MCCXXII

UNIVERSITÀ DEGLI STUDI DI PADOVA
SEDE AMMINISTRATIVA: UNIVERSITÀ DEGLI STUDI DI PADOVA
DIPARTIMENTO DI FISICA E ASTRONOMIA "G. GALILEI"

SCUOLA DI DOTTORATO DI RICERCA IN FISICA
CICLO XVII

TESI DI DOTTORATO

Searches for signatures of an extended Higgs sector in final states with leptons and Higgs to $b\bar{b}$ decays at CMS

DIRETTORE DELLA SCUOLA: Ch.mo Prof. Andrea Vitturi

SUPERVISORE: Dott. Tommaso Dorigo

CO-SUPERVISORE: Dott. Stefano Lacaprara

DOTTORANDO: Alberto Zucchetta

CONTENTS

Riassunto

Il modello standard (SM) delle particelle elementari, verificato sperimentalmente ad alta precisione, prevede l'esistenza di una particella scalare, il bosone di Higgs, grazie al quale avviene la rottura di simmetria elettrodebole ed è possibile spiegare le masse non nulle dei fermioni. La ricerca del bosone di Higgs è stata perseguita in passato da diversi esperimenti ai collisori LEP II e Tevatron, ma sempre con esito negativo. Solo nell'anno 2011, all'acceleratore LHC gli esperimenti ATLAS ed CMS hanno cominciato ad osservare le prime evidenze dell'elusiva particella, per poi confermane definitivamente la scoperta nel 2012. Tuttavia, pur essendo l'esistenza e le caratteristiche osservate del bosone di Higgs l'ennesima verifica sperimentale della validità del modello standard, esistono fenomeni naturali che esso non è in grado di spiegare, come ad esempio la massa dei neutrini o la materia oscura. Diverse estensioni del modello standard sono state proposte; diverse tra le più accreditate prevedono non uno, ma più bosoni.

Con l'entrata a regime di LHC nel 2011 e le prime evidenze di un bosone di Higgs, una delle domande fondamentali a cui dare risposta era se il segnale che si stava osservando fosse dovuto al bosone di Higgs dello SM, oppure al più leggero dei diversi bosoni previsti dalle teorie oltre il modello standard. La risposta poteva venire sia dalla misura degli accoppiamenti dell'Higgs alle particelle dello SM, sia osservando processi e decadimenti non previsti dallo SM, o misurati con una frequenza maggiore dell'atteso. La produzione di un bosone di Higgs in associazione con quarks b, e il suo successivo decadimento in una coppia di quark b, è un processo difficilmente osservabile nel modello standard, pertanto la sua osservazione averebbe significato la presenza di nuova fisica. Precedenti ricerche a Tevatron hanno evidenziato un lieve eccesso in questo canale.

In seguito alla scoperta del bosone di Higgs nel Luglio 2012 da parte degli esperimenti CMS e ATLAS, la conoscenza delle proprietà di questa particella, ivi compresa la sua massa, diventano informazioni che possono essere utilizzate per aumentare la sensibilità a ricerche di processi di nuova fisica. Uno dei canali che beneficia di questa informazione è la ricerca di un bosone pseudoscalare A, in particolare nel suo decadimento in un bosone Z e un bosone di Higgs leggero h, che si assume essere quello di massa 125 GeV recentemente scoperto. Lo stato finale consiste in una coppia di elettroni o muoni originati dal decadimento del bosone Z, e una coppia di quark b dal decadimento dell'Higgs. Questo canale, non previsto dallo SM, permette di sondare una regione dello spazio dei parametri di teorie oltre il modello standard per certi versi complementare a quella del canale con molti b nello stato finale.

Questa tesi è il risultato del lavoro svolto in tre anni di analisi dati a CMS.

Nel Capitolo 1 viene introdotto brevemente il modello standard e il bosone di Higgs, senza tralasciare le ricerche sperimentali nel passato ai collisori e la sua scoperta a LHC.

Il Capitolo 2 spiega i motivi per cui il modello standard non può essere considerato la teoria definitiva, e sono illustrate alcune delle sue estensioni più accreditate che permettono di superarne alcune limitazioni.

Nel Capitolo 3 viene descritto l'apparato sperimentale, che consiste nell'acceleratore LHC e nel rivelatore CMS, e su come vengono ricostruite e misurate le particelle prodotte dalle collisioni.

Il Capitolo 4 descrive la ricerca di un bosone di Higgs in associazione con quarks b, e il suo successivo decadimento in una coppia di quark b con i dati 2011 di CMS. Il mio contributo al canale semileptonico di questa analisi consiste nello sviluppo delle tecniche di stima degli eventi di fondo con il metodo delle matrici di b-tagging, e sull'ottimizzazione del discriminatore multivariato per la selezione di una regione di controllo.

Il Capitolo 5 illustra la ricerca del processo $\mathrm{A} \to \mathrm{Zh} \to \ell\ell\mathrm{b\bar{b}}$, che mette a frutto le competenze acquisite con i jet da b quark e le teorie oltre il modello standard, e mi ha visto coinvolto nello sviluppo della strategia di analisi, in particolar modo nella caratterizzazione del segnale, nel controllo dei fondi attraverso regioni di controllo opportunamente definite, nel miglioramento della sensibilità con lo sviluppo e l'ottimizzazione di algoritmi multivariati, della stima delle incertezze sistematiche usate e dei limiti sulla sezione d'urto di produzione.

Abstract

The Standard Model (SM) of elementary particles, a model experimentally verified to an unprecedented level of accuracy, foresees the existence a scalar particle, the so-called Higgs boson, that breaks the electroweak symmetry and explains the non-null value of fermion masses. The pursue of the Higgs boson has been carried out by several experiments at the LEP II and Tevatron colliders in the past, but unsuccessfully. However, during the year 2011, the ATLAS and CMS experiments at LHC began to observe some evidence of the elusive particle, finally confirming the discovering in 2012. Even if the Higgs boson and its observed properties represent a triumph for the Standard Model, there are many unresolved phenomena that the SM can not explain. Several extensions have been proposed; some of the most relevant foresee several Higgs bosons instead of only one.

After the commissioning of the LHC collider in 2011 and the first evidences of a light scalar boson, one of the fundamental questions to be answered was whether the new particle was the SM Higgs boson, or the lightest among the bosons foreseen by the theories extending the SM. The confirmation could come from the measurements of the Higgs couplings to the SM particles, or alternatively from the observation of new production processes and decays not predicted by the SM, or measured at a different rate. The associated production with b quarks of a Higgs boson, and its decay into pairs of b quarks, is a process with a small predicted cross section in the SM, and its observation would imply the presence of new physics. Previous searches at the Tevatron collider reported a slight excess in this channel.

