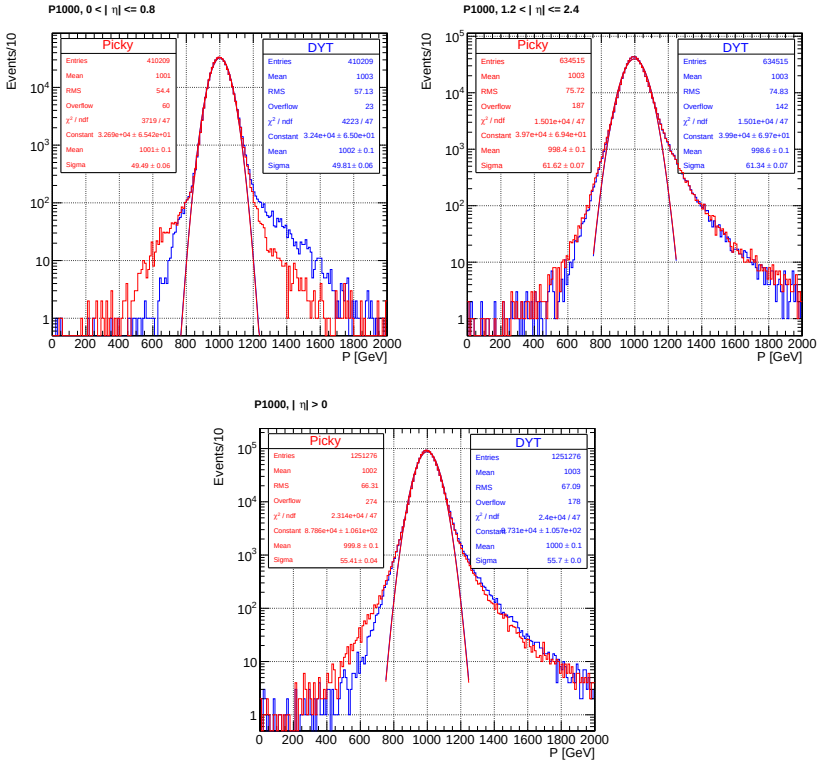

Figure 5.11: Comparison of the muon reconstructed momentum obtained by the DYT algorithm (*blue distribution*) and the Picky algorithm (*red distribution*) using a sample of muons generated with the Particle Muon Gun [179]. The muons are generated in the full acceptance region of the CMS detector ($\eta \in [-2.5, 2.5]$ and $\phi \in [-3.14, 3.14]$), requiring a momentum of 1 TeV $\pm$ 0.01%. Three different regions in η are reported: the barrel region (*Top left*), the endcap region (*Top right*) and finally the entire η region (*Bottom*). The Gaussian fit is performed in the region defined by $\mathrm{p}_{sim} \pm 25\%$.

The next result that will be presented show the normalized relative residual distribution (pull). The pull is defined as:

$$P(q/p_{reco}) = \frac{q/p_{rec} - q/p_{sim}}{\sigma(q/p_{rec})} \tag{5.4}$$

where p_{sim} is the muon generated momentum, whereas the p_{rec} is the reconstructed momentum. The main feature of this variable lies in the fact that the associated events are expected to distribute according to a normalized Gaussian centered on zero with standard deviation equal to 1. The Gaussian fit on the core of the dis-

tribution gives important informations about possible errors on the reconstructed momentum and its uncertainty. The more the values measured by the fit vary from the expected parameters, the more the systematic errors in the muon track reconstruction are pronounced. In particular, if the standard deviation of the pull is close to unity, then the track errors are well estimated. On the contrary, in the case the standard deviation is bigger then one it means that the errors on the momentum are underestimated, whereas if the standard deviation is lower then one the errors on the momentum are overestimated. Finally, if the mean value differs from zero, it means that the track momentum is systematically wrongly estimated.

Figure 5.12 shows the comparison of the pull distribution between the Picky and the DTY algorithm. Both the algorithms have very similar performance and they don't present any bias in the muon momentum estimate. In particular, the standard deviation has a maximum deviation from unity of about 4% in the the endcap region for the DYT, whereas the mean value presents a shift of 5% along the negative x-axis.

These results support the DYT algorithm strategy. However the levels of misalignment of the muon chambers, the accuracy of the estimation of the alignment position errors, the accuracy of the inner tracker track parameters, or of the muon segments could vary with the muon direction and momentum. These considerations have charted the way to develop and improve the DYT algorithm by using a dynamic threshold defined as a function of both the muon direction and momentum.

In addition, these conclusions are furthermore corroborated by the results obtained on the momentum spectrum of muons generated at 2 TeV reported in Fig. 5.13. Indeed, the promising performances shown at 1 TeV couldn't be regained as soon as the test moves away from the tuned region at 1 TeV.

Indeed, in the barrel region the number of the outliers events and the RMS values are respectively 164% and 6% worst then the values predicted by Picky. On the other hand in the endcap the DYT predicts 25% less events in the overflow region then Picky and the improvement in the RMS estimation amounts to 3%.

5.2.3 Second version of the DYT

The second version of the DYT is characterized by a threshold on the estimator value that changes dynamically with the momentum and the direction of the reconstructed muon. The double parametrization is achieved in two different steps:

- *DYT Tuning at Different Momentum Values*: the tuning of the DYT follows the same strategy adopted in the first version of the algorithm but extends the procedure to a few values of momentum in different pseudorapidity regions. In order to optimize the DYT in the muon momentum spectrum covered by the CMS analyse, the muon momentum values (p_{sim}) used in the tuning step are

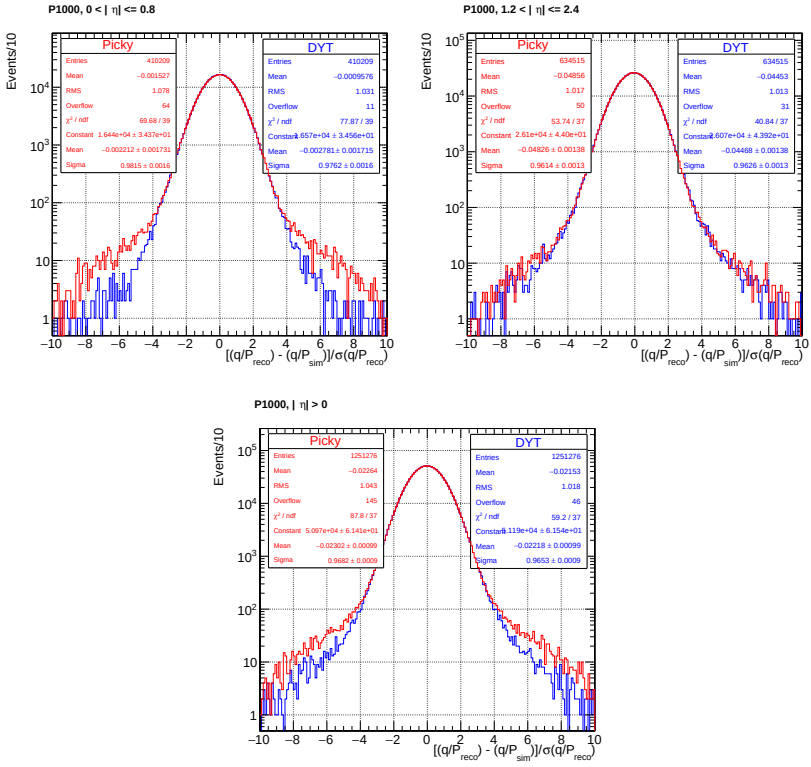

Figure 5.12: Comparison of the pull distribution of the reconstructed momentum obtained by the DYT algorithm (*blue distribution*) and the Picky algorithm (*red distribution*) using a sample of muons generated with the Particle Muon Gun [179]. The muons are generated in the full acceptance region of the CMS detector ($\eta \in [-2.5, 2.5]$ and $\phi \in [-3.14, 3.14]$), requiring a momentum of 1 TeV $\pm$ 0.01%. Three different regions in η are reported: the barrel region (*Top left*), the endcap region (*Top right*) and finally the entire η region (*Bottom*). The Gaussian fit is performed in the region defined by [-2,+2]· RMS.

50 GeV, 400 GeV, 1 TeV, 2 TeV and 3 TeV. The CMS pseudorapidity region is divided in five sub-regions: $0 < |\eta| < 0.8$, $0.8 < |\eta| < 1.2$, $1.2 < |\eta| < 2.0$, $2.0 < |\eta| < 2.2$, and $2.2 < |\eta| < 2.4$. The range of the tested threshold values is [1,95]. For each value of p_{sim}, the comparison among the different thresholds is performed in each pseudorapidity region by checking the number of simulated muons whose DYT estimated momentum falls outside the window [0,2]· p_{sim}. The threshold value with the lowest number of outliers is chosen. In the case two threshold values have the same minimum number of outliers,

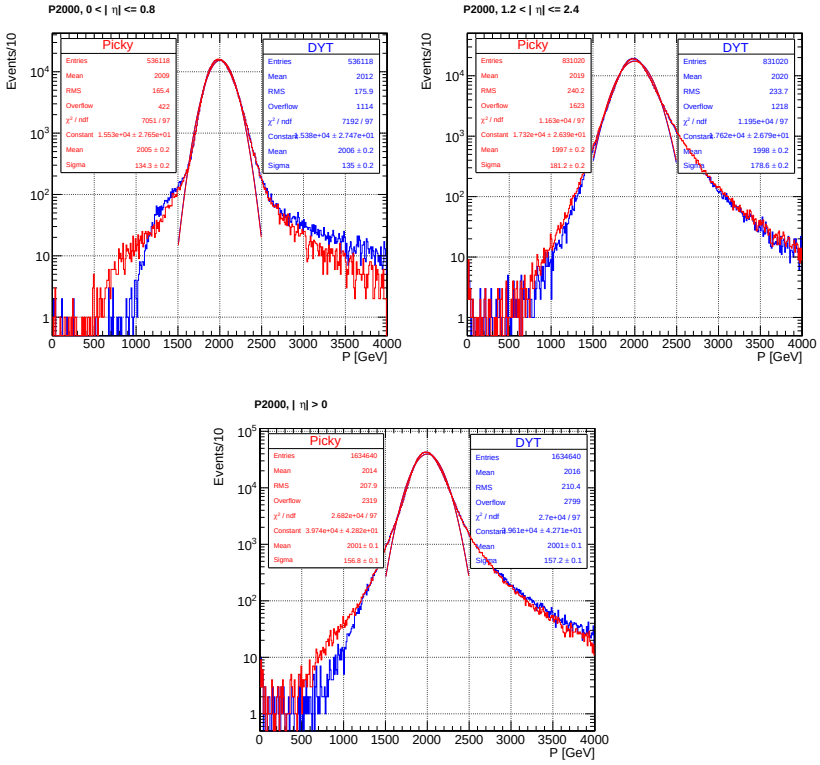

Figure 5.13: Comparison of the muon reconstructed momentum obtained by the DYT algorithm and the Picky algorithm using a sample of muons generated with the Particle Muon Gun [179]. The muons are generated in the full acceptance region of the CMS detector ($\eta \in [-2.5, 2.5]$ and $\phi \in [-3.14, 3.14]$), requiring a momentum defined by the interval 2 TeV $\pm 0.01\%$. Three different regions in η are reported: the barrel region (*Top left*), the endcap region (*Top right*) and finally the entire η region (*Bottom*). The Gaussian fit is performed in the region defined by $\mathrm{p}_{sim} \pm 25\%$.

the threshold corresponding to the momentum with the lowest RMS is chosen. At the end of the tuning procedure each pseudorapidity region is characterized by a set of threshold values, one for each p_{sim}. As an example, the amount of outliers as a function of the threshold values in the entire pseudorapidity region ($0 < |\eta| < 2.4$) for different values of p_{sim} is reported in Fig. 5.14. The comparison highlights all the properties of the dependency of the overflow on the threshold value and the nominal p_{sim} value. Indeed on one side the overflow events at a fixed thresholds value tend always to increase with the increase of the generated momentum, and on the other side at a fixed p_{sim} the number of

outliers decreases until reaching a global minimum value and restarts to increase again after it.

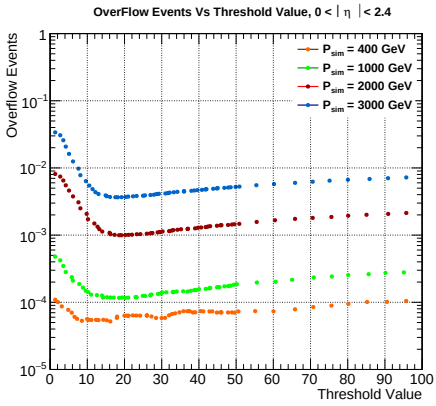

Figure 5.14: Example of the overflow values distributions used to tune the second version of the DYT as a function of the different values of the threshold for different generated muon momentum values (400 GeV, 1000 GeV, 2000 GeV, 3000 GeV). The distribution reports the amount of outliers for the entire pseudorapidty acceptance of the CMS detector ($0 < |\eta| < 2.4$).

- *Fitting the different threshold values*: the goal of the fit is to find for each η region a continuos parametrization in momentum in such a way that the DYT adopts the most optimal threshold value to use for it. The parametric equation of the fitting function is found empirically following two principles:
  - the same parametric equation has to be used preferably in all the pseudorapidity regions
  - the parametric equation has to be defined by a minimum number of free parameters

  The parametric equation of the fitting function is:

$$Thrs(p) = Thrs_{\eta}(50GeV) \cdot (1 + \alpha \cdot p_{\mu} + p_{\mu}^{\beta}) \tag{5.5}$$

  where α and β are the free parameters of the fit and they are specific for each pseudorapidity region, Thrs$_\eta$(50 GeV) is the optimal threshold associated to each η region for muons with the generated momentum equal to 50 GeV, finally p_μ is the momentum of the muon. To apply properly the fitting function and

identify the best threshold value to use for each muon particle, the DYT has to know in advance a preliminary estimate of its direction and its momentum. In the current version of the DYT, these informations are taken from the global muon track.

The fitting functions in all the η regions are reported in Fig. 5.15.

In the second part of this section the performance of the second version of the DYT will be presented. The first group of results shows a comparison between the two versions of the DYT, whereas the second group is a comparison between the DYT and the Picky algorithm.