Afterwards the discovery of the Higgs boson in July 2012 by the CMS and ATLAS experiments, the knowledge of the properties of the Higgs boson and its mass become a valuable handle to increase the sensitivity of new physics searches. One of the most beneficial channels is the search of a heavy pseudoscalar A, and its decay into a Z boson and a light Higgs boson h, the latter assumed to be the recently-discovered 125 GeV boson. The Z boson is sought in its decay into a pair of electrons or muons, and the Higgs into a pair of b quarks. This channel, not predicted by the SM, probes a region in the parameter phase-space of beyond the SM theories which is, to some extent, complementary to the one of the multi-b search.

This thesis is the result of three years of data analysis at CMS.

In Chapter 1 the standard model and the Higgs boson are introduced, mentioning the past searches at colliders and its discovery at LHC.

Chapter 2 explains why the SM can not be considered as the ultimate theory, and describes some of its most accredited extensions which allow to overcome some of the critical issues.

In Chapter 3 the experimental apparatus is described, consisting in the LHC accelerator and the CMS detector, focusing on the reconstruction and measurements of the particles produced by the collisions.

Chapter 4 describes the search for a Higgs boson decaying into a pair of b quarks and produced in association with other b quarks with 2011 data. My contribution to the semi-leptonic channel consists in the development of multijet background prediction techniques with the b-tagging matrix method, and in the optimization of the multivariate discriminator for the determination of a control region.

In Chapter 5 the search of the process $\mathrm{A} \to \mathrm{Zh} \to \ell\ell\mathrm{b}\bar{\mathrm{b}}$ is presented, capitalizing the expertise in b jets and beyond the SM theories acquired. I am involved in the planning of the analysis strategy, and particularly in the signal characterization, estimating the backgrounds in appropriate control regions, improving the analysis sensitivity developing and tuning multivariate algorithms, estimating the systematic uncertainties and setting the model-independent upper limits.

Chapter 1

The standard model and the Higgs boson

The known matter is capable of experiencing four kinds of fundamental interactions, which are electromagnetic, weak, strong, and gravitational. The first three are described by a single theory, the standard model (SM) of fundamental interactions, which provides a classification of the known particles and explains the interactions among them with an impressively high precision. The standard model also includes a mechanism of spontaneous symmetry breaking, called Higgs mechanism, due to the presence of a massive scalar particle, the Higgs boson, which explains how the mediator vectors of weak interaction gain masses. The Higgs boson was discovered at LHC in 2012, confirming once more the validity of the SM. This chapter describes the standard model of particle physics, the Higgs mechanism, and the experimental searches that led to the discovery of the Higgs boson.

1.1 THE STANDARD MODEL

The standard model (SM) of electroweak and strong interactions is the simplest theory describing the fundamental constituents of matter and their interactions in agreement with all known experimental data. The SM does not include gravity, but at the subatomic scale its effects can be neglected with respect to the electromagnetic, weak, and strong fundamental forces.

The SM is based on the symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$. The $SU(3)_C$ group is the basis of Quantum Chromo-Dynamics (QCD), which describes the strong interactions, and is assumed to be an exact symmetry. During the 1960s, Glashow [1], Weinberg [2], and Salam [3] combined electromagnetic interactions, based on the Quantum Electrodynamics Theory (QED) [4] with weak interactions into the $SU(2)_L \otimes U(1)_Y$ gauge theory, which describes all the electroweak interactions between matter fields by the exchange of massless vector bosons. The symmetry is not manifest, and is said to be spontaneously broken into $U(1)_{em}$. The electroweak gauge bosons are the mediator of the electromagnetic interactions, the massless and electrically-neutral photon (γ), and the massive mediators of the weak interactions, the neutral Z boson and the two charged $\mathrm{W}^{\pm}$ bosons. The mediators of the colour-flow are eight massless spin-one gauge bosons, called gluons, self-interacting and with a non-null colour charge.

In 1964, Gell-Mann [5] and Zweig [6] independently postulated the existence of the truly fundamental matter fields. These, named quarks, are fermions with $\frac{1}{2}\hbar$-spin and fractional charge, and could be used to describe the large number of observed particles at the time. The SM has six flavours of quarks (u, d, s, c, b, t) and six leptons (e, μ, τ, ν_e, ν_μ, ν_τ), together with their corresponding antiparticles, and the left- and right-handed components of the quark and lepton fields are assigned to different representations of the group because of the chiral structure of the weak interactions. Quarks, as leptons, are grouped in three generations, in doublets and singlets of a representation of the $SU(2)$ symmetry group. The matter constituents of the SM, and their quantum numbers under the SM symmetry group, are reported in Tab. 1.1.

Generation	First	Second	Third	Q	$U(1)_Y$	$SU(2)_L$	$SU(3)$
Leptons	$\begin{pmatrix}\nu_e \\ e\end{pmatrix}_L$	$\begin{pmatrix}\nu_\mu \\ \mu\end{pmatrix}_L$	$\begin{pmatrix}\nu_\tau \\ \tau\end{pmatrix}_L$	0 -1	$-1/2$	2	1
	ν_{eR}	$\nu_{\mu R}$	$\nu_{\tau R}$	0	0	1	1
	e_R	μ_R	τ_R	-1	-1	1	1
Quarks	$\begin{pmatrix}u \\ d\end{pmatrix}_L$	$\begin{pmatrix}c \\ s\end{pmatrix}_L$	$\begin{pmatrix}t \\ b\end{pmatrix}_L$	$+2/3$ $-1/3$	$+1/6$	2	3
	u_R	c_R	t_R	$+2/3$	$+2/3$	1	3
	d_R	s_R	b_R	$-1/3$	$-1/3$	1	3

Table 1.1: The SM fermions quantum numbers under the $SU(3) \times SU(2)_L \times U(1)_Y$ symmetry group, for the left- and right-handed components.

Quarks and antiquarks are the simplest representations of the $SU(3)_C$ group, and from their product follows a singlet bilinear and an octect, corresponding to two different bound states, *mesons* (built of a quark and an anti-quark $q\bar{q}$), and *baryons* (consist of three quarks qqq or three anti-quarks $\bar{q}\bar{q}\bar{q}$), collectively denoted as *hadrons*. In the high-energy regime, QCD is well described in terms of weakly-interacting quarks and gluons (the effect is called *asymptotic freedom*), allowing a perturbative description of the strong interactions. However, the strength of the attractive force between coloured particles increases with the distance, leading to non-perturbative dynamics (*quark confinement*). The characteristic energy scale at which these effects become important is $\Lambda_{QCD} \sim 200\,\text{MeV}$, approximately the light hadron mass scale. A coloured object produced in a collider, due to the colour confinement, cannot exist individually. Pairs of quarks and anti-quarks are produced from the vacuum and combined with the original quark until formation of colourless objects such as hadrons (*hadronization*). Since the energy scale at which these effects take place may be equal or lower than Λ_{QCD}, Monte Carlo models, tuned on experimental results, are usually applied in order to describe the hadronization process. The production of an high-energy coloured object results in a narrow cone of charged and neutral hadrons and other particles, called a *jet*.

1.2 THE ELECTROWEAK SYMMETRY BREAKING AND THE HIGGS BOSON

The invariance of the standard model Lagrangian under the electroweak symmetry group requires the presence of four gauge bosons. Furthermore, in the SM there is no way to describe massive gauge bosons without breaking the gauge invariance. The electromagnetic mediator, the photon, is massless, but on the other hand the limited range of the weak interactions implied that theirs mediators were massive. The observation of the $W^\pm$ and Z bosons confirmed that they are not massless [7].

In 1964, three independent groups of theorists proposed a mechanism through which a complex scalar field with non-zero vacuum expectation value was introduced into the Lagrangian, results in the breaking of electroweak symmetry. This mechanism would become known as the

Englert-Brout-Higgs-Guralnik-Hagen-Kibble mechanism, or the Higgs mechanism [8–10]. As a consequence of the electroweak symmetry breaking, a new scalar particle is posited, called the *Higgs boson*.

To see how the Higgs mechanism works mathematically, the Higgs scalar field, consisting in a complex $SU(2)$ doublet, can be written in the Lagrangian as:

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

The potential $V(\phi)$ of the Higgs sector is:

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \lambda |\phi^\dagger \phi|^2$$

choosing $\mu^2 > 0$ and $\lambda > 0$ the minimum of the potential is not found for a single value of ϕ, and the choice of the ground state is arbitrary. In other terms, after the choice of a state in the local minima of the potential, the system state is not symmetric anymore under $SU(2)$, and the chosen coordinate is not invariant under transformations in the (ϕ^+, ϕ^0) space; this is referred as the *spontaneous symmetry breaking*. By fixing the ground state to be along ϕ^0, the vacuum expectation of the Higgs field is:

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \qquad v = -\frac{\mu^2}{\lambda}$$

By rewriting the Φ field in a generic gauge in terms of its vacuum expectation value, Φ becomes:

$$\Phi = \frac{1}{\sqrt{2}} e^{\frac{i}{v} \phi^a t_a} \begin{pmatrix} 0 \\ h + v \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \phi_4 \end{pmatrix}, \qquad a = 1 \dots 3$$

Before the spontaneous symmetry breaking of the electroweak gauge symmetry group there are four massless scalar fields ϕ, known as Goldstone bosons [11], each one with one degree of freedom. After the symmetry breaking, only one real scalar field remains. The degree of freedom is token by the massive scalar Higgs boson, with mass $m_H = \sqrt{2\lambda v^2}$; since the parameter λ is unknown in the theory, the mass of the Higgs is an unknown parameter of the SM. The remaining three missing degrees of freedom give mass to the $W^\pm$ and Z bosons, which acquire a mass terms in the Lagrangian. Additionally, the fermions can be made to couple to the Higgs field, allowing them to have a mass term as well.

1.3 PHENOMENOLOGY OF THE HIGGS BOSON

1.3.1 SM HIGGS BOSON PRODUCTION

The standard model Higgs boson can be produced at hadron colliders mainly through four different processes, whose Feynman diagrams are reported in Fig. 1.1. Some are specific to hadronic colliders, some are possible both at hadronic and electron colliders. The relative magnitude and the cross section values are a function of Higgs mass, of the type of particles used at the collider and the center-of-mass energy.

Gluon-gluon fusion (ggF) : the gluon coupling to the Higgs boson in the SM is mediated by triangular loops of top and bottom quarks. This is the dominant production process

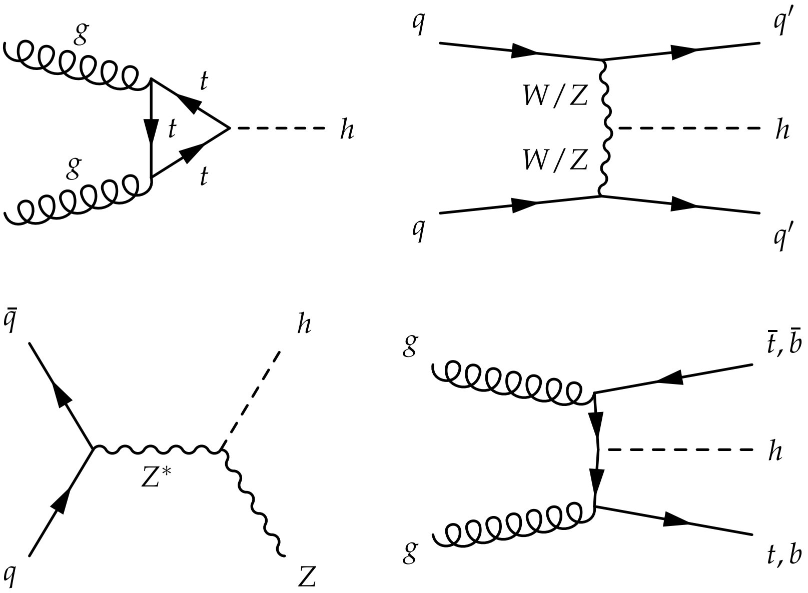

Figure 1.1: Feynman diagrams of the four Higgs production processes at hadron colliders. Clockwise from top left: gluon-gluon fusion, vector-boson fusion, associated production with vector bosons (or Higgsstrahlung), associated production with top or bottom quarks.

at hadron colliders, given the high partonic luminosity of the gluons. The proton remnants often escape undetected along the beampipe, leaving no visible signatures in the detector.

Vector-boson fusion (VBF) : the Higgs boson is produced through its couplings to vector bosons W and Z, emitted by the interacting quarks in the beam protons. Due to the mass of the mediators, a sizable fraction of the transverse momentum is transmitted to the hadronic fragments of the proton, originating a pair of energetic jets with a small angle with respect to the beam axis. The VBF cross section is about one order of magnitude lower than the gluon-gluon fusion, at least for Higgs masses up to 1 TeV. This production mechanism is viable also at electron-positron colliders.

Associated production with vector bosons (VH) : this production process is distinctive because it involves a massive SM particle that can yield a clear signature in the detector. The Higgs boson is emitted via a *Higgsstrahlung* from an off-shell W or Z boson. Even if its cross section is not large at LHC, this process is particularly interesting if the Higgs mass is relatively low and its decays have large backgrounds, or at $p\bar{p}$ colliders (Tevatron), due to the relative abundance of anti-quarks, or at lepton colliders (LEP) where the Z boson is produced by e^+e^- pairs.

Associated production with top or bottom quarks (ttH/bbH) : despite the smaller cross section, the ttH production process provides distinctive signatures such as a large amount of hadronic activity in the collision or possibly one or more leptons from semileptonic top decays. This production process is sizable at the LHC, given the large center-of-mass required to produce three massive particles. In the SM, the cross section of the associated production with b quarks is slightly higher than ttH, but the background is also larger by several orders of magnitude, making this process extremely difficult to be detected at a hadron collider.