The comparison of the muon momentum reconstruction performance between the first version and the second version of the DYT algorithm shows that the benefits of the dynamic estimation of the threshold values are shared with all the generated momentum points. Actually, the improvements become more and more significant with the increase of the momentum value. Indeed, focusing on the momentum distribution, the events of the overflow are always reduced in the second DYT version. In particular the reduction of the outliers in the barrel region amounts to 61% for muons generated with momentum of 1 TeV and 81% for muons generated at 2 TeV. On the contrary, the reduction of the overflow events in the endcap is limited to 7% for muons with generated momentum of 1 TeV and 5% for muons with generated momentum at 2 TeV. At 400 GeV, the reduction of the outliers in the tail of the momentum distribution doesn't exceed 7%. Crossing the previous considerations with the RMS values, it can be asserted that the outliers recovered in the barrel region populate mainly the core of the momentum spectrum. Indeed in the barrel region, the RMS values associated to the muon momentum spectrum reconstructed by the second version of the DYT are systematically lower then the one estimated with the first version of the algorithm. In particular the RMS value is reduced by 9% at 1 TeV and 12% at 2 TeV. No significant improvements are measured for muons generated at 400 GeV. In the endcap region, a slight deterioration of the RMS (less the 1%) is measured for all the momentum values used in this study. These conclusions are supported by the performance on the resolution distributions. The comparison of the muon momentum spectrum at 400 GeV, 1 TeV and 2 TeV is reported in the Fig. 5.16, 5.19, and 5.22. The muon momentum

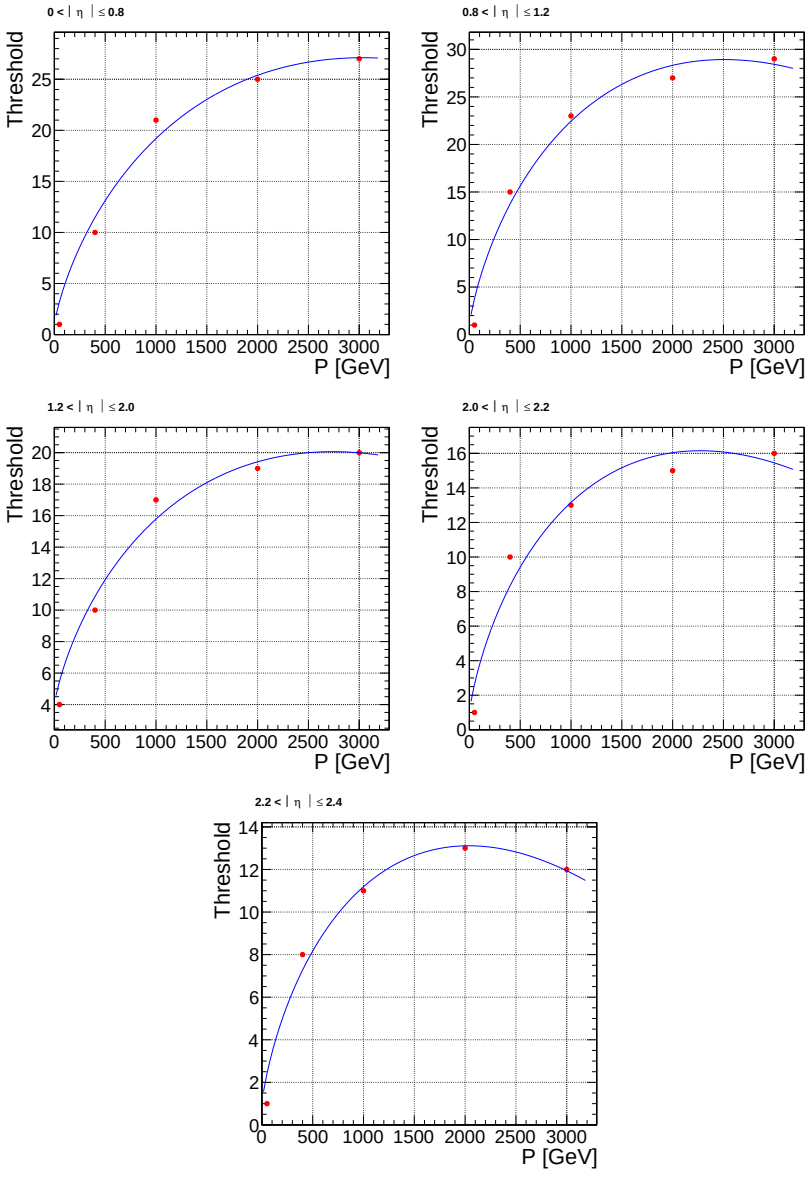

Figure 5.15: Threshold values (red dots) as a function of the generated momentum in the different pseudorapidity regions used in the tuning step. The parametric equation of the fitting function (blue curve) is reported in the Eq. 5.5.

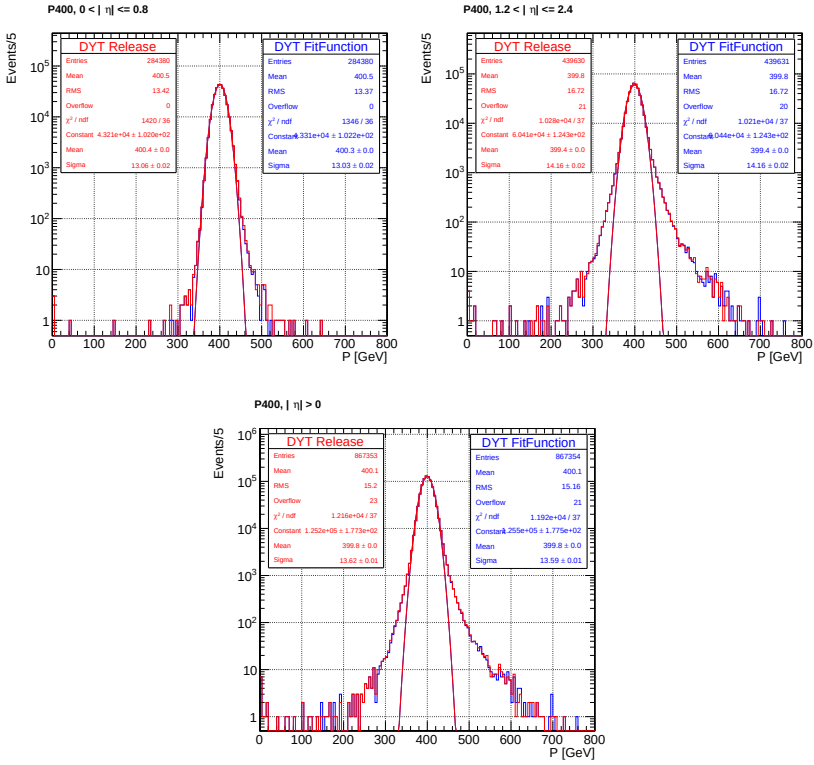

Figure 5.16: Comparison of the performances of the reconstructed muon momentum between the first version of the DYT (red distribution) and the new one (blue distribution) for muons generated with the Particle Muon Gun with momentum at 400 GeV. The results on the top left report the momentum distribution in the barrel, on the top right the momentum distribution in the endcap, finally on the bottom the entire pseudorapidity spectrum. The Gaussian fit is performed in the region defined by $Mean \pm 25\%$.

resolution for the different momentum values is shown in Fig. 5.17, 5.20 and 5.23, whereas their pull distribution is reported in Fig. 5.18, 5.21 and 5.24.

The comparison of the muon reconstruction performance between the DYT and Picky algorithms also profits from the dynamic estimation of the threshold values. Actually, the performance results at 2 TeV reported in Fig 5.25 show a substantial improvement in the reduction of both the events in the overflow region and the RMS. In particular, the outliers are reduced by about 56% and 29% respectively in the barrel and in the endcap. The improvement of the RMS amounts to 6% in the barrel and 2% in the endcap. The resolution distributions are reported in the Fig. 5.26. In particular

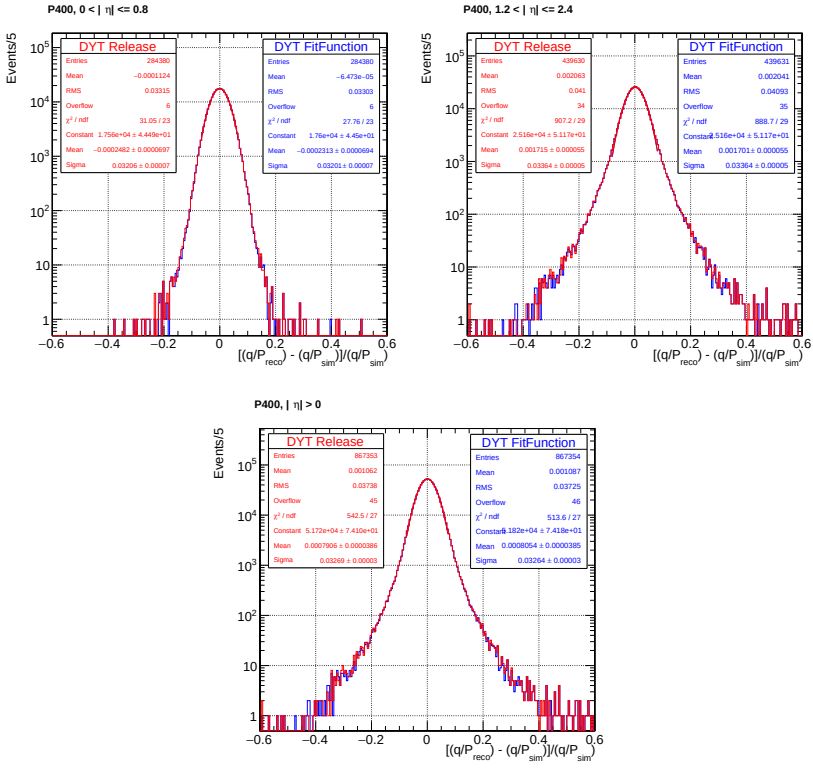

Figure 5.17: Comparison of the muon momentum resolution between the first version of the DYT (red distribution) and the new one (blue distribution) for muons generated with the Particle Muon Gun with momentum at 400 GeV. The results on the top left report the momentum distribution in the barrel, on the top right the momentum distribution in the endcap, finally on the bottom the entire pseudorapidity spectrum. The Gaussian fit is performed in the region defined by $Mean \pm 2 \cdot RMS$.

the benefits are localized in the tails of the resolution distribution where the improvements are of the order of 2%-4% respectively in the barrel region and in the endcap region. This variation is propagated on the mean value of the resolution. Indeed, a shift of the order of 15%-22% is measured respect to the mean value predicted by Picky. However, this effect is negligible considering that the precision at which the mean value is estimated is lower then 0.5%. The pull distributions are presented in Fig. 5.27. The performance of the DYT algorithm are in line with the ones obtained by Picky. In the barrel the mean value presents a shift in the negative region lower then 1% and the underestimation of the error is about 4%-5%. In the endcap the shift of the mean value is more significant and it is about 5%-6%, whereas the error presents an

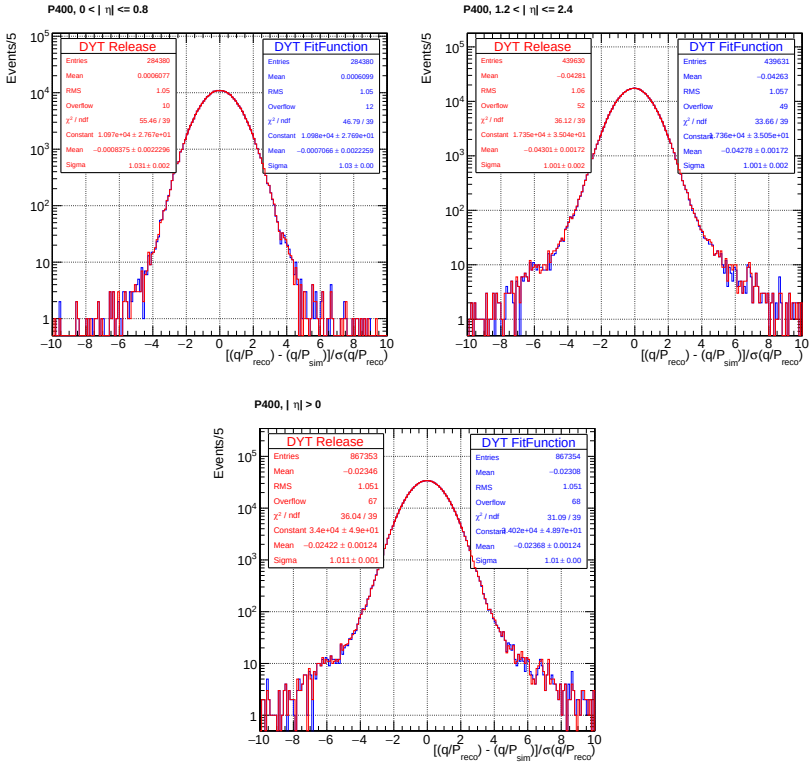

Figure 5.18: Comparison of the pull distibution between the first version of the DYT (red distribution) and the new one (blue distribution) for muons generated with the Particle Muon Gun with momentum at 400 GeV. The results on the top left report the momentum distribution in the barrel, on the top right the momentum distribution in the endcap, finally on the bottom the entire pseudorapidity spectrum. The Gaussian fit is performed in the region defined by $Mean \pm 2 \cdot RMS$.

overestimation of about 4%. For the sake of completeness the distributions of the momentum, the resolution, and the pull for muons generated with p_{sim} equal to 400 GeV, 1000 GeV are reported from Fig. 5.28 to Fig. 5.33 but they will not be commented.

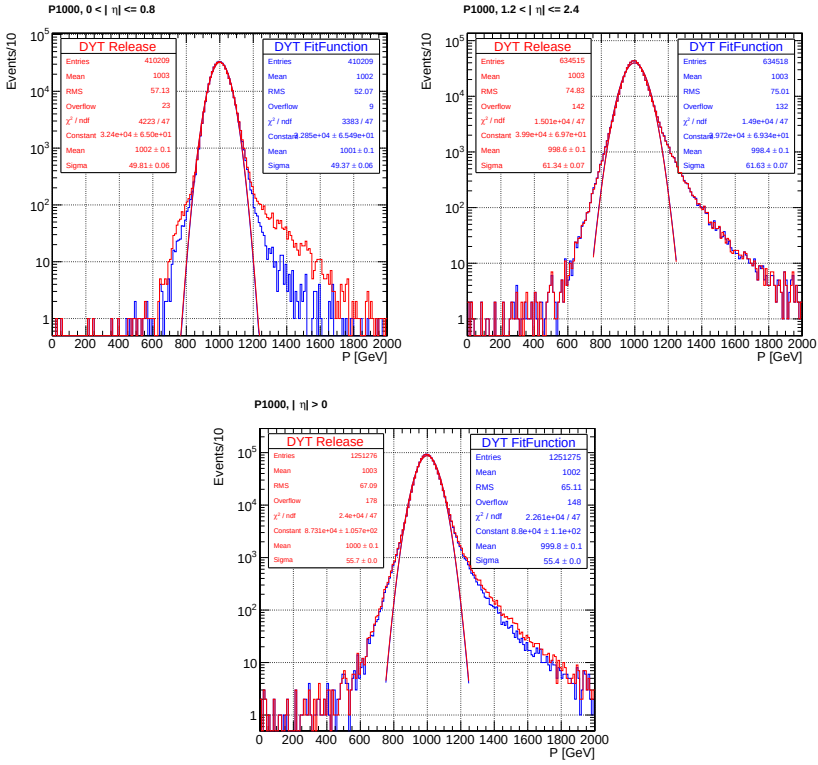

Figure 5.19: Comparison of the performances of the reconstructed muon momentum between the first version of the DYT (red distribution) and the new one (blue distribution) for muons generated with the Particle Muon Gun with momentum at 1 TeV. The results on the top left report the momentum distribution in the arrel, on the top right the momentum distribution in the endcap, finally on the bottom the entire pseudorapidity spectrum. The Gaussian fit is performed in the region defined by $Mean \pm 25\%$.

5.2.4 Characterization of the OverFlow in the Endcap Region

The results presented in the previous paragraph have shown the benefits of the new DYT approach. However, the same results show that the performance in the endcap region can be hardly improved. The distribution 5.34 gives the contribution to the overflow from all the pseudorapidity regions of CMS in the case of muons generated with the Particle Muon Gun with nominal momentum at 2 TeV after the

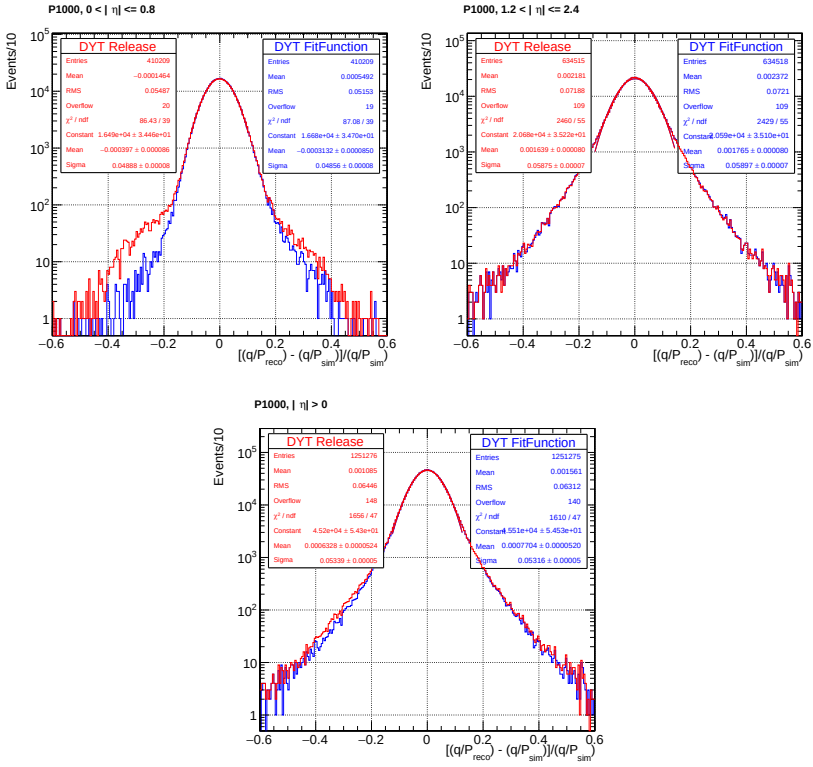

Figure 5.20: Comparison of the muon momentum resolution between the first version of the DYT (red distribution) and the new one (blue distribution) for muons generated with the Particle Muon Gun with momentum at 1 TeV. The results on the top left report the momentum distribution in the barrel, on the top right the momentum distribution in the endcap, finally on the bottom the entire pseudorapidity spectrum. The Gaussian fit is performed in the region defined by $Mean \pm 2 \cdot RMS$.

tuning procedure. The contribution of the pseudorapidity region between 2.0 and 2.4 contributes to 56% of the entire amount of events that fall inside the overflow region.

Can this effect be explained by considering only the variegated structure of the endcap region, or is it hiding a more complex and subtle reconstruction problem? To answer this question a detailed comparison between the segments reconstructed in the muon chamber (RecoSeg) used to refit the muon track (RecoTrack) and the segments (SimSeg) of the muon track simulated by Geant4 (SimTrack) is performed. It should be noted that Geant4 besides multiple scattering and ionization energy losses, properly models also the radiative energy losses. Matching each RecoTrack to its corresponding SimTrack in the local reference frame of the muon chamber allow any systematic

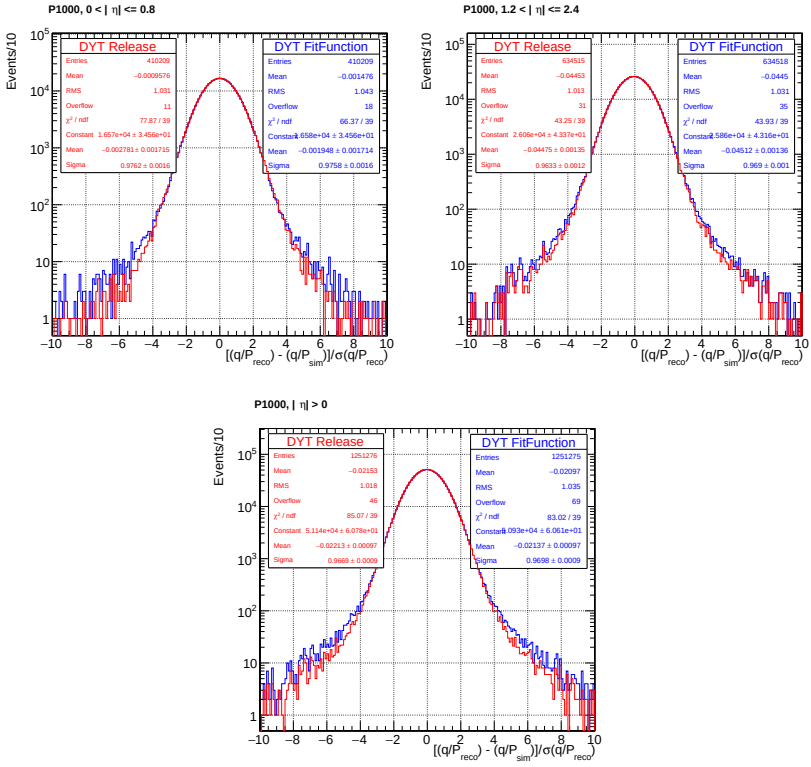

Figure 5.21: Comparison of the pull distribution between the first version of the DYT (red distribution) and the new one (blue distribution) for muons generated with the Particle Muon Gun with momentum at 1 TeV. The results on the top left report the momentum distribution in the barrel, on the top right the momentum distribution in the endcap, finally on the bottom the entire pseudorapidity spectrum. The Gaussian fit is performed in the region defined by $Mean \pm 2 \cdot RMS$.

error in the DYT reconstruction to be easily identified. The SimSeg has to be built starting from the Simulated hits (SimHit) associated to the real reconstructed hits of the RecoSeg. As soon as the SimHit collection is identified, the SimSeg is obtained solving the parametric equation of a straight line in the 3D space. Both the RecoSeg and the SimSeg are defined by a 4-vector v in the plane $z = 0$ of the muon chamber ($v = (x, y, dx/dz, dy/dz)$). The comparison between RecoSeg and SimSeg is performed considering the distribution of the resolution and the pull variables. In the endcap region the magnetic field bends muons the r-ϕ plane of the CMS detector, which coincides with the local x-z plane of the muon endcap region. The measurement of the muon momentum is directly related to the curvature of the muon track, for

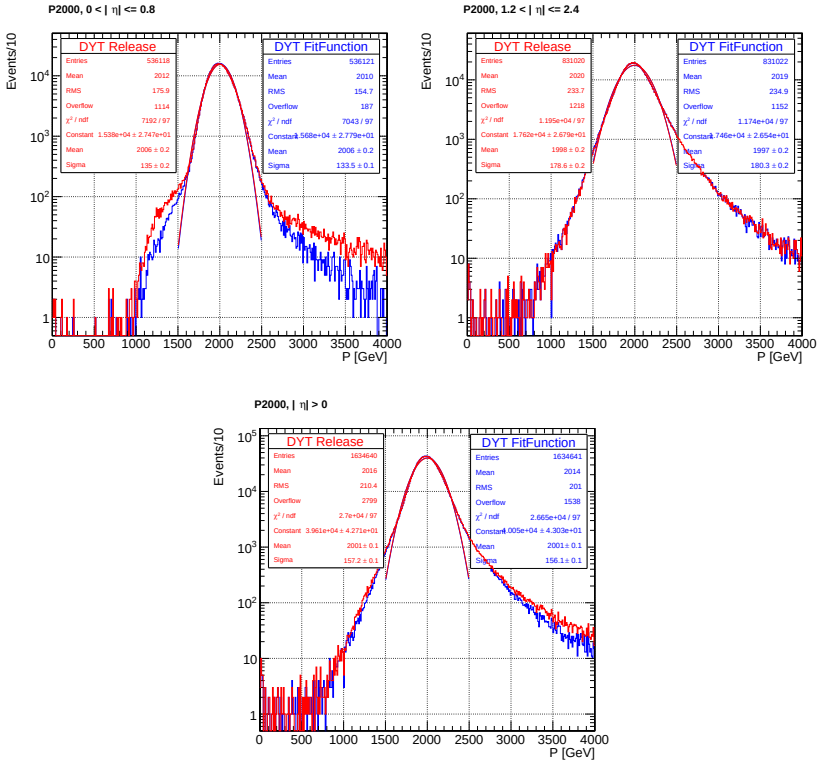

Figure 5.22: Comparison of the performances of the reconstructed muon momentum between the first version of the DYT (red distribution) and the new one (blue distribution) for muons generated with the Particle Muon Gun with momentum at 2 TeV. The results on the top left report the momentum distribution in the barrel, on the top right the momentum distribution in the endcap, finally on the bottom the entire pseudorapidity spectrum. The Gaussian fit is performed in the region defined by $Mean \pm 25\%$.

this reason the following study focuses on the local x and dx/dz components of the v vector. The resolution is defined by

$$R(v^i_{rec}) = v^i_{rec} - v^i_{sim} \tag{5.6}$$

where v^i refers to the ith component of the vector that define the segment in the local frame of the chamber. v_{rec} identifies the reconstructed segment, whereas v_{sim} identifies the simulated segment. The pull of the coordinate v^i is defined as the resolution of v^i divided by its associated uncertainty.

$$P(v^i_{rec}) = (v^i_{rec} - v^i_{sim})/\delta(v^i_{rec}) \tag{5.7}$$

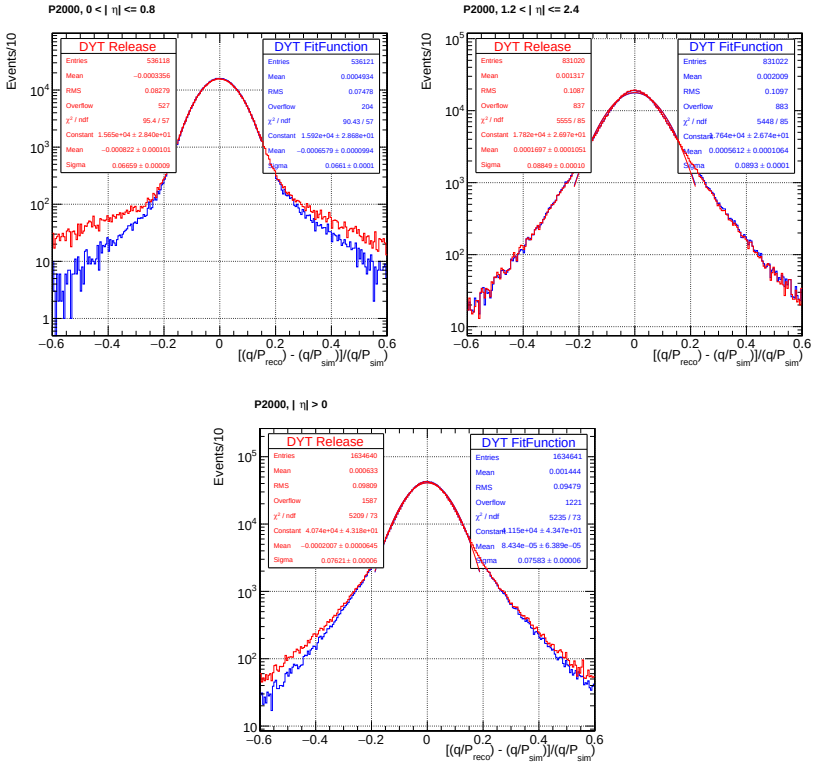

Figure 5.23: Comparison of the muon momentum resolution between the first version of the DYT (red distribution) and the new one (blue distribution) for muons generated with the Particle Muon Gun with momentum at 2 TeV. The results on the top left report the momentum distribution in the barrel, on the top right the momentum distribution in the endcap, and finally on the bottom the entire pseudorapidity spectrum. The Gaussian fit is performed in the region defined by $Mean \pm 2 \cdot RMS$.

The investigation is performed in two steps: the different distributions are first presented by stations and then, within a station, by rings. The first rough comparison doesn't reveal any reconstruction problem, wheres the second iteration has shown some systematic reconstruction problems in the CSC chambers of the first station. All the results that followes are obtained considering muons generated with momentum equal to 2 TeV.

Figure 5.35, reports the resolution of the x_{rec} and the $(dx/dz)_{rec}$ coordinates for all the segments reconstructed in all the chambers belonging to the first and second station separately. The agreement between the reconstructed and the simulated coordinates is very good. The majority of the events populates the core of the distribution,

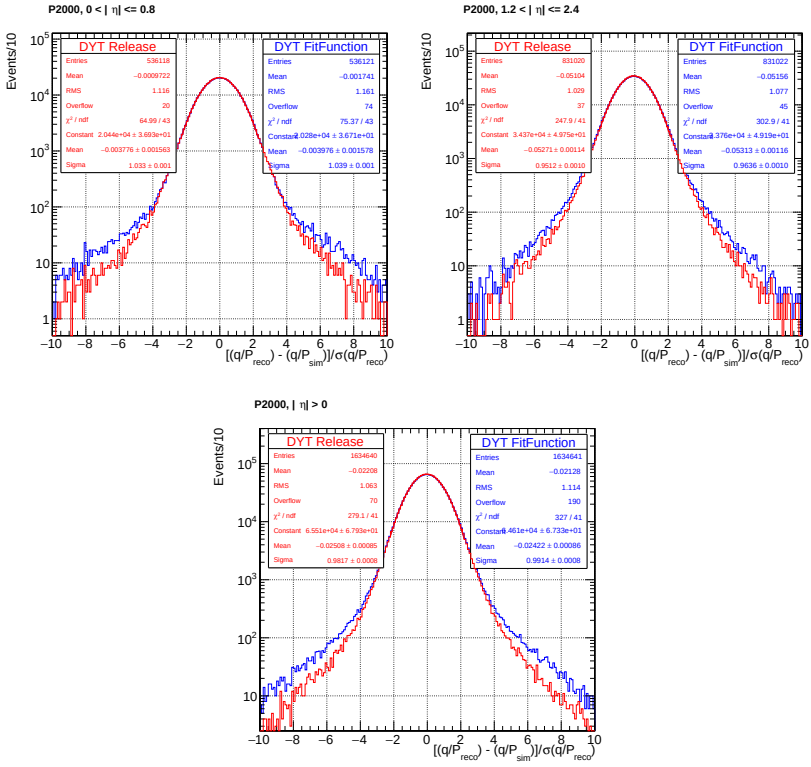

Figure 5.24: Comparison of the pull distribution between the first version of the DYT (red distribution) and the new one (blue distribution) for muons generated with the Particle Muon Gun with momentum at 2 TeV. The results on the top left report the momentum distribution in the barrel, on the top right the momentum distribution in the endap, and finally on the bottom the entire pseudorapidity spectrum. The Gaussian fit is performed in the region defined by $Mean \pm 2 \cdot RMS$.

the x_{rec} mean value of the distribution shifts from zero by a factor of the order of one micron, and the standard deviation amounts to about 100 μm. For the $(dx/dz)_{rec}$ direction the mean value is shifted from zero by a factor of the order of 10^{-4}-10^{-5}, and the RMS is lower then 0.5. The Gaussian fit on the core of the distribution defined by $Mean \pm 2 \cdot RMS$ draws to the same conclusions.

The pull distributions of the same variables are reported in Fig. 5.36. In both stations no bias in the estimate of the mean value is measured which is slightly shifted by a factor lower than 0.5%. On the contrary, the width of the Gaussian fit presents a reduction of 30% with respect to the expected value associated to a Normal Gaussian

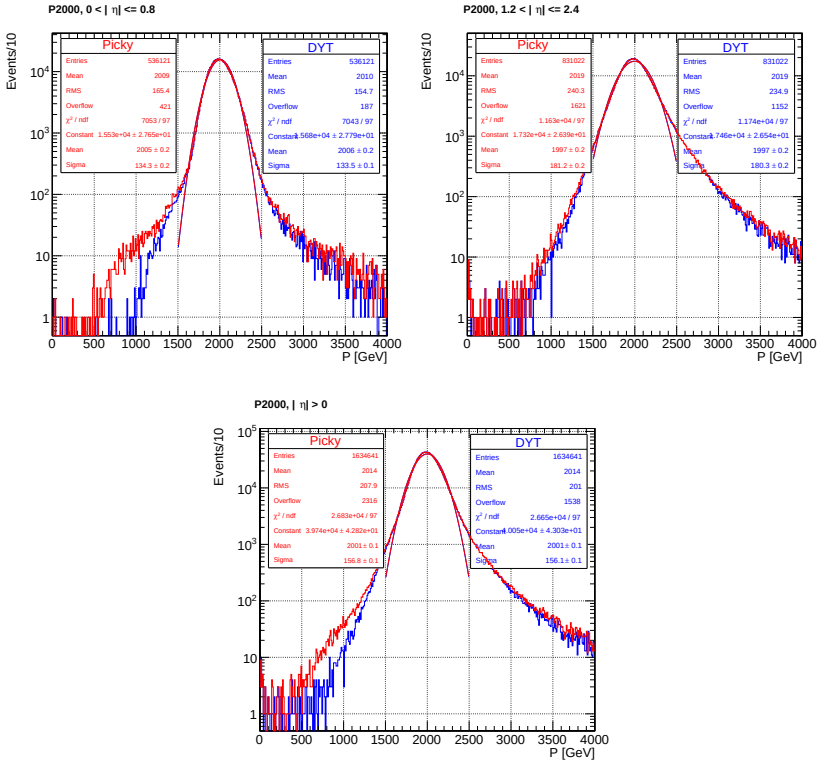

Figure 5.25: Comparison of the performances of the reconstructed muon momentum between the second version of the DYT (blue distribution) and Picky (red distribution) for muons generated with the Particle Muon Gun with momentum at 2 TeV. The results on the top left report the momentum distribution in the barrel, on the top right the momentum distribution in the endcap, and finally on the bottom the entire pseudorapidity spectrum. The Gaussian fit is performed in the region defined by $Mean \pm 25\%$.

distribution, indicating an over-estimation of the error associated to both the coordinates.