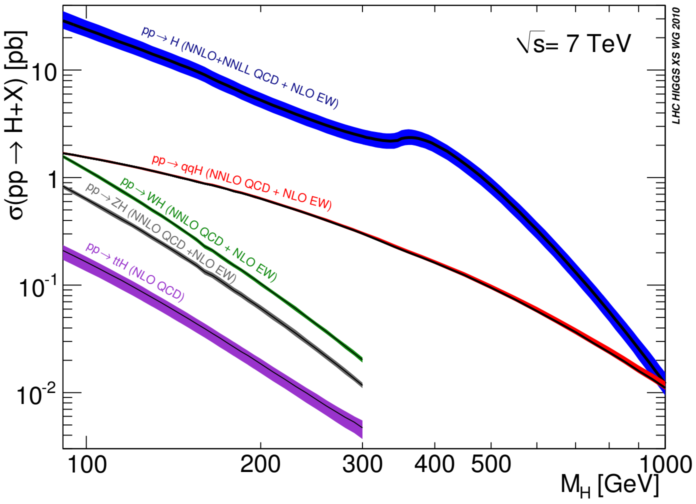

Figure 1.2: Standard model Higgs cross sections at LHC as a function of Higgs mass at $\sqrt{s} = 7\,\mathrm{TeV}$ [12].

1.3.2 SM HIGGS BOSON DECAYS

The Higgs branching ratios ($\mathcal{B}r$) directly depend on the Higgs partial widths, that is, the inverse decay rates of each kinematically allowed final state. In the case of the Higgs decaying into fermion pairs, the width at tree-level is:

$$\Gamma_{f\bar{f}} = \frac{N_c G_F \beta^3 m_{\mathrm{h}} m_f^2}{4\sqrt{2}\pi}, \qquad \beta = \sqrt{1 - \frac{4m_f^2}{m_{\mathrm{h}}^2}}\,.$$

The partial width to fermions is linear in m_{h}, modulo the cubic β dependence, which steepens the ascent with m_{h} near threshold. While the total Higgs width above fermion thresholds grows with Higgs mass (Fig. 1.3), Higgs total widths below W W pair threshold are quite narrow, of the order of tens of MeV. The only complicated partial width to fermions is that of top quarks, which have to be treated as virtual particles up to $2m_t$.

From the experimental point of view, decays to light fermions are dominant if the Higgs boson is light ($m_{\mathrm{h}} \lesssim 135\,\mathrm{GeV}$). The most common decay modes are $\mathrm{h} \to \mathrm{b}\bar{\mathrm{b}}$ and $\mathrm{h} \to \tau\tau$, and the ratio of the two is proportional to the square of the fermion masses, but QCD corrections and the colour factor for b quarks have to be taken into account. Both the decay channels share a relatively poor experimental mass resolution, about $10 \sim 15\%$ for the $\mathrm{b}\bar{\mathrm{b}}$ and $\sim 20\%$ for the $\tau\tau$, the first limited by the difficulty of estimating the jet energy, the latter because of the presence of neutrinos in the final state. The $\mathrm{b}\bar{\mathrm{b}}$ advantage in terms of branching ratios is compromised by a much larger background at hadron colliders, making this search virtually impossible without the presence of other well identified particles in the event.

For an heavier Higgs boson ($m_{\mathrm{h}} > 135\,\mathrm{GeV}$) the decays to weak bosons become dominant. Near the W and Z masses thresholds, the decay into vector bosons can be treated as on-shell final states, making the partial width calculation easier:

$$\Gamma_{VV} = \frac{G_F m_{\mathrm{h}}^3}{16\sqrt{2}\pi}\delta_V \beta \left(1 - x + \frac{3}{4}x^2\right)$$

The factor $\beta = \sqrt{1-x}$ comes from the phase space, and $x 4 m_V / m_\mathrm{h}$ from the matrix element. The constant factor δ_V is 2 for $\mathrm{h} \to \mathrm{WW}$ and 1 for $\mathrm{h} \to \mathrm{ZZ}$ decays. The partial width is dominantly cubic in m_h^3, leading to a very rapid total width growth with m_h, which reaches 1 GeV around $m_\mathrm{h} = 190\,\mathrm{GeV}$. As a consequence, the Higgs couplings to vector bosons are stronger compared to fermions in the high-mass region ($m_\mathrm{h} \gtrsim 160\,\mathrm{GeV}$). Even if the top quark has a larger mass than the W or Z, the on-shell partial width and thus $\mathcal{B}r$ of the Higgs decay into $\mathrm{t\bar{t}}$ are smaller than those of vector bosons. The partial widths to vector bosons are non-negligible below threshold, in the $135 \lesssim m_\mathrm{h} \lesssim 160\,\mathrm{GeV}$ mass range: the W and Z are unstable particles and therefore have finite widths, and they may be produced off-shell, making them suitable also for Higgs searches at low mass. The possibilities of final states due to W and Z decays is large, but only few are experimentally interesting. The Higgs decay into two Z bosons, and their consequent decay into a pair of electrons and muons, has several experimental advantages, in spite of the small $\mathcal{B}r$: the presence of four leptons entails an almost background-free environment at hadron colliders, while the Higgs mass can be fully reconstructed with an excellent resolution due to the good performance of the lepton measurements. The other relevant channel is the $\mathrm{h} \to \mathrm{WW}^* \to \ell\nu\ell\nu$, which has an higher branching ratio, relatively small backgrounds but the presence of neutrinos in the final state does not allow a precise determination of the Higgs mass.

The Higgs decays can also involve massless gauge bosons, such as pairs of gluons (gg), photons ($\gamma\gamma$) or a Z boson and a photon ($\mathrm{Z}\gamma$). These decays are possible due to loop-induced transitions that can occur at higher orders in perturbation theory. Due to the strength of their couplings, loops of top quark and electroweak gauge bosons are most important. For $\mathrm{h} \to gg$, that means only top quarks contribute to the loop, since the W,Z do not couple with gluons, while for the $\mathrm{h} \to \gamma\gamma$ and $\mathrm{h} \to \mathrm{Z}\gamma$ both the top quark and W loops (interfering destructively) contribute to the decay width. Despite the small branching ratio, the decay into a pair of photons is one of the most interesting channel to detect a light Higgs boson in spite the large, smoothly falling QCD diphoton production, due to the very good mass resolution of photon pairs. The $\mathcal{B}r$ and the total width of the Higgs boson as a function of its mass is shown in Fig. 1.3.

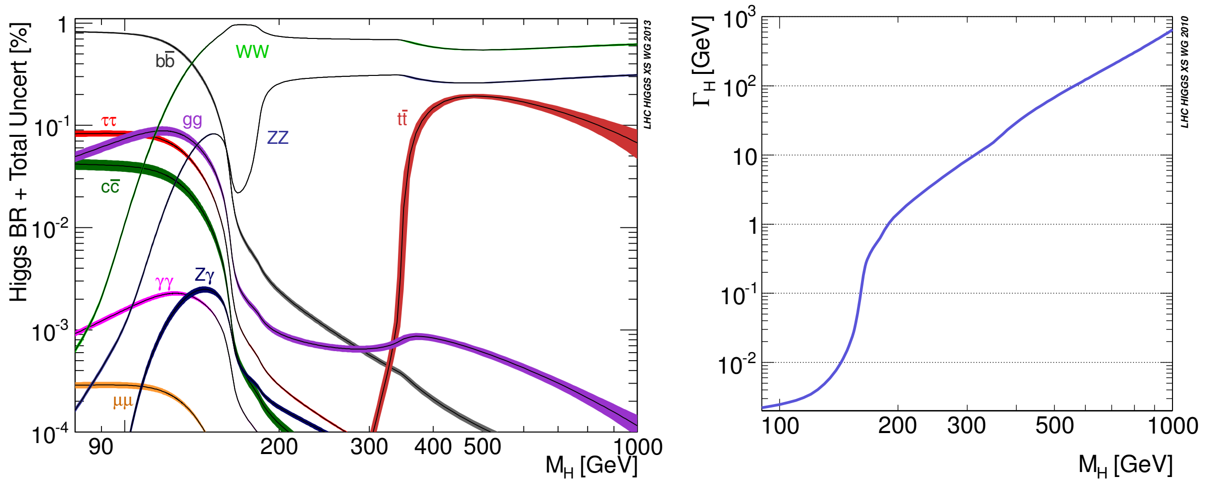

Figure 1.3: Standard model Higgs boson branching ratios (left) and total width (right) as a functions of its mass [12].

1.4 EARLY SEARCHES FOR THE SM HIGGS BOSON

The Standard Model describes remarkably well a large number of processes. Until 2012 the Higgs boson was the most important piece of the SM missing experimental evidence, and its mass was unknown. Early results on the Higgs boson mass began to be published in the 1970s, considering the possible effect of Higgs exchange on neutron- and deuteron-electron scattering [13], Higgs emission from neutron stars [14], nuclear transitions [15, 16] and neutron-nucleus scattering [17], setting for the first time experimental limits to the Higgs boson mass, even if quite far from the region of interest: $m_H > 15\,\mathrm{MeV}$ was the strongest limit obtained in that period. Those searches were motivated by the possibility that the Higgs mass was entirely due to quantum corrections, which would have yielded $m_H \sim 10\,\mathrm{GeV}$ in the absence of heavy fermions, being the top quark discovered decades later.

Searches for the Higgs boson at the electroweak scale began at the LEP e^+e^- collider starting from 1989 [18]. There were three important processes for producing the Higgs boson: in Z decay ($\mathrm{Z} \rightarrow \mathrm{h} + \bar{f}f$, with $\sqrt{s} \approx m_\mathrm{Z}$ at LEP I), in association with the Z ($e^+e^- \rightarrow \mathrm{Z} + \mathrm{h}$, with $\sqrt{s} \gtrsim m_\mathrm{Z}$ at LEP II), and via $\mathrm{W}^+\mathrm{W}^-$ or ZZ fusion ($e^+e^- \rightarrow \bar{\nu}\mathrm{h}\nu, e^+\mathrm{h}e^-$); the direct process $e^+e^- \rightarrow \mathrm{h}$ is negligible because of the small Yukawa couplings to e^+e^- at $\sqrt{s} = 91\,\mathrm{GeV}$. The dominant process, the Higgsstrahlung off a Z boson, has a cross section in the range of hundreds of femtobarns, sharply decreasing as the sum of Z and Higgs boson masses exceed the total center-of-mass energy. The other processes are much harder to detect because of the even smaller cross section. Because of this, the non-appearance of the Higgs boson in LEP I searches in Z decays implied $m_H > 58\,\mathrm{GeV}$ [19]. In 2000 the center-of-mass energy increased to 206 GeV, and the four LEP II experiments could investigate the range of Higgs masses below approximately 115 GeV.

A small excess of events observed by the ALEPH experiment [20] in the very highest-energy runs was not confirmed by the other experiments. The statistical significance of the signal was 1.7 standard deviations (Fig. 1.4), a value too small to justify an extended high-energy run. LEP II could exclude the existence of a SM Higgs boson with mass below 114.4 GeV, at 95% confidence level [19]. Because of the strong dependence of the cross section curve for Higgs boson on the LEP II center-of-mass energy, such a lower limit is very strict, and the probability of a failed discovery is essentially negligible for $m_\mathrm{h} \lesssim 112\,\mathrm{GeV}$.

The LEP experiments found no signatures of the Higgs boson [19]. However, even if direct searches proved to be unsuccessful, the electroweak constraints from experimental data obtained at LEP [21, 22] together with the SLC results and the top quark discovering [23] at Tevatron made it possible to set a $m_\mathrm{h} < 186\,\mathrm{GeV}$ upper limit at 95% CL to the mass of the SM Higgs boson [24].

1.5 SEARCHES FOR THE SM HIGGS BOSON AT TEVATRON

During roughly the same years, the Tevatron, a proton-antiproton ($p\bar{p}$) machine was also gathering data at Fermilab in the USA, and the lead in Higgs searches was taken by the CDF and D0 experiments. The production of the Higgs boson at hadron colliders is more problematic than in e^+e^- collisions. The backgrounds from other physical processes are large, and the direct production is small, because quarks which constitute the protons and antiprotons are light, and consequently their couplings to the Higgs are also small. The dominant production mechanism at the Tevatron for the Higgs boson is via gluon-gluon fusion, but the associated production with a $\mathrm{W}^\pm$ or Z offers a cleaner signature due to the leptonic decays of the vector bosons. These allow an efficient triggering of the events, making easier to distinguish the decays of the Higgs

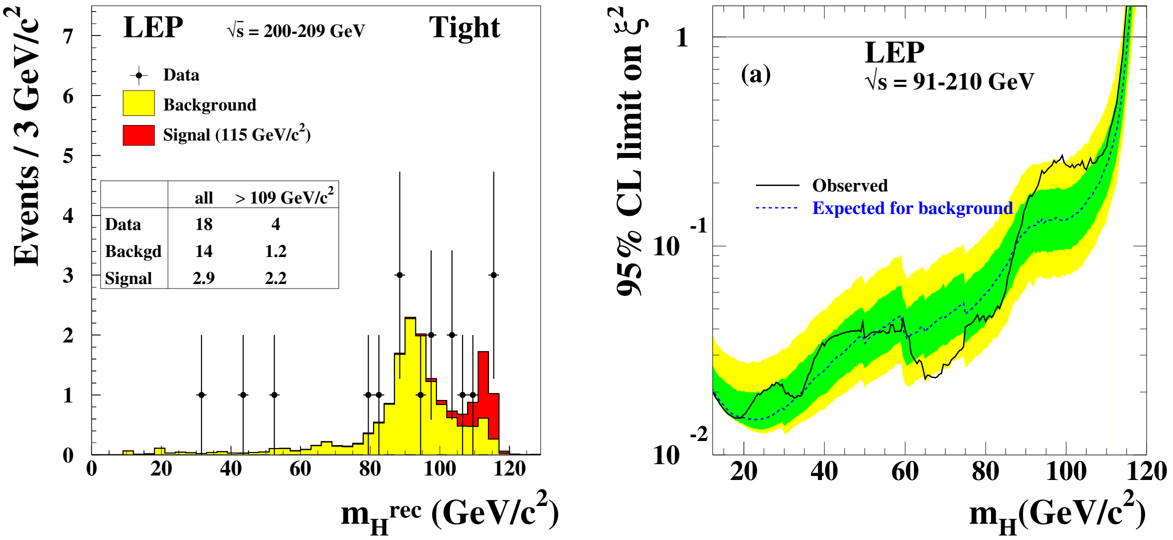

Figure 1.4: Left: data and simulation distributions of the reconstructed Higgs boson mass m_H^{rec} with the full LEP II luminosity. Right: observed and expected 95% CL upper limits on the Higgs boson production ratio $\xi = \sigma/\sigma_{SM}$ from the combination of the four LEP II experiments. The green and yellow shaded areas around the median expected line correspond to the 68% and 95% probability bands, and the horizontal line represent the Standard Model coupling [19].