The results presented up to this point even if they refer to the performance of the muon reconstruction in the first and second station of the CSC sub-detector are representative also of the performance obtained in the third and fourth station that are not presented for the sake of conciseness.

Each CSC station considered as a whole is likely not to reveal any pathological problem in the local reconstruction of the segment in the x_{rec} and $(dx/dz)_{rec}$ coordinates, if these problems happen in particular regions of the stations. For this reason

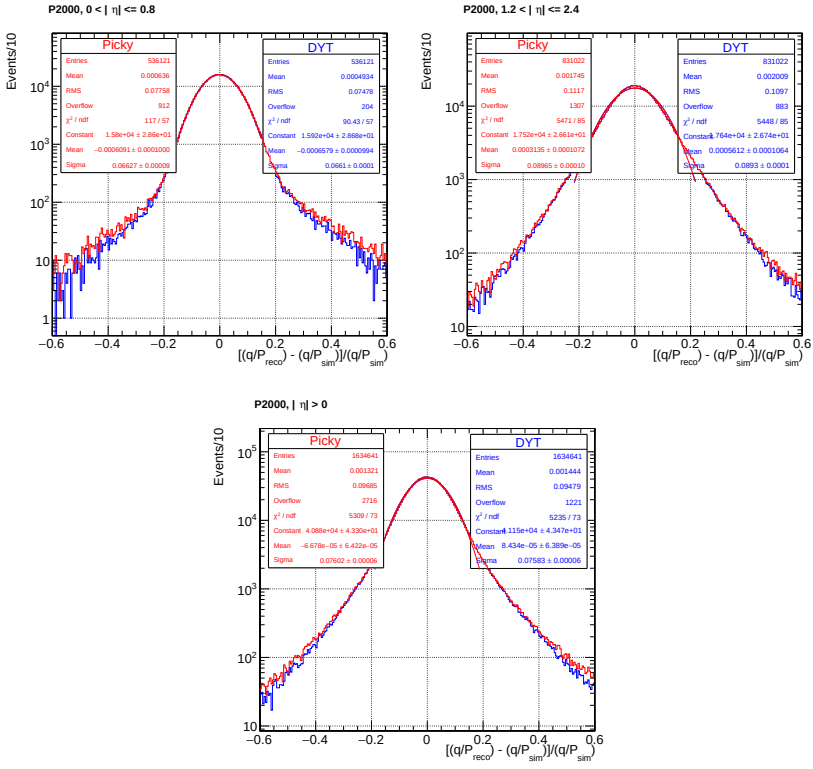

Figure 5.26: Comparison of the muon momentum resolution between the second version of the DYT (blue distribution) and Picky (red distribution) for muons generated with the Particle Muon Gun with momentum at 2 TeV. The results on the top left report the momentum distribution in the barrel, on the top right the momentum distribution in the endcap, and finally on the bottom the entire pseudorapidity spectrum. The Gaussian fit is performed in the region defined by $Mean \pm 2 \cdot RMS$.

a more detailed study considering also the ring sub-structure of each CSC station was performed.

Fig. 5.37 shows both the resolution and the pull distribution of the x_{rec} coordinate focusing on the first ring of the first station (ME1/1) and the first ring of the second station (ME2/1). Both the resolution distribution in ME1/1 and in ME2/1 are squeezed around the peak. Indeed the mean value is slightly shifted by a factor of the order of 1 μm respect to zero, whereas the maximum value for the RMS is about 0.5. The results obtained with the Gaussian fit on the core of the resolution distribution are in line with the one just described above. The Gaussian mean value of the pull distribution is shifted from zero by less then 1% in ME1/1 and about 1% in the ME2/1.

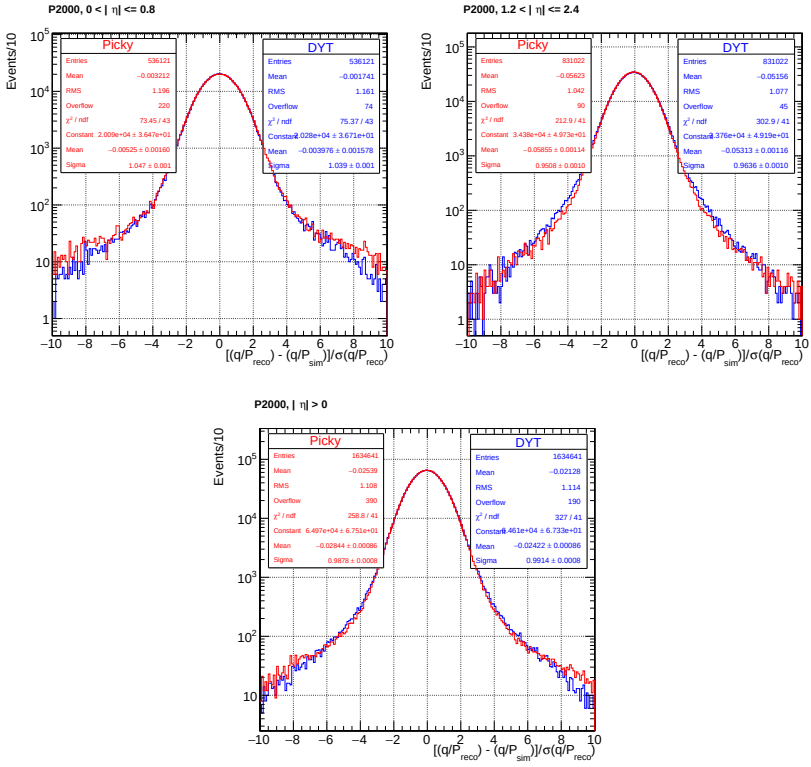

Figure 5.27: Comparison of the pull distribution between the second version of the DYT (blue distribution) and Picky (red distribution) for muons generated with the Particle Muon Gun with momentum at 2 TeV. The results on the top left report the momentum distribution in the barrel, on the top right the momentum distribution in the endcap, and finally on the bottom the entire pseudorapidity spectrum. The Gaussian fit is performed in the region defined by $Mean \pm 2 \cdot RMS$.

In both the cases an overestimation of the local errors associated to the local segment reconstruction is measured.

The resolution and the pull distribution for the dx/dz_{rec} direction are reported in Fig. 5.38. The first ring of the first station presents less events in the tail region with respect to ME2/1 and indeed the muon segment measurement are squeezed around the peak. Indeed both the RMS and the standard deviation associated to the Gaussian fit are less then 1%, in addition the mean value presents a negligible shift respect to zero of the order of 10^{-5}. The pull distribution reveals an anomaly in the ME1/1 which is not present in the ME2/1. Indeed the pull of the dx/dz coordinate in the ME1/1

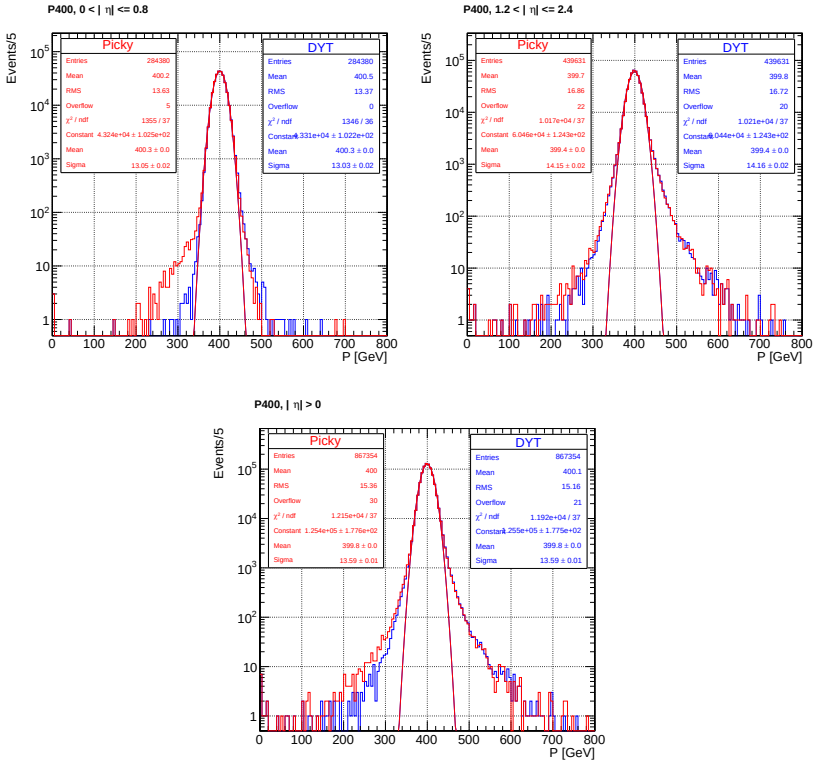

Figure 5.28: Comparison of the performances of the reconstructed muon momentum between the second version of the DYT (blue distribution) and Picky (red distribution) for muons generated with the Particle Muon Gun with momentum at 400 GeV. The results on the top left report the momentum distribution in the barrel, on the top right the momentum distribution in the endcap, and finally on the bottom the entire pseudorapidity spectrum. The Gaussian fit is performed in the region defined by $Mean \pm 25\%$.

present a double peak on the right and on the left side of zero, whereas the pull of the dx/dz direction for the segment that cross the first ring of the second station is centered in zero.

To better understand the source of the bias in the pull distribution of the dx/dz_{rec} variable in the chambers of the ME1/1 station, the correlation between the dx/dz_{rec} coordinate and its error $\delta(dx/dz_{rec})$ was studied in both the ME1/1A and the ME1/1B chambers. The results are reported in Fig. 5.39. The results associated to ME1/1A present some peculiarities if compared to the other distributions reported in the same figure. Indeed, the segment measurements are distributed along two bands that move

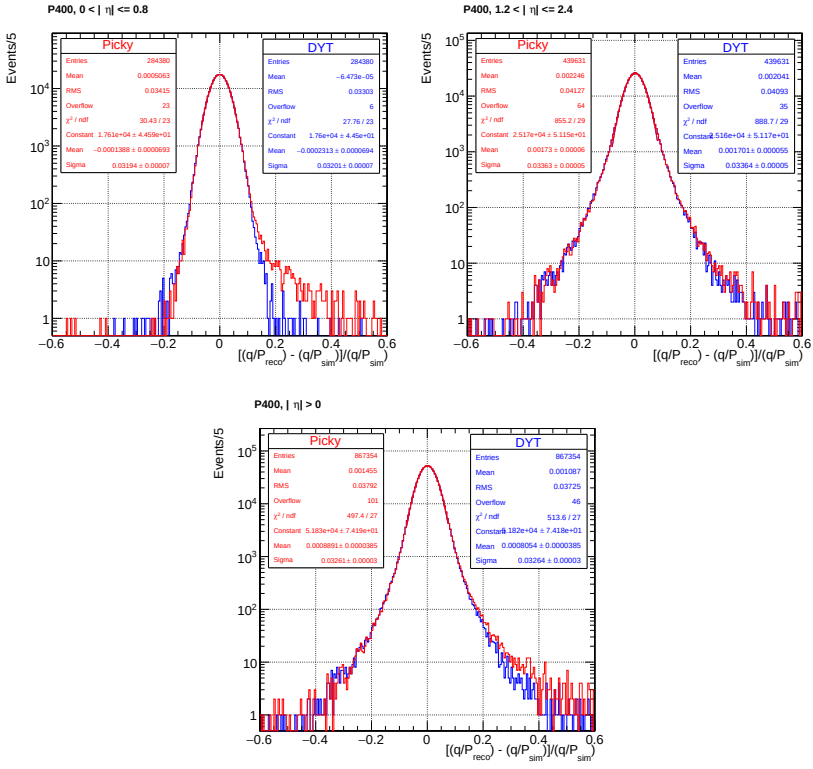

Figure 5.29: Comparison of the muon momentum resolution between the second version of the DYT (blue distribution) and Picky (red distribution) for muons generated with the Particle Muon Gun with momentum at 400 GeV. The results on the top left report the momentum distribution in the barrel, on the top right the momentum distribution in the endcap, and finally on the bottom the entire pseudorapidity spectrum. The Gaussian fit is performed in the region defined by $Mean \pm 2 \cdot RMS$.

in the opposite direction starting from the origin of the resolution axis. On the contrary in both ME1/1B and ME2/1 the measurements are uniformly spread around the mean value of the resolution and the range between 0.003 and 0.01 for the error associated to the dx/dz_{rec} direction.

The segments that cross the ME1/1A chambers and populate the two opposite bands of the correlation distribution previously described are further isolated according to the sign of their resolution value and they are used to study the correlation between their local coordinate x and y. The results are reported in Fig. 5.40 and they are divided by considering both the sign of the $(dx/dz)_{rec}$ resolution variable and

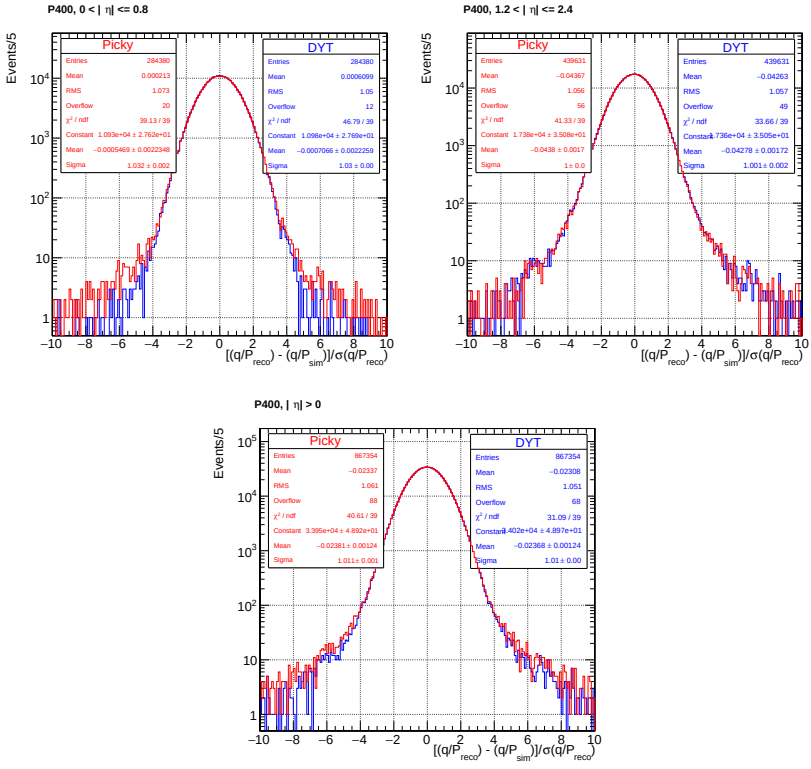

Figure 5.30: Comparison of the pull distribution between the second version of the DYT (blue distribution) and Picky (red distribution) for muons generated with the Particle Muon Gun with momentum at 400 GeV. The results on the top left report the momentum distribution in the barrel, on the top right the momentum distribution in the endcap, and finally on the bottom the entire pseudorapidity spectrum. The Gaussian fit is performed in the region defined by $Mean \pm 2 \cdot RMS$.

the endcap to which they belong. All the correlation results present a well defined diagonal pattern which mimic the structure of the wires in the CSC chambers of the first ring. However, to a more careful look this structure is not continuos on the entire x-range of the chamber but rather, the x-range is divided in two halves which matched with two diagonal staggered pattern. In addition the results in the x-y plane associated to the two opposite sign region of the resolution distribution present an overlap between the empty region that are present in the 2D distribution and the filled region of the 2D plots obtained with the segments associated to the opposite sign of the resolution variable.