boson into pairs of b jets from the Higgs decay from the much larger multijet background. Due to the small cross section, varying from 400 fb to 100 fb for m_h in the 100 $\sim$ 140 GeV range, the Tevatron experiments started to be sensitive to the signal after collecting an integrated luminosity of several inverse femtobarns. The most recent combined limits on the rate of Higgs boson production by CDF and DZERO are still unable to tighten the low mass exclusion bound set by LEP II in 2002. Interestingly, there is a small excess of Higgs candidate events in a range around 120 to 140 GeV compatible with the mass (Fig. 1.5) and the cross section expected for the SM Higgs boson (Fig. 1.6).

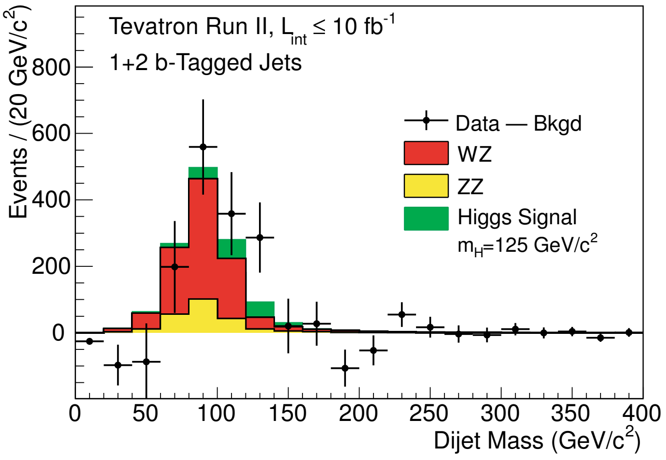

Figure 1.5: Dijet invariant mass in the $V\mathrm{h} \to \mathrm{b\bar{b}}$, showing data with expected background subtracted (excluding diboson production) for the CDF and D0 collaborations [25].

For higher Higgs boson masses, a sizable fraction of the decays are into pairs of W bosons. In the dileptonic channel, CDF and DZERO excluded an intermediate range of Higgs boson mass between $156 < m_\mathrm{h} < 177\,\mathrm{GeV}$ [25], due to the distinctive signature of two leptons and missing energy in the final state, which is easy to trigger and less demanding to separate from backgrounds. However, because of the presence of neutrinos, the mass of the Higgs boson cannot be reconstructed as precisely as in the $\mathrm{h} \to \mathrm{b\bar{b}}$ decay. Other channels, like the $\mathrm{h} \to \gamma\gamma$ and $\mathrm{h} \to \tau\tau$ are considered, but their sensitivity is generally lower.

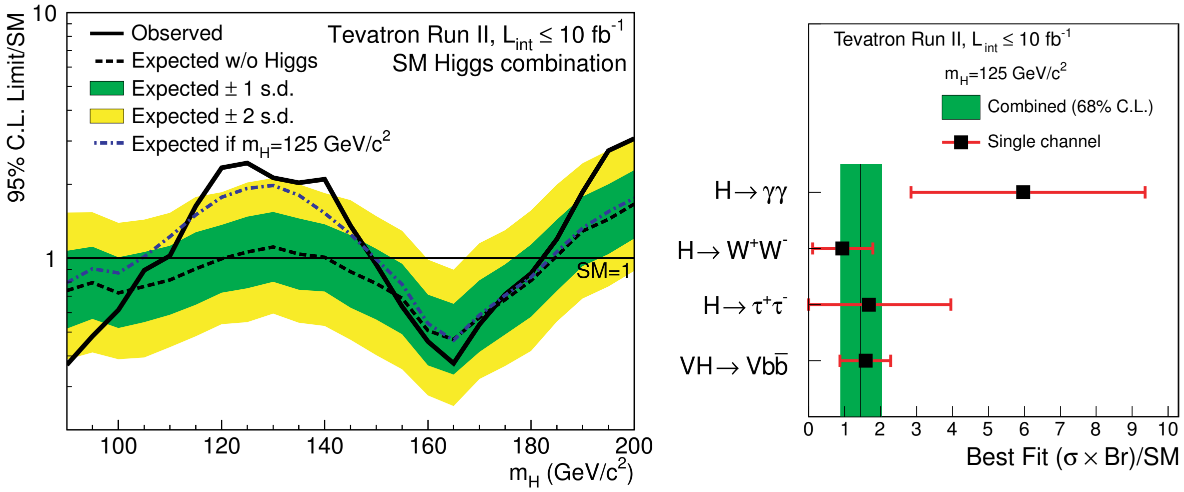

Figure 1.6: Left: expected and observed exclusion limit for a SM Higgs boson as a function of its mass. This results is the combination of the $V\mathrm{h} \to \mathrm{b\bar{b}}$, $\mathrm{h} \to \mathrm{WW}$, $\mathrm{h} \to \tau\tau$, $\mathrm{h} \to \gamma\gamma$ for the CDF and D0 collaborations, with the full Tevatron luminosity [25]. Right: best-fitted cross section divided by the SM Higgs boson cross section for the four search channels [25].

1.6 THE DISCOVERY OF THE HIGGS BOSON AT THE LHC

Searches for the Higgs boson finally came to an end in the Summer 2012, with the simultaneous announcements from the CMS and ATLAS experiments of the observation of a new boson with a mass close to 125 GeV with a statistical significance of 5σ, after collecting approximately $5\,\mathrm{fb}^{-1}$ of 2011 data at $\sqrt{s} = 7\,\mathrm{TeV}$ and $5\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 8\,\mathrm{TeV}$ in 2012. The trustworthiness of this claim was supported by event excesses observed in data in several different channels sensitive to the decay of a scalar boson, in the high-resolution (4ℓ, $\gamma\gamma$) and low-resolution channels ($2\ell2\nu$, $\tau\tau$, $\mathrm{b\bar{b}}$) final states. The first two have been particularly useful for a very early yet precise determination of the Higgs boson mass, and later of its properties, such as spin and parity. The observed CMS significance was 4.9σ [26, 27], and analogously the ATLAS experiment reported 6.0σ [28]. These claims induced CERN to publicly announce the discovery of a new particle on July 4, 2012.