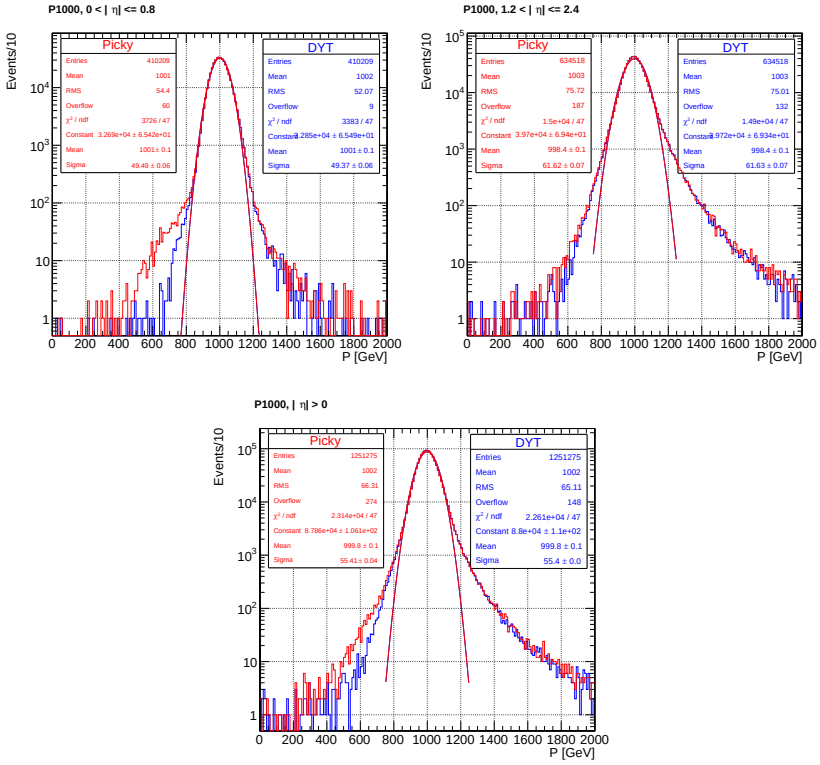

Figure 5.31: Comparison of the performances of the reconstructed muon momentum between the second version of the DYT (blue distribution) and Picky (red distribution) for muons generated with the Particle Muon Gun with momentum at 1 TeV. The results on the top left report the momentum distribution in the barrel, on the top right the momentum distribution in the endcap, and finally on the bottom the entire pseudorapidity spectrum. The Gaussian fit is performed in the region defined by $Mean \pm 25\%$.

The same distributions are presented for the ME1/1B chambers and are reported in Fig. 5.41. Also in this case the diagonal pattern related to the diagonal wires structure characterizing the ME1/1 chambers is evident. However, the different local position associated to each segment are more uniformly distributed in the whole chambers.

For illustrative purpose, the x-y scatter distributions for the chambers at the first ring of the second station are reported in the Fig. 5.42.

These results show a clear bias in the reconstruction of the segments in the CSC chambers ME1/1A and ME1/1 B. This bias could not be tracked down further in the context of this thesis because of lack of time, but its future correction is likely to bring

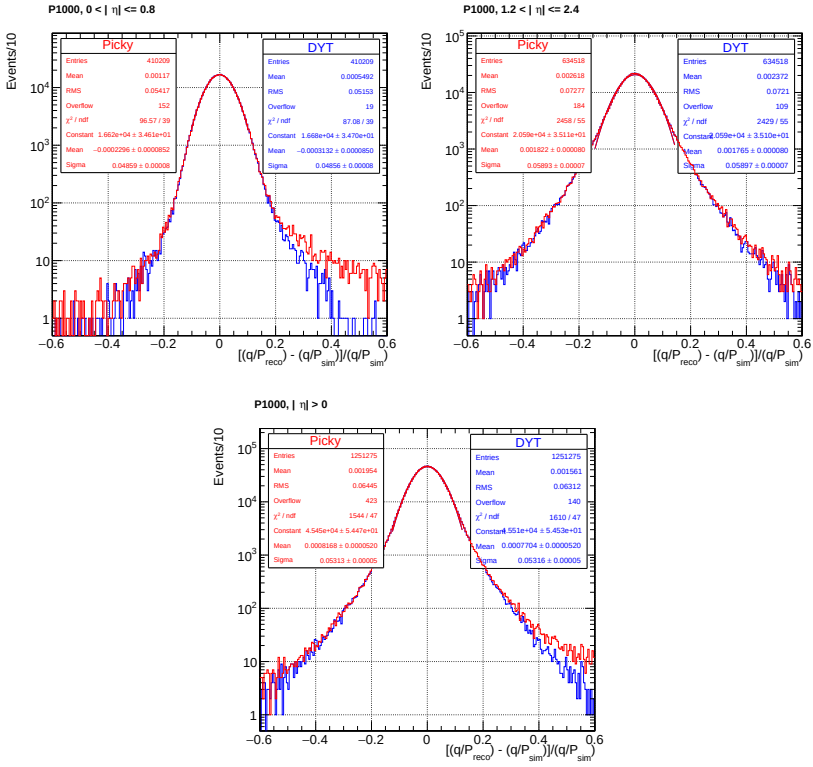

Figure 5.32: Comparison of the muon momentum resolution between the second version of the DYT (blue distribution) and Picky (red distribution) for muons generated with the Particle Muon Gun with momentum at 1 TeV. The results on the top left report the momentum distribution in the barrel, on the top right the momentum distribution in the endcap, and finally on the bottom the entire pseudorapidity spectrum. The Gaussian fit is performed in the region defined by $Mean \pm 2 \cdot RMS$.

a significant improvement in the muon reconstruction in the eta region [2, 2.4]. Once the correction will be applied the comparison of the DYT performance with the other muon TeV refitters will also have to be performed.

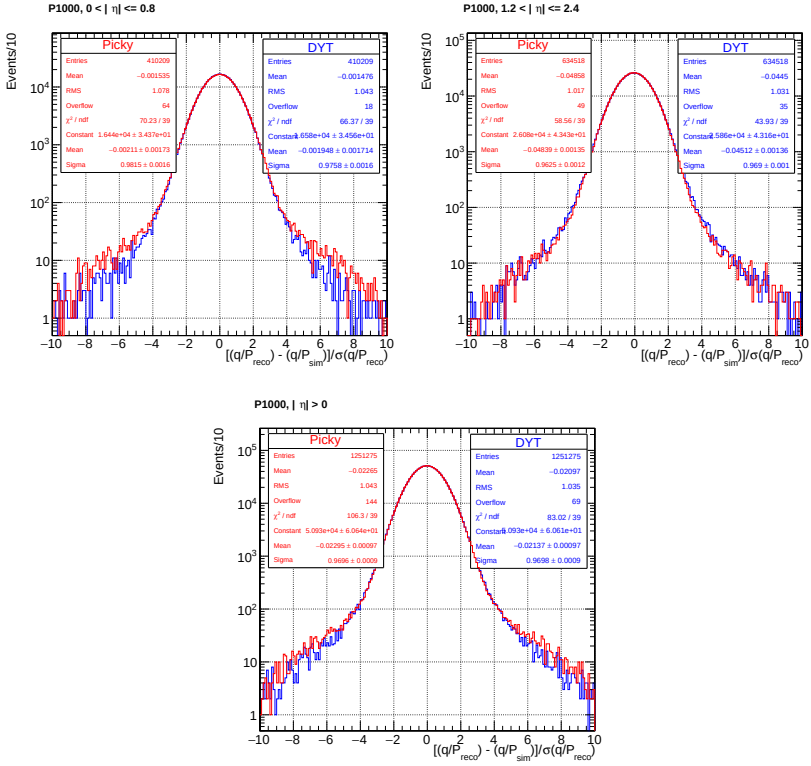

Figure 5.33: Comparison of the pull distribution between the second version of the DYT (blue distribution) and Picky (red distribution) for muons generated with the Particle Muon Gun with momentum at 1 TeV. The results on the top left report the momentum distribution in the barrel, on the top right the momentum distribution in the endcap, and finally on the bottom the entire pseudorapidity spectrum. The Gaussian fit is performed in the region defined by $Mean \pm 2 \cdot RMS$.

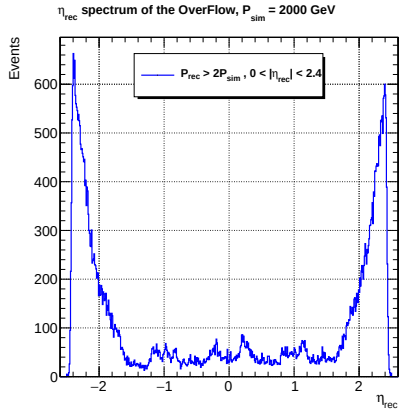

Figure 5.34: Distribution in η of the overflow events for muons generated with Particle Muon Gun with nominal momentum at 2000 GeV. The events are reconstructed using the second version of the DYT, using the optimized threshold values. The region $2.0 < |\eta| < 2.4$ contributes to the 56% of the entire amount of the events in the distribution.

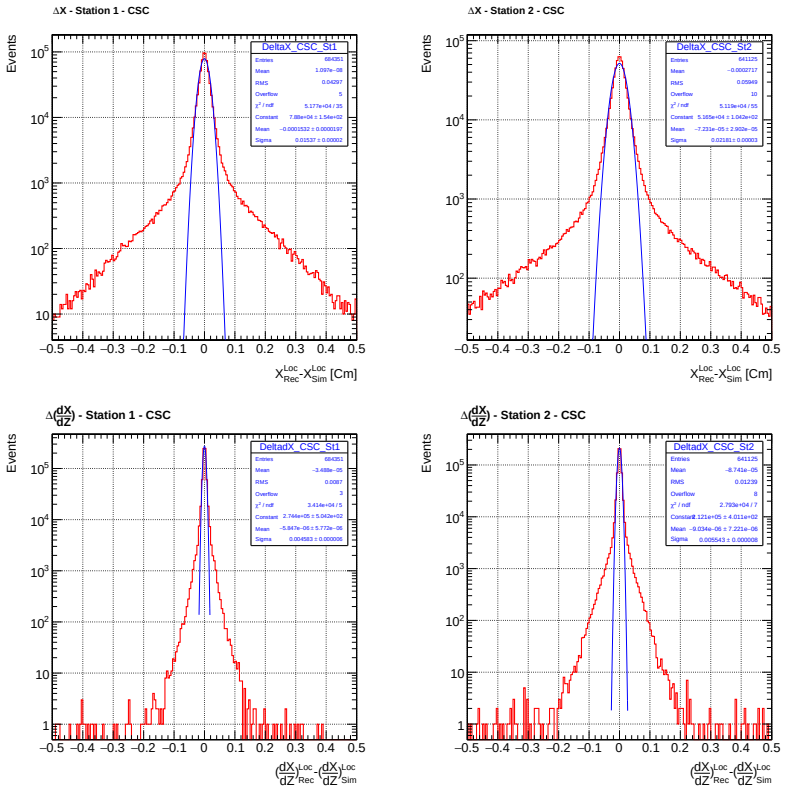

Figure 5.35: Resolution distributions for the x and dx/dz coordinates of the reconstructed segments associated to muons generated with nominal momentum of 2 TeV. The results refer to the chambers within the first (*left*) and the second (*right*) station of the endcap. The Gaussian fit is performed in the region defined by $Mean \pm 2 \cdot RMS$.

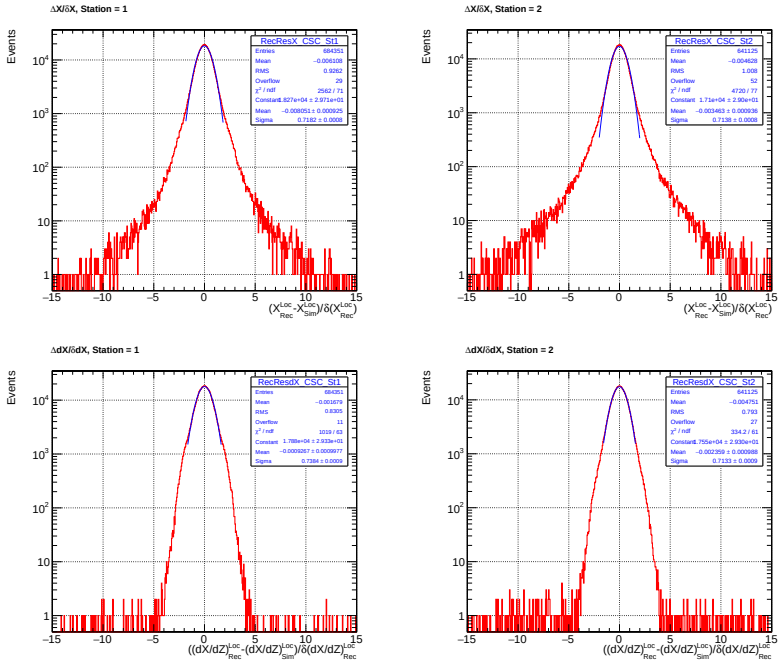

Figure 5.36: Pull distributions for the x and dx/dz coordinates of the reconstructed segments associated to muons generated with nominal momentum of 2 TeV. The results refer to the chambers within the first (*left*) and the second (*right*) station of the endcap. The Gaussian fit is performed in the region defined by $Mean \pm 2 \cdot RMS$.

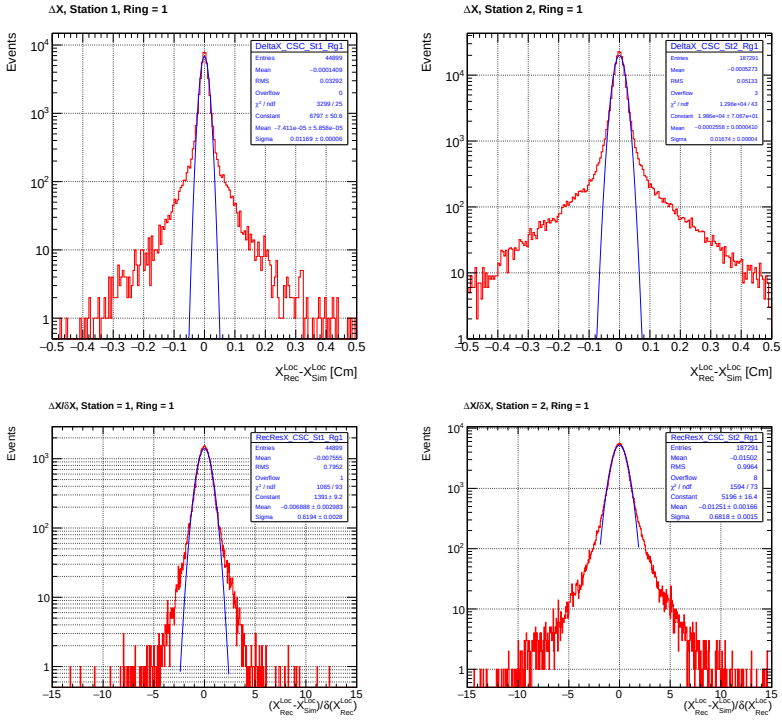

Figure 5.37: Resolution and pull distribution for the x coordinate of the reconstructed segments associated to muons generated with nominal momentum of 2 TeV. The results refer to the chambers of the first ring within the first (*left*) and the second (*right*) station of the endcap. The Gaussian fit is performed in the region defined by $Mean \pm 2 \cdot RMS$.

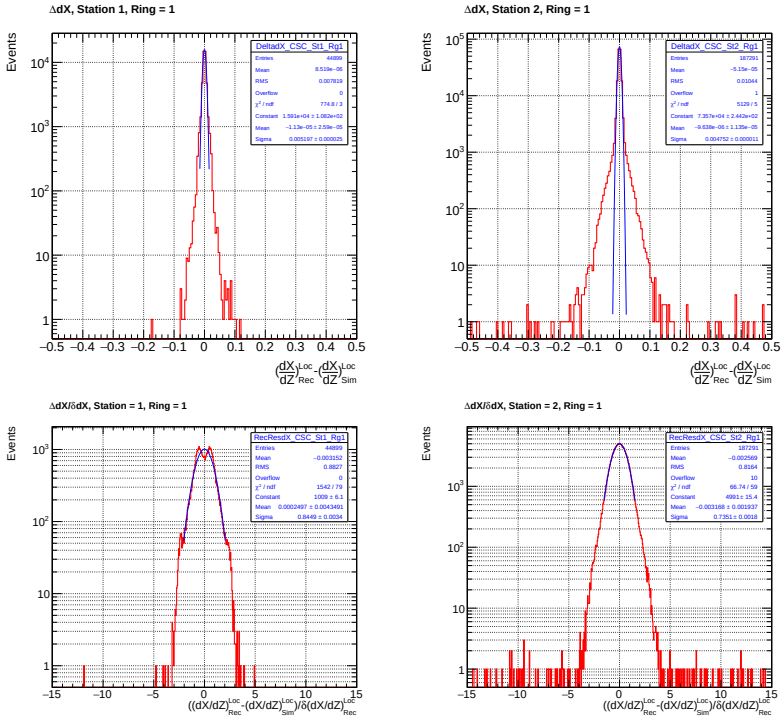

Figure 5.38: Resolution and pull distribution for the dx/dz direction of the reconstructed segments associated to muons generated with nominal momentum of 2 TeV. The results refer to the chambers of the first ring within the first (*left*) and the second (*right*) station of the endcap. The Gaussian fit is performed in the region defined by $Mean \pm 2 \cdot RMS$.

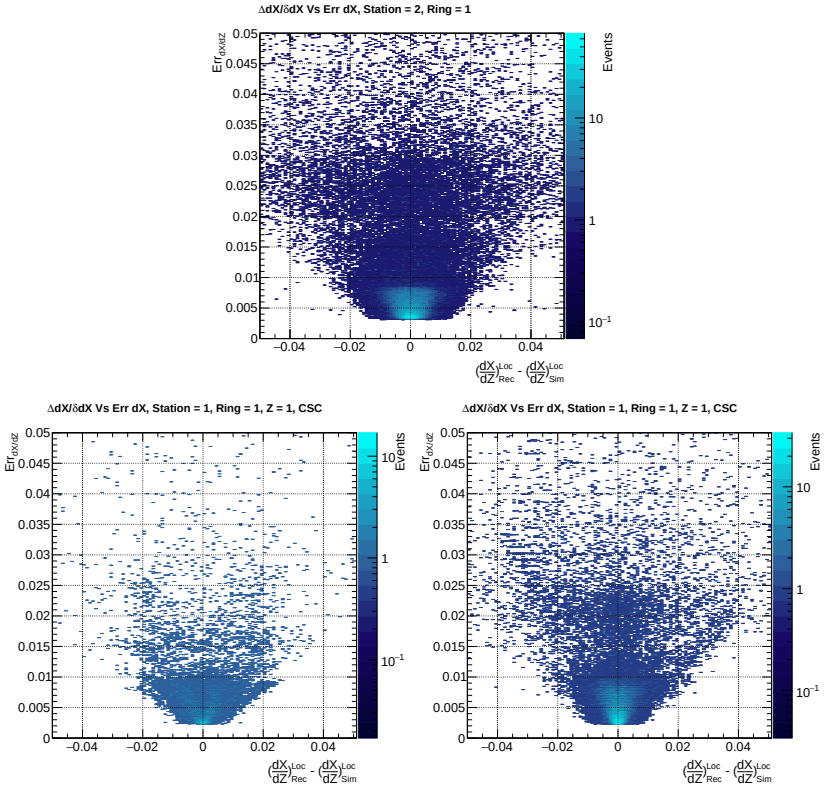

Figure 5.39: Correlation distributions for the dx/dz direction and its error for the segments associated to muon generated with nominal momentum of 2 TeV. The results refer to the chambers of the first ring within the first (*bottom*) and the second (*top*) stations of the endcap identified by $z = 1$. The chambers of the first ring within the first station are further divided in the ME1/1A (*left*) and M1/1B (*right*) chambers.

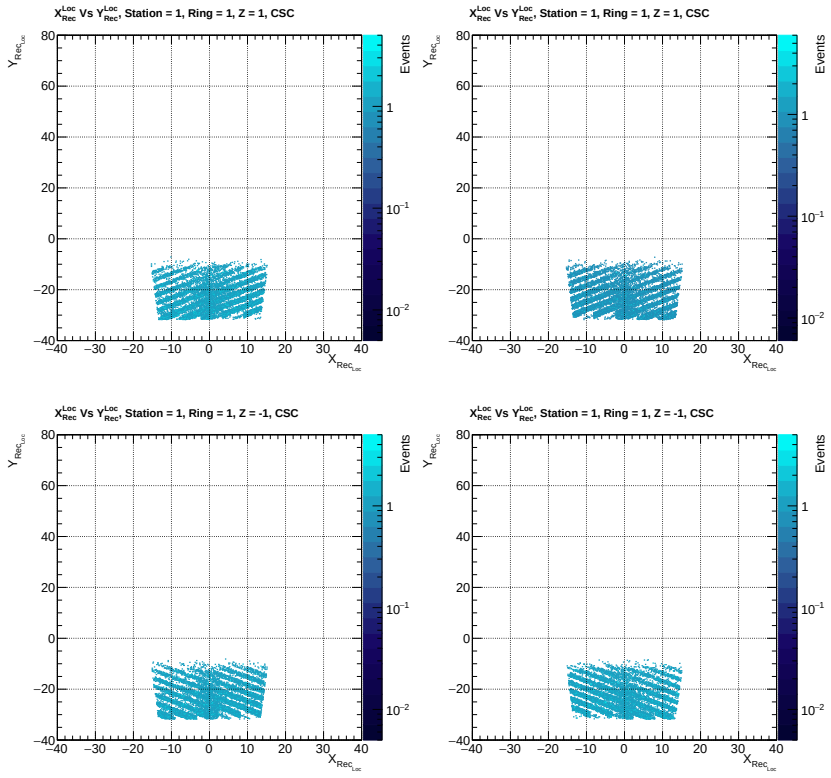

Figure 5.40: Distributions of the local x and y coordinates of the reconstructed segments associated to muon generated with nominal momentum of 2 TeV. The results refer to ME1/1A chambers in the two endcaps region identified by $z = \pm 1$. The distributions report two cases: the x and y coordinates in the case the resolution of the corrispective dx/dz direction is positive (*right*) and in the case it is negative (*left*).

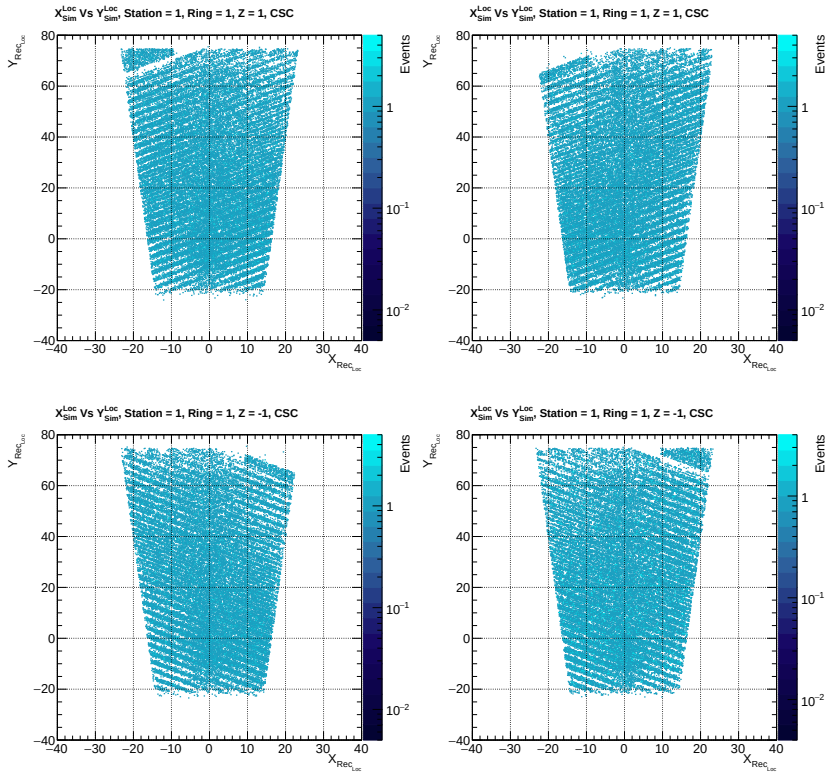

Figure 5.41: Distributions of the local x and y coordinates of the reconstructed segments associated to muon generated with nominal momentum of 2 TeV. The results refer to ME1/1B chambers in the two endcaps region identified by $z = \pm 1$. The distributions report two cases: the x and y coordinates in the case the resolution of the correspective dx/dz direction is positive (*right*) and in the case it is negative (*left*).

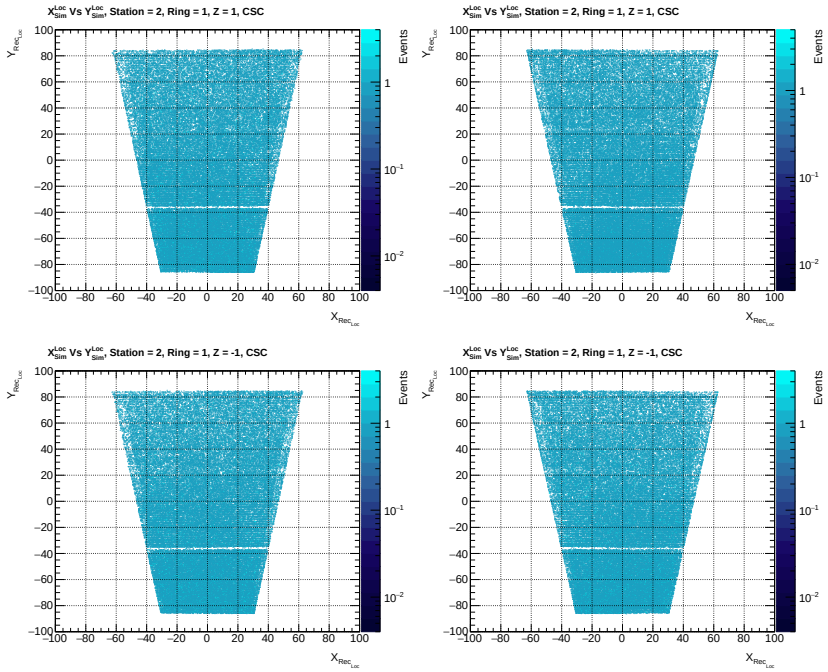

Figure 5.42: Distributions of the local x and y coordinates of the reconstructed segments associated to muon generated with nominal momentum of 2 TeV. The results refer to the first ring within the second chamber of the two endcaps region identified by $z = \pm 1$. The distribution report two cases: the x and y coordinates in the case the resolution of the correspective dx/dz direction is positive (*right*) and in the case it is negative (*left*).

Conclusion

The 2012 marks a milestone in high energy physics, indeed both the ATLAS and the CMS experiment announced the discovery of the Higgs boson, closing a chapter opened fifty years before. Thanks to this dowel, the Standard Model is finally complete, however a lot of experimental and theoretical inputs show that the SM is not yet the ultimate model of the *infinitely small* and reclaim its extension.

The analysis presented in this thesis has covered a big part of the PhD work and it was driven by the desire to contribute at this new challenge. Motivated by some theoretical predictions, which try to extend the SM with the introduction of a new scalar particles as in the 2HDM, the search has targeted a generic scalar resonance in the mass range between 200 GeV and 3 TeV using the $ZZ \rightarrow 2l2\nu$ decay channel. The work was done in collaboration with the "Université Libre de Bruxelles", the "University of Delhi", and the "Beihang University". The results presented in this thesis were preceded by limits set on the Electroweak Singlet Model and the 2HDM. This work was performed by the same analysis group but using the 2015 data only. Their results were not presented in this manuscript but can be consulted here [47].

In addition to the development of the analysis framework, the main contribution to these results was the generation and the study of the signal, the implementation and validation of the reweigthing procedure to take into account the interference contributions between the signal and the background with MELA package, and finally the the development of the statistical model used in order to set limits on the production cross section of the heavy scalar resonance both in ggF and in VBF. The analysis has exploited the data collected by CMS experiment during the 2016 for a total luminosity of 35.9 fb^{-1}. No excess was measured in data and limits were set on the production cross section of an heavy scalar resonance in the ZZ channel decay both in gluon fusion and in vector boson fusion. The results were further combined with other two searches belonging to the ZZ family, $ZZ \rightarrow 4l$ and $ZZ \rightarrow 2l2q$, to set limits on the total production cross section of the ZZ channel. This work is public and its results are

available in this document [128]. The combined limits of the three different channels are reported in the Fig. 5.43. They show that the limits are guided by the $ZZ \rightarrow 4l$ final state in the low mass region below 500 GeV and by $ZZ \rightarrow 2l2q$ above. Apparently, this result is in contrast with the one obtained in the combination of the Di-boson channels performed by CMS at the end of the Run-I [71] where the $2l2\nu$ channel were guiding the limits in the very high mass region. The loss of efficiency in $2l2\nu$ is understandable considering that from one side this channel is strongly affected by the lack of data in the Transverse Mass shape region above 1 TeV, on the other side the improvement in the $2l2q$ channel resides in the introduction of a more performing selection up to values of mass of the order of 4 TeV which enhances the $ZZ \rightarrow 2l2q$ performances. Finally, the limits were set using a totally different benchmark: an heavy SM-Like scenario during Run-I, and a generic scalar boson with a generic width during Run-II. Furthermore, the $ZZ \rightarrow 2l2\nu$ can still reach a better performance also in the low mass region dominated by $ZZ \rightarrow 4l$ as long as a new preliminary ri-optimization is performed. The plots reported in the Fig. 5.44 point out exactly this aspect.

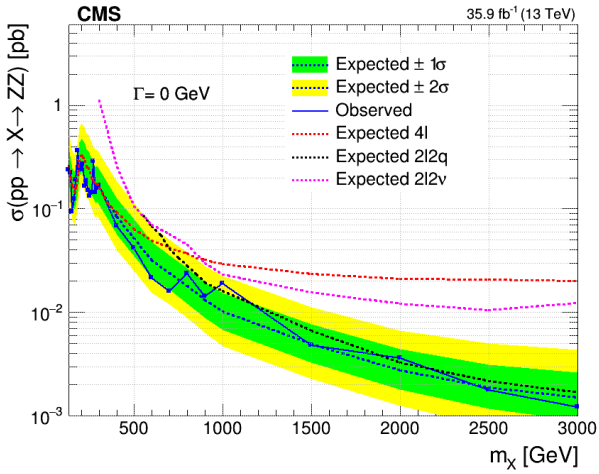

Figure 5.43: Combined limits on the Cross Section for a generic scalar boson with width of 0 GeV in ZZ channel. The results merged all the expected limits in each sub-channel.

In parallel with the work carried in the search mentioned above, the PhD working period was spent to develop and study the performances of one of the "Tev-Muons" refitters used in the CMS collaboration to reconstruct the track of high energetic muons, the *Dynamic Truncation Algorithm* (DYT). The performance of the algorithm are sig-

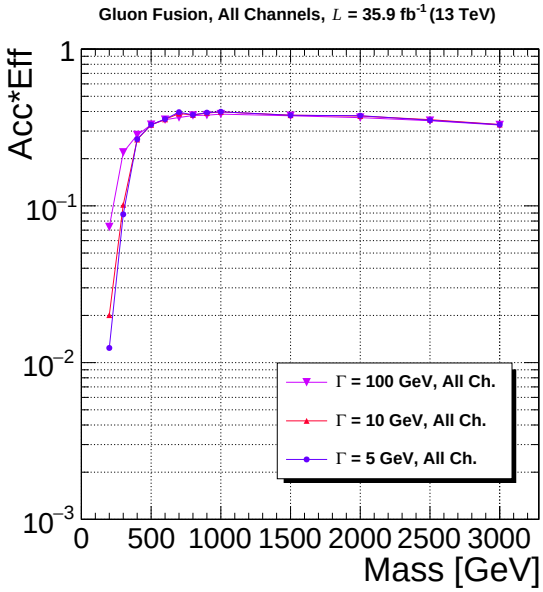

Figure 5.44: Efficiency of the signal selection as a function of the heavy scalar boson mass and its width. As reported in the distribution, the low mass region is affected by a low efficiency selection strategy that needs to be rioptimized.

nificantly improved in the Barrel region of the Muon spectrometer, whereas the situation in the EndCap didn't achieve the results expected.