Immediately after the discovery, the searches became measurements. In this section a summary of the main SM Higgs channels at CMS are described, updated to the full 2011 and 2012 integrated luminosity (about $5\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 7\,\mathrm{TeV}$ and $20\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 8\,\mathrm{TeV}$).

1.6.1 HIGGS BOSON DECAY IN THE FOUR LEPTON FINAL STATE

The $\mathrm{h} \to \mathrm{ZZ}^* \to 4\ell$ channel, with the Z decay into pairs of electrons and muons, has always been considered the “golden” channel in the Higgs boson searches at LHC [29]. The presence of four isolated leptons suppresses the large QCD and electroweak background processes at an hadron collider, except for the irreducible Z Z production, where the two bosons are non resonating. The excellent invariant mass resolution ($1-2\%$ depending on the channel) of the four-body system also allowed a further separation from the background and a precise determination of the resonance mass. The main drawback of the 4ℓ channel was the very small branching fraction of this decay, which required a rather large amount of data to observe a statistically significant signal. The experimental challenge was to detect four leptons with small transverse momentum, maximizing acceptance and reconstruction efficiency.

The CMS analysis is inclusive, aiming to detect the $\mathrm{h} \rightarrow \mathrm{ZZ}^* \rightarrow 4\ell$ decay regardless of the Higgs boson production mode. In order to be sensitive also for a Higgs boson with a mass below the kinematical threshold, where the $\mathcal{B}r(\mathrm{h} \rightarrow \mathrm{ZZ})$ is still sizable (see Sec. 1.3), one or both the Z boson candidates are allowed to be *off-shell*. The sensitivity of the search is increased by building a multivariate likelihood discriminator using the masses and the production and decay angles of the two Z bosons. The results for the combined 2011 and 2012 datasets are extracted through a two-dimensional fit on the four-lepton invariant mass ($m_{4\ell}$, shown in Fig. 1.7), and the discriminator output. The observed excess around 125 GeV has a statistical significance of 6.5 standard deviations.

1.6.2 HIGGS BOSON DECAYS IN THE DIPHOTON CHANNEL

The $\mathrm{h} \rightarrow \gamma\gamma$ decay is an interesting channel for probing a low mass Higgs boson [30]. Even though the decay is quite rare, having a branching ratio of about $\sim 0.1\%$, the two high-energy photons in the final state provide a clear signature against the background processes. The excellent photon reconstruction and energy measurement capabilities at CMS allow a precise determination of the Higgs boson mass. The analysis strategy is to look for narrow peaks over a smoothly-falling continuum of the diphoton invariant mass spectrum $m_{\gamma\gamma}$, which is by far the most discriminant variable against QCD and fake photons backgrounds. The analysis sensitivity is enhanced by dividing events in exclusive categories, using detector regions and multivariate discriminators to define classes with different signal purity and mass resolution. Additional categories are defined as a function of the number of jets in the event, to exploit also the VBF production mode (see Sec. 1.3). The diphoton invariant mass is shown in Fig. 1.7, by weighting each category by the observed S/(S+B) ratio. The results are extracted with a combined fit on the background distribution, modeled and normalized a-posteriori on the $m_{\gamma\gamma}$ spectrum. The excess observed for $m_{\gamma\gamma} \approx 125$ GeV, has a local significance of 4.0σ.

1.6.3 HIGGS BOSON DECAYS IN THE $2\ell 2\nu$ CHANNEL

The $\mathrm{h} \rightarrow \mathrm{WW}$ channel is the most sensitive for a Higgs boson in the mass range near the W and Z boson kinematical thresholds ($160-190$ GeV) where decays to pairs of on-shell W W are possible, and the Z Z channel is not completely open [31]. The final states with two leptonic W decays are the most important for the Higgs search, due to the small electroweak backgrounds, coming from mainly $\mathrm{t\bar{t}}$ dileptonic decays and the irreducible W W dibosons production. Due to the presence of two undetected neutrinos in the final state, the invariant mass of the system cannot be fully reconstructed, making this channel suitable for a discovery but not for a precise mass determination. Despite this flaw, a good discriminating power can be achieved exploiting the different process kinematics. Events are categorized according to the flavour of the two leptons (ee, $\mu\mu$ and separately $e\mu$, due to the different background composition) and as a function of the number of reconstructed jets. Selection cuts are applied to discriminate against reducible backgrounds regardless of the Higgs boson mass hypothesis, and finally the signal is extracted using the output distributions of multivariate classifiers trained individually for each Higgs boson mass hypothesis. Data show a broad excess in a mass region around 120 GeV, as expected, with a significance of about 4.7σ for a 125 GeV Higgs boson hypothesis.

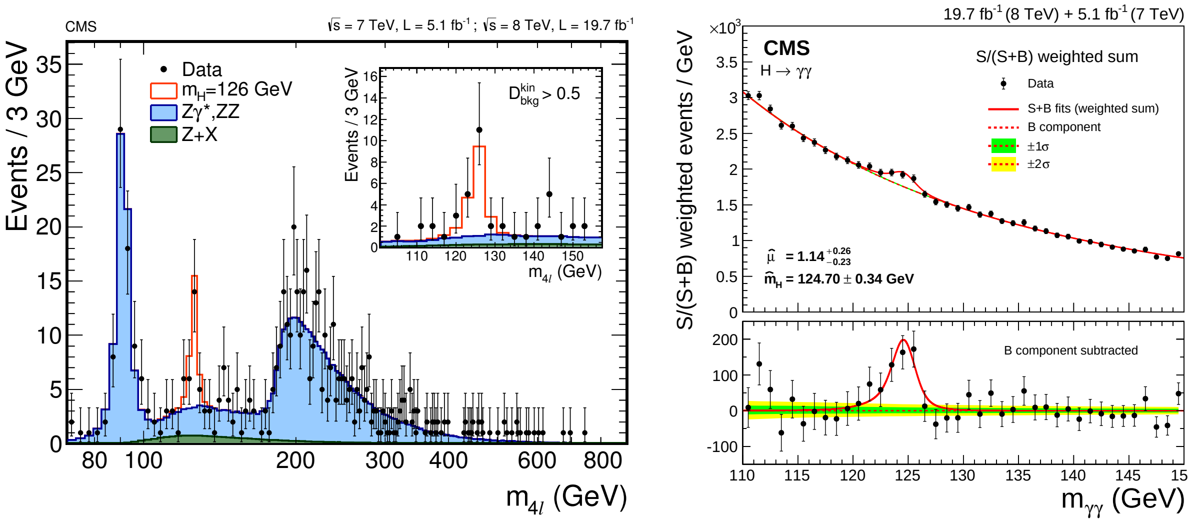

Figure 1.7: Left: four-lepton reconstructed mass in full mass range for the sum of the $4e$, 4μ, and $2e2\mu$ channels. The insert shows the low mass region, with a cut on the likelihood discriminant $\mathcal{D} > 0.5$ to reject the ZZ background. Right: diphoton invariant mass distribution with each event weighted by the S/(S+B) value of its category. The inset shows the central part of the unweighted invariant mass distribution.