The surprising results in the CSCs chambers obliged to a deeper understanding of the local reconstruction of the muon track in that region of the detector. The comparison between data and simulations allowed the identification of anomalies in the reconstruction of the segments in the first ring of the first station of the CSC detector. At the moment of the drafting of this manuscript it was not possible to find a valid solution to fix the problem and consequently to test the real improvement of the code in the forward region.

Lastly, the work done during this PhD period was crowned during the last year and a half by two different positions of responsibility inside the ZZ analysis group, one as the share main contact of the $ZZ \rightarrow 2l2\nu$ group analysis and the second one as responsible for the entire Monte Carlo generation of the ZZ group analysis.

Bibliography

[1] Vardan Khachatryan et al., "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV", *Eur. Phys. J.*, vol. C75, no. 5, pp. 212, 2015.

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

[3] Lyndon Evans, "The large hadron collider", *New Journal of Physics*, vol. 9, no. 9, pp. 335, 2007.

[4] S. L. Glashow, "Partial Symmetries of Weak Interactions", *Nucl. Phys.*, vol. 22, pp. 579–588, 1961.

[5] Steven Weinberg, "A Model of Leptons", *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, 1967.

[6] Abdus Salam, "Weak and Electromagnetic Interactions", *Conf. Proc.*, vol. C680519, pp. 367–377, 1968.

[7] G. Arnison et al., "Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV", *Phys. Lett.*, vol. 122B, pp. 103–116, 1983, [,611(1983)].

[8] M. Banner et al., "Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider", *Phys. Lett.*, vol. 122B, pp. 476–485, 1983.

[9] R. W. McAllister and R. Hofstadter, "Elastic scattering of 188-mev electrons from the proton and the alpha particle", *Phys. Rev.*, vol. 102, pp. 851–856, May 1956.

[10] E. D. Bloom, D. H. Coward, H. DeStaebler, J. Drees, G. Miller, L. W. Mo, R. E. Taylor, M. Breidenbach, J. I. Friedman, G. C. Hartmann, and H. W. Kendall, "High-energy inelastic $e-p$ scattering at 6° and 10°", *Phys. Rev. Lett.*, vol. 23, pp. 930–934, Oct 1969.

[11] T. Eichten et al., "Measurement of the neutrino-nucleon and antineutrino-nucleon total cross sections", *Physics Letters B*, vol. 46, no. 2, pp. 274 – 280, 1973.

[12] Murray Gell-Mann, "A Schematic Model of Baryons and Mesons", *Phys. Lett.*, vol. 8, pp. 214–215, 1964.

[13] G. Zweig, "An SU(3) model for strong interaction symmetry and its breaking. Version 1", 1964.

[14] H. Fritzsch, Murray Gell-Mann, and H. Leutwyler, "Advantages of the Color Octet Gluon Picture", *Phys. Lett.*, vol. 47B, pp. 365–368, 1973.

[15] G. Hanson et al., "Evidence for jet structure in hadron production by e^+e^- annihilation", *Phys. Rev. Lett.*, vol. 35, pp. 1609–1612, Dec 1975.

[16] R. Brandelik et al., "Evidence for Planar Events in e^+e^- Annihilation at High-Energies", *Phys. Lett.*, vol. 86B, pp. 243–249, 1979.

[17] Otto Eberhardt, Geoffrey Herbert, Heiko Lacker, Alexander Lenz, Andreas Menzel, Ulrich Nierste, and Martin Wiebusch, "Impact of a Higgs boson at a mass of 126 GeV on the standard model with three and four fermion generations", *Phys. Rev. Lett.*, vol. 109, pp. 241802, 2012.

[18] Eric Kuflik, Yosef Nir, and Tomer Volansky, "Implications of Higgs searches on the four generation standard model", *Phys. Rev. Lett.*, vol. 110, no. 9, pp. 091801, 2013.

[19] F. Abe et al., "Observation of top quark production in $\bar{p}p$ collisions", *Phys. Rev. Lett.*, vol. 74, pp. 2626–2631, 1995.

[20] S. Abachi et al., "Observation of the top quark", *Phys. Rev. Lett.*, vol. 74, pp. 2632–2637, 1995.

[21] Harald Fritzsch and Murray Gell-Mann, "Current algebra: Quarks and what else?", *eConf*, vol. C720906V2, pp. 135–165, 1972.

[22] Roel Aaij et al., "Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays", *Phys. Rev. Lett.*, vol. 115, pp. 072001, 2015.

[23] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons", *Phys. Rev. Lett.*, vol. 13, pp. 321–323, 1964.

[24] Peter W. Higgs, "Broken symmetries, massless particles and gauge fields", *Phys. Lett.*, vol. 12, pp. 132–133, 1964.

[25] Peter W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons", *Phys. Rev. Lett.*, vol. 13, pp. 508–509, 1964.

[26] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global Conservation Laws and Massless Particles", *Phys. Rev. Lett.*, vol. 13, pp. 585–587, 1964.

[27] Peter W. Higgs, "Spontaneous Symmetry Breakdown without Massless Bosons", *Phys. Rev.*, vol. 145, pp. 1156–1163, 1966.

[28] T. W. B. Kibble, "Symmetry breaking in nonAbelian gauge theories", *Phys. Rev.*, vol. 155, pp. 1554–1561, 1967.

[29] Emmy Noether, "Invariant Variation Problems", *Gott. Nachr.*, vol. 1918, pp. 235–257, 1918, [Transp. Theory Statist. Phys.1,186(1971)].

[30] Stephen P. Martin, "A Supersymmetry primer", 1997, [Adv. Ser. Direct. High Energy Phys.18,1(1998)].

[31] Tania Robens and Tim Stefaniak, "LHC Benchmark Scenarios for the Real Higgs Singlet Extension of the Standard Model", *Eur. Phys. J.*, vol. C76, no. 5, pp. 268, 2016.

[32] Rose Natalie Lerner and John McDonald, "Gauge singlet scalar as inflaton and thermal relic dark matter", *Phys. Rev.*, vol. D80, pp. 123507, 2009.

[33] Jorge de Blas, Marco Ciuchini, Enrico Franco, Satoshi Mishima, Maurizio Pierini, Laura Reina, and Luca Silvestrini, "Electroweak precision observables and Higgs-boson signal strengths in the Standard Model and beyond: present and future", *JHEP*, vol. 12, pp. 135, 2016.

[34] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, Marc Sher, and Joao P. Silva, "Theory and phenomenology of two-Higgs-doublet models", *Phys. Rept.*, vol. 516, pp. 1–102, 2012.

[35] Martin Bauer, Ulrich Haisch, and Felix Kahlhoefer, "Simplified dark matter models with two Higgs doublets: I. Pseudoscalar mediators", *JHEP*, vol. 05, pp. 138, 2017.

[36] Mark Trodden, "Electroweak baryogenesis", *Rev. Mod. Phys.*, vol. 71, pp. 1463–1500, 1999.

[37] John F. Gunion and Howard E. Haber, "Conditions for CP-violation in the general two-Higgs-doublet model", *Phys. Rev.*, vol. D72, pp. 095002, 2005.

[38] Jihn E. Kim, "Light Pseudoscalars, Particle Physics and Cosmology", *Phys. Rept.*, vol. 150, pp. 1–177, 1987.

[39] John F. Gunion and Howard E. Haber, "The CP conserving two Higgs doublet model: The Approach to the decoupling limit", *Phys. Rev.*, vol. D67, pp. 075019, 2003.

[40] Baradhwaj Coleppa, Felix Kling, and Shufang Su, "Constraining Type II 2HDM in Light of LHC Higgs Searches", *JHEP*, vol. 01, pp. 161, 2014.

[41] H. S. Cheon and Sin Kyu Kang, "Constraining parameter space in type-II two-Higgs doublet model in light of a 126 GeV Higgs boson", *JHEP*, vol. 09, pp. 085, 2013.

[42] P. M. Ferreira, Rui Santos, Marc Sher, and Joao P. Silva, "Could the LHC two-photon signal correspond to the heavier scalar in two-Higgs-doublet models?", *Phys. Rev.*, vol. D85, pp. 035020, 2012.

[43] Sanghyeon Chang, Sin Kyu Kang, Jong-Phil Lee, Kang Young Lee, Seong Chan Park, and Jeonghyeon Song, "Comprehensive study of two Higgs doublet model in light of the new boson with mass around 125 GeV", *JHEP*, vol. 05, pp. 075, 2013.

[44] Cheng-Wei Chiang and Kei Yagyu, "Implications of Higgs boson search data on the two-Higgs doublet models with a softly broken Z_2 symmetry", *JHEP*, vol. 07, pp. 160, 2013.

[45] LEP Electroweak Working Group, "Precision Electroweak Measurements and Constraints on the Standard Model", 2010.

[46] Robert V. Harlander, Stefan Liebler, and Hendrik Mantler, "SusHi: A program for the calculation of Higgs production in gluon fusion and bottom-quark annihilation in the Standard Model and the MSSM", *Comput. Phys. Commun.*, vol. 184, pp. 1605–1617, 2013.

[47] "Search for a heavy scalar boson decaying into a pair of Z bosons in the $2\ell 2\nu$ final state", Tech. Rep. CMS-PAS-HIG-16-001, CERN, Geneva, 2016.

[48] S. Dittmaier et al., "Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables", 2011.

[49] Vardan Khachatryan et al., "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV", *Eur. Phys. J. C*, vol. 75, no. CMS-HIG-14-009. CERN-PH-EP-2014-288. CMS-HIG-14-009, pp. 212. 75 p, Dec 2014.

[50] J. Stirling., "Private communication - hep, imperial", Tech. Rep.

[51] O.P. Pavlenko, A.M. Snigirev, and G.M. Zinovjev, "The gribov-lipatov relation in perturbative qcd", *Physics Letters B*, vol. 126, no. 3, pp. 267 – 270, 1983.

[52] L. N. Lipatov, "The parton model and perturbation theory", *Sov. J. Nucl. Phys.*, vol. 20, pp. 94–102, 1975, [Yad. Fiz.20,181(1974)].

[53] G. Altarelli and G. Parisi, "Asymptotic freedom in parton language", *Nuclear Physics B*, vol. 126, no. 2, pp. 298 – 318, 1977.

[54] Yuri L. Dokshitzer, "Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics.", *Sov. Phys. JETP*, vol. 46, pp. 641–653, 1977, [Zh. Eksp. Teor. Fiz.73,1216(1977)].

[55] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, "Parton distributions for the LHC", *Eur. Phys. J.*, vol. C63, pp. 189–285, 2009.

[56] John C. Collins, Davison E. Soper, and George F. Sterman, "Factorization of Hard Processes in QCD", *Adv. Ser. Direct. High Energy Phys.*, vol. 5, pp. 1–91, 1989.

[57] Tania Robens and Tim Stefaniak, "Status of the Higgs Singlet Extension of the Standard Model after LHC Run 1", *Eur. Phys. J.*, vol. C75, pp. 104, 2015.

[58] Suyong Choi, Sunghoon Jung, and P. Ko, "Implications of LHC data on 125 GeV Higgs-like boson for the Standard Model and its various extensions", *JHEP*, vol. 10, pp. 225, 2013.

[59] Csaba Csaki, "The Minimal supersymmetric standard model (MSSM)", *Mod. Phys. Lett.*, vol. A11, pp. 599, 1996.

[60] D. de Florian et al., "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector", 2016.

[61] Paolo Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP*, vol. 11, pp. 040, 2004.

[62] Stefano Frixione, Paolo Nason, and Carlo Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method", *JHEP*, vol. 11, pp. 070, 2007.

[63] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP*, vol. 06, pp. 043, 2010.

[64] E. Bagnaschi, G. Degrassi, P. Slavich, and A. Vicini, "Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM", *JHEP*, vol. 02, pp. 088, 2012.

[65] Paolo Nason and Carlo Oleari, "NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG", *JHEP*, vol. 02, pp. 037, 2010.

[66] Massimiliano Grazzini, "NNLO predictions for the Higgs boson signal in the $H \to WW \to l\nu l\nu$ and $H \to ZZ \to 4l$ decay channels", *JHEP*, vol. 02, pp. 043, 2008.

[67] Stefano Catani and Massimiliano Grazzini, "An NNLO subtraction formalism in hadron collisions and its application to Higgs boson production at the LHC", *Phys. Rev. Lett.*, vol. 98, pp. 222002, 2007.

[68] Massimiliano Grazzini and Hayk Sargsyan, "Heavy-quark mass effects in Higgs boson production at the LHC", *JHEP*, vol. 09, pp. 129, 2013.

[69] Michael Spira, "Higgs Boson Production and Decay at Hadron Colliders", *Prog. Part. Nucl. Phys.*, vol. 95, pp. 98–159, 2017.

[70] A. Denner, S. Heinemeyer, I. Puljak, D. Rebuzzi, and M. Spira, "Standard Model Higgs-Boson Branching Ratios with Uncertainties", *Eur. Phys. J.*, vol. C71, pp. 1753, 2011.

[71] Vardan Khachatryan et al., "Search for a Higgs boson in the mass range from 145 to 1000 GeV decaying to a pair of W or Z bosons", *JHEP*, vol. 10, pp. 144, 2015.

[72] Nikolas Kauer, "Interference effects for H $\to$ WW/ZZ $\to \ell\bar{\nu}_\ell\bar{\ell}\nu_\ell$ searches in gluon fusion at the LHC", *JHEP*, vol. 12, pp. 082, 2013.

[73] Nikolas Kauer and Claire O'Brien, "Heavy Higgs signal-background interference in $gg \to VV$ in the Standard Model plus real singlet", *Eur. Phys. J.*, vol. C75, pp. 374, 2015.

[74] John M. Campbell and R. K. Ellis, "MCFM for the Tevatron and the LHC", *Nucl. Phys. Proc. Suppl.*, vol. 205-206, pp. 10–15, 2010.

[75] John M. Campbell, R. Keith Ellis, and Ciaran Williams, "Bounding the Higgs width at the LHC using full analytic results for $gg \to e^- e^+ \mu^- \mu^+$", *JHEP*, vol. 04, pp. 060, 2014.

[76] Yanyan Gao, Andrei V. Gritsan, Zijin Guo, Kirill Melnikov, Markus Schulze, and Nhan V. Tran, "Spin determination of single-produced resonances at hadron colliders", *Phys. Rev.*, vol. D81, pp. 075022, 2010.

[77] Sara Bolognesi, Yanyan Gao, Andrei V. Gritsan, Kirill Melnikov, Markus Schulze, Nhan V. Tran, and Andrew Whitbeck, "On the spin and parity of a single-produced resonance at the LHC", *Phys. Rev.*, vol. D86, pp. 095031, 2012.

[78] Ian Anderson et al., "Constraining anomalous HVV interactions at proton and lepton colliders", *Phys. Rev.*, vol. D89, no. 3, pp. 035007, 2014.

[79] Andrei V. Gritsan, Raoul Röntsch, Markus Schulze, and Meng Xiao, "Constraining anomalous Higgs boson couplings to the heavy flavor fermions using matrix element techniques", *Phys. Rev.*, vol. D94, no. 5, pp. 055023, 2016.

[80] "Measurements of properties of the Higgs boson and search for an additional resonance in the four-lepton final state at $\sqrt{s} = 13$ TeV", Tech. Rep. CMS-PAS-HIG-16-033, CERN, Geneva, 2016.

[81] "Search for new diboson resonances in the dilepton + jets final state at $\sqrt{s} = 13$ TeV with 2016 data", Tech. Rep. CMS-PAS-HIG-16-034, CERN, Geneva, 2017.

[82] "Search for high mass Higgs to WW with fully leptonic decays using 2015 data", Tech. Rep. CMS-PAS-HIG-16-023, CERN, Geneva, 2016.

[83] "Search for new physics in high mass diphoton events in 3.3 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV and combined interpretation of searches at 8 TeV and 13 TeV", Tech. Rep. CMS-PAS-EXO-16-018, CERN, Geneva, 2016.

[84] "Search for high-mass resonances in $Z\gamma \to e^+e^-\gamma/\mu^+\mu^-\gamma$ final states in proton-proton collisions at $\sqrt{s} = 13$ TeV", Tech. Rep. CMS-PAS-EXO-16-034, CERN, Geneva, 2016.

[85] G. L. Bayatian et al., "CMS physics: Technical design report", 2006.

[86] S. Chatrchyan et al., "The CMS Experiment at the CERN LHC", *JINST*, vol. 3, pp. S08004, 2008.

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

[88] Oliver Sim Brüning, Paul Collier, P Lebrun, Stephen Myers, Ranko Ostojic, John Poole, and Paul Proudlock, *LHC Design Report*, CERN Yellow Reports: Monographs. CERN, Geneva, 2004.

[89] G. Aad et al., "The ATLAS Experiment at the CERN Large Hadron Collider", *JINST*, vol. 3, pp. S08003, 2008.

[90] A. Augusto Alves, Jr. et al., "The LHCb Detector at the LHC", *JINST*, vol. 3, pp. S08005, 2008.

[91] K. Aamodt et al., "The ALICE experiment at the CERN LHC", *JINST*, vol. 3, pp. S08002, 2008.

[92] Fabienne Marcastel, "CERN's Accelerator Complex. La chaîne des accélérateurs du CERN", Oct 2013, General Photo.

[93] P. Lujan, "Public cms luminosity information", Twiki page, 2017.

[94] S Chatrchyan et al., "Precise Mapping of the Magnetic Field in the CMS Barrel Yoke using Cosmic Rays", *JINST*, vol. 5, pp. T03021, 2010.

[95] *The CMS muon project: Technical Design Report*, Technical Design Report CMS. CERN, Geneva, 1997.

[96] Thomas Lenzi, "Development and Study of Different Muon Track Reconstruction Algorithms for the Level-1 Trigger for the CMS Muon Upgrade with GEM Detectors", Master's thesis, U. Brussels (main), 2013.

[97] Serguei Chatrchyan et al., "The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC", *JINST*, vol. 8, pp. P11002, 2013.

[98] Serguei Chatrchyan et al., "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST*, vol. 9, no. 10, pp. P10009, 2014.

[99] R. Fruhwirth, "Application of Kalman filtering to track and vertex fitting", *Nucl. Instrum. Meth.*, vol. A262, pp. 444–450, 1987.

[100] J. R. Cash and Alan H. Karp, "A variable order runge-kutta method for initial value problems with rapidly varying right-hand sides", *ACM Trans. Math. Softw.*, vol. 16, no. 3, pp. 201–222, September 1990.

[101] K. Rose, "Deterministic annealing for clustering, compression, classification, regression and related optimisation problems", *Proceedings of the IEEE*, pp. 2210 – 2239.

[102] Alan James Bell, "The design and construction of the beam scintillation counter for CMS", 2008, Presented on 2008.

[103] C. Ohm, "Phase and Intensity Monitoring of the Particle Beams at the ATLAS Experiment", Master's thesis, Linkoping U., 2007.

[104] E. Gatt et al., "Analysis of the position resolution in centroid measurements in mwpc", *Nucl.Instrum. and Methods*, vol. 188, pp. 327–346, 1981.

[105] Giacomo Bruno, Peter Timothy Cox, Stefano Lacaprara, Norbert Neumeister, Bart Van de Vyver, Stefano Villa, and Richard Wilkinson, "Local Reconstruction in the Muon Detectors", Tech. Rep. CMS-NOTE-2002-043, CERN, Geneva, Nov 2002.

[106] A. M. Sirunyan et al., "Particle-flow reconstruction and global event description with the cms detector", *JINST*, vol. 12, pp. P10003, 2017.

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

[108] Vardan Khachatryan et al., "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s}$ = 8 TeV", *JINST*, vol. 10, no. 06, pp. P06005, 2015.

[109] W Adam, R Frühwirth, A Strandlie, and T Todorov, "Reconstruction of electrons with the gaussian-sum filter in the cms tracker at the lhc", *Journal of Physics G: Nuclear and Particle Physics*, vol. 31, no. 9, pp. N9, 2005.

[110] Andreas Hocker et al., "TMVA - Toolkit for Multivariate Data Analysis", *PoS*, vol. ACAT, pp. 040, 2007.

[111] Vardan Khachatryan et al., "Performance of Photon Reconstruction and Identification with the CMS Detector in Proton-Proton Collisions at sqrt(s) = 8 TeV", *JINST*, vol. 10, no. 08, pp. P08010, 2015.

[112] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez, "The Anti-k(t) jet clustering algorithm", *JHEP*, vol. 04, pp. 063, 2008.

[113] "Jet Energy Corrections determination at 7 TeV", Tech. Rep. CMS-PAS-JME-10-010, CERN, Geneva, 2010.

[114] Matteo Cacciari and Gavin P. Salam, "Pileup subtraction using jet areas", *Phys. Lett.*, vol. B659, pp. 119–126, 2008.

[115] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez, "The Catchment Area of Jets", *JHEP*, vol. 04, pp. 005, 2008.

[116] "Pileup Jet Identification", Tech. Rep. CMS-PAS-JME-13-005, CERN, Geneva, 2013.

[117] Henning Kirschenmann, "Jet performance in CMS", Tech. Rep. CMS-CR-2013-325, CERN, Geneva, Oct 2013.

[118] "Jet energy scale and resolution performances with 13 TeV data".

[119] "Identification of b quark jets at the CMS Experiment in the LHC Run 2", Tech. Rep. CMS-PAS-BTV-15-001, CERN, Geneva, 2016.

[120] "Performance of heavy flavour identification algorithms in proton-proton collisions at 13 TeV at the CMS experiment", May 2017.

[121] The CMS collaboration, "Performance of the cms missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8$ tev", *Journal of Instrumentation*, vol. 10, no. 02, pp. P02006, 2015.

[122] "Performance of Jet Algorithms in CMS", Tech. Rep. CMS-PAS-JME-07-003, CERN, Geneva.

[123] "Studies of Tracker Material", Tech. Rep. CMS-PAS-TRK-10-003, 2010.

[124] "Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector", Tech. Rep. CMS-PAS-JME-16-004, CERN, Geneva, 2016.

[125] D. Bertolini, P. Harris, M. Low, and N. Tran, "Pileup per Particle Identification", *JHEP*, vol. 10, pp. 059, 2014.

[126] "Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV", 2010.

[127] G. Petrucciani, A. Rizzi, and C. Vuosalo, "Mini-AOD: A New Analysis Data Format for CMS", *J. Phys. Conf. Ser.*, vol. 664, no. 7, pp. 072052, 2015.

[128] Albert M Sirunyan et al., "Search for a new scalar resonance decaying to a pair of Z bosons in proton-proton collisions at $\sqrt{s} = 13$ TeV", 2018.

[129] A. J. Barr, T. J. Khoo, P. Konar, K. Kong, C. G. Lester, K. T. Matchev, and M. Park, "Guide to transverse projections and mass-constraining variables", *Phys. Rev.*, vol. D84, pp. 095031, 2011.

[130] Ciulli V., "Summary table of samples produced for the 1 billion campaign, with 25 ns bunch-crossing", Twiki page, 2017.

[131] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP*, vol. 07, pp. 079, 2014.

[132] Emanuele Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method", *Eur. Phys. J.*, vol. C71, pp. 1547, 2011.

[133] Tom Melia, Paolo Nason, Raoul Rontsch, and Giulia Zanderighi, "W^+W^-, WZ and ZZ production in the POWHEG BOX", *JHEP*, vol. 11, pp. 078, 2011.

[134] Barbara Jager, Andreas von Manteuffel, and Stephan Thier, "Slepton pair production in the POWHEG BOX", *JHEP*, vol. 10, pp. 130, 2012.

[135] Paolo Nason and Giulia Zanderighi, "W^+W^- , WZ and ZZ production in the POWHEG-BOX-V2", *Eur. Phys. J.*, vol. C74, no. 1, pp. 2702, 2014.

[136] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands, "A Brief Introduction to PYTHIA 8.1", *Comput. Phys. Commun.*, vol. 178, pp. 852–867, 2008.

[137] Rikkert Frederix and Stefano Frixione, "Merging meets matching in MC@NLO", *JHEP*, vol. 12, pp. 061, 2012.

[138] Johan Alwall et al., "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions", *Eur. Phys. J.*, vol. C53, pp. 473–500, 2008.

[139] Richard D. Ball et al., "Parton distributions for the LHC Run II", *JHEP*, vol. 04, pp. 040, 2015.

[140] S. Agostinelli et al., "GEANT4: A Simulation toolkit", *Nucl.Instrum.Meth.*, vol. A506, pp. 250–303, 2003.

[141] Simone Alioli, Keith Hamilton, Paolo Nason, Carlo Oleari, and Emanuele Re, "Jet pair production in POWHEG", *JHEP*, vol. 04, pp. 081, 2011.

[142] The Egamma ID Group, "Tag and probe methodology for analyses using electrons and photons", CMS Note 2012/146, 2012.

[143] Brun H., "Muon HLT Efficiency Measurement with the "Reference-Trigger" Method", Twiki page, 2013.

[144] Serguei Chatrchyan et al., "Missing transverse energy performance of the CMS detector", *JINST*, vol. 6, pp. P09001, 2011.

[145] Artur Apresyan et all., "Identification and mitigation of anomalous signals in cms hadronic calorimeter", *Journal of Physics: Conference Series*, vol. 404, no. 1, pp. 012044, 2012.

[146] Dalfonso, M. and others, "Met optional filters for run ii", Twiki, 2015.

[147] L.. Soffi, "Cut based electron id for run 2", Twiki page, 2017.

[148] J. Bendavid, "Tools for conversion rejection (electron id) and electron vetoing (photon id)", Twiki page, 2012.

[149] I. Kravchenko, "Electron efficiencies and scale factors", Twiki page, 2017.

[150] "Search for diboson resonances in the $2l2\nu$ final state", Tech. Rep. CMS-PAS-B2G-16-023, CERN, Geneva, 2017.

[151] I. Kravchenko et al., "Cut based photon id for run 2", Twiki page, 2015.

[152] Nicolas Postiau and Pascal Vanlaer, "γ-induced electroweak corrections to the $pp \to w^{\pm}z + x$ process at $\sqrt{s} = 13\,\text{TeV}$, and associated uncertainties", CMS Note 2016/431, 2016.

[153] Anastasiya Bierweiler, Tobias Kasprzik, and Johann H. Kühn, "Vector-boson pair production at the LHC to $\mathcal{O}(\alpha^3)$ accuracy", *JHEP*, vol. 12, pp. 071, 2013.

[154] Stefan Gieseke, Tobias Kasprzik, and Johann H. Kühn, "Vector-boson pair production and electroweak corrections in HERWIG++", *Eur. Phys. J.*, vol. C74, no. 8, pp. 2988, 2014.

[155] Aneesh Manohar, Paolo Nason, Gavin P. Salam, and Giulia Zanderighi, "How bright is the proton? A precise determination of the photon parton distribution function", *Phys. Rev. Lett.*, vol. 117, no. 24, pp. 242002, 2016.

[156] S. Frixione, V. Hirschi, D. Pagani, H. S. Shao, and M. Zaro, "Weak corrections to Higgs hadroproduction in association with a top-quark pair", *JHEP*, vol. 09, pp. 065, 2014.

[157] S. Frixione, V. Hirschi, D. Pagani, H. S. Shao, and M. Zaro, "Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons", *JHEP*, vol. 06, pp. 184, 2015.

[158] "Measurement of the differential cross section of Z boson production in association with jets in proton-proton collisions at $\sqrt{s} = 13$ TeV", Tech. Rep. CMS-PAS-SMP-15-010, CERN, Geneva, 2015.

[159] Jon Butterworth et al., "PDF4LHC recommendations for LHC Run II", *J. Phys.*, vol. G43, pp. 023001, 2016.

[160] Iain W. Stewart and Frank J. Tackmann, "Theory Uncertainties for Higgs and Other Searches Using Jet Bins", *Phys.Rev.*, vol. D85, pp. 034011, 2012.

[161] John M. Campbell, R. Keith Ellis, and Ciaran Williams, "Vector boson pair production at the LHC", *JHEP*, vol. 07, pp. 018, 2011.

[162] "CMS Luminosity Measurements for the 2016 Data Taking Period", Tech. Rep. CMS-PAS-LUM-17-001, CERN, Geneva, 2017.

[163] J.. Piedra, "Rochester correction", Twiki page, 2017.

[164] "Missing Transverse Energy Performance in Minimum-Bias and Jet Event from Proton-Proton Collisions at $\sqrt{s}$", Tech. Rep. CMS-PAS-JME-10-004, CERN, Geneva, 2010.

[165] M Marionneau, "Uncertainties and met : Technicalities", `https://indico.cern.ch/event/510570/contributions/1190890/attachments/1246661/1836294/JME_METUnc_210316.pdf`, 2016.

[166] Serguei Chatrchyan et al., "Measurement of the wγ and zγ inclusive cross sections in pp collisions at $\sqrt{s} = 7$ tev and limits on anomalous triple gauge boson couplings", *Phys. Rev. D*, vol. 89, pp. 092005, May 2014.

[167] "Measurement of inclusive W and Z boson production cross sections in pp collisions at $\sqrt(s) = 13$ TeV", Tech. Rep. CMS-PAS-SMP-15-004, CERN, Geneva, 2015.

[168] Ansgar Denner, Stefan Dittmaier, Markus Hecht, and Christian Pasold, "NLO QCD and electroweak corrections to $W + \gamma$ production with leptonic W-boson decays", *JHEP*, vol. 04, pp. 018, 2015.

[169] Ansgar Denner, Stefan Dittmaier, Markus Hecht, and Christian Pasold, "NLO QCD and electroweak corrections to $Z + \gamma$ production with leptonic Z-boson decays", *JHEP*, vol. 02, pp. 057, 2016.

[170] A L Read, "Presentation of search results: the cl(s) technique", *Journal of Physics G: Nuclear and Particle Physics*, vol. 28, no. 10, pp. 2693, 2002.

[171] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J.*, vol. C71, pp. 1554, 2011, [Erratum: Eur. Phys. J.C73,2501(2013)].

[172] J. S. Conway, "Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra", in *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland 17-20 January 2011*, 2011, pp. 115–120.

[173] Georges Aad et al., "Constraints on the off-shell Higgs boson signal strength in the high-mass ZZ and WW final states with the ATLAS detector", *Eur. Phys. J.*, vol. C75, no. 7, pp. 335, 2015.

[174] C. Patrignani et al., "Review of Particle Physics", *Chin. Phys.*, vol. C40, no. 10, pp. 100001, 2016.

[175] K. A. Olive et al., "Review of Particle Physics", *Chin. Phys.*, vol. C38, pp. 090001, 2014.

[176] Grzegorz Wrochna, "Muon Trigger of the CMS Detector for LHC", Tech. Rep. CMS-NOTE-1997-096, CERN, Geneva, Nov 1997.

[177] CMS Collaboration, "Measurement of the muon stopping power in lead tungstate", *Journal of Instrumentation*, vol. 5, no. 03, pp. P03007, 2010.

[178] G. Acquistapace et al., "CMS, the magnet project: Technical design report", 1997.

[179] J. Yarba, "Particle Guns", Twiki page, 2010.