1.6.4 HIGGS BOSON DECAYS INTO τ PAIRS

The observation of the Higgs boson decays into fermion pairs is extremely important, because it represents a direct test the Yukawa couplings of the Higgs boson to quarks and leptons. Higgs boson decays into fermions are suitable only for a low mass Higgs boson, as the branching fraction steeply decreases for $m_{\mathrm{h}} \gtrsim 135\,\mathrm{GeV}$ when the decays into vector bosons becomes sizable. The natural candidates for this search are the $\mathrm{h} \to \mathrm{b\bar{b}}$ and $\mathrm{h} \to \tau\tau$ channels, that have the largest branching ratio $\mathcal{B}r \approx 60\%$ and $\approx 5-8\%$ respectively. Other decay modes, such as $\mathrm{h} \to \mu\mu$, could be observed only with a much more larger data samples. Due to the presence in the final state of two hadronic jets for the $\mathrm{h} \to \mathrm{b\bar{b}}$ channel, or two or more neutrinos for the $\mathrm{h} \to \tau\tau$ channel, the Higgs boson invariant mass resolution is worse than other channels, being about 10% for the $\mathrm{b\bar{b}}$ and 20% for the $\tau\tau$ final states.

The $\mathrm{h} \to \tau\tau$ channel can also play an important role in the analyses for a light SM Higgs boson, thanks to the excellent CMS τ reconstruction and trigger performances [32]. Four ditau decay final states are used: $e\tau_h$, $\mu\tau_h$, $e\mu$ and $\mu\mu$, where τ_h denotes a tau decaying hadronically. Events are separated into different categories also according to the final state topology: a VBF category, including events with two leading jets at large rapidity separation; a boosted category, in which at least one hadronic jet is required; and a zero-jet category, characterized by larger event yields but lower purity due to the larger $\mathrm{Z} \to \tau\tau$ background. In the categories of events with zero or one jet, where a signal contribution is expected mostly from the gluon fusion production mode, events are further separated according to the p_{T} of the hadronic τ or the muon, depending on the final state. The invariant mass $m_{\tau\tau}$ is reconstructed from the leptons and the missing energy in the event using a likelihood technique based on a model of the τ-decay phase-space and the missing energy resolution, which results into a typical 20% mass resolution. Different signal extraction strategies are adopted for each category, due to the varying background composition. The combined observed excess for a 125 GeV Higgs boson is 3.8σ.

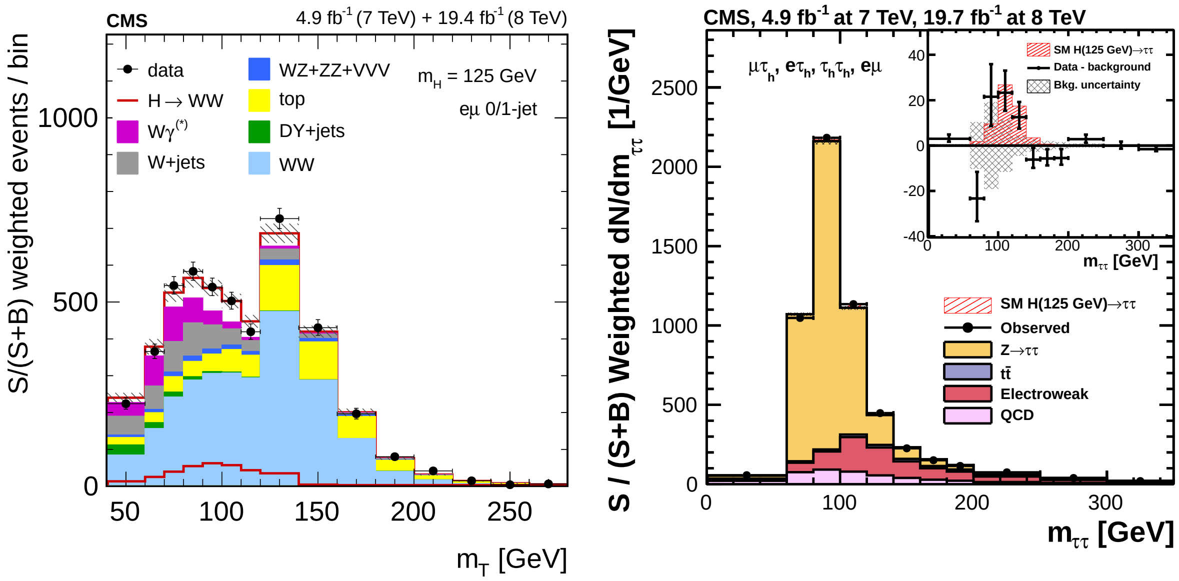

Figure 1.8: Left: transverse mass m_T distribution in the h $\rightarrow WW \rightarrow e\nu\mu\nu$ final state for the 0-jet and 1-jet categories combined. Right: combined observed and predicted $m_{\tau\tau}$ distributions for the $\mu\tau_h$, $e\tau_h$, $\tau_h\tau_h$, and $e\mu$ channels. The inset shows the corresponding difference between the observed data and expected background distributions, together with the signal distribution for a SM Higgs boson at $m_\mathrm{h} = 125\,\mathrm{GeV}$. In both plots, events are weighted according to the observed S/(S+B) ratio in each category.

1.6.5 HIGGS BOSON DECAYS INTO A PAIR OF B QUARKS

The search for a Higgs boson decaying into a pair on b quarks [33] is performed relying on the associated production with a vector boson (Vh, with V=Z,W) due to the overwhelming QCD multijet background in the gluon fusion and vector boson fusion production modes. Five vector boson decay modes are considered: W $\rightarrow e\nu$, W $\rightarrow \mu\nu$, Z $\rightarrow ee$, Z $\rightarrow \mu\mu$, and Z $\rightarrow \nu\nu$. In the latter, no high-momentum leptons are present, but it is detected as a large unbalance in the transverse momentum of the event due to the neutrinos. The main backgrounds in this channel are the production of vector bosons in association with b quark jets (called Zbb̄, Wbb̄) and light-flavoured jets mistagged as b jets. The top quark pair and the diboson production are also important backgrounds, given the larger cross sections and the topology close to the signal. These backgrounds are reduced by performing the search in the phase space region where the vector boson and the dijet system are produced with large transverse momenta ($p_\mathrm{T}^V > 50\,\mathrm{GeV}$) in a back-to-back topology. This kinematic regime is also characterized by a more accurate jet energy reconstruction. In each channel, event categories are defined for different boost intervals of the vector boson. To improve the sensitivity, two multivariate techniques are used. The first consists of a energy regression on the momentum of the two b jets, improving by $10-20\%$ the dijet invariant mass resolution. The second is a BDT discriminator trained on simulated signal and background events for a number of Higgs boson masses, since the discriminating power offered by the reconstructed mass peak is not enough to obtain a good signal sensitivity without any prior knowledge of the background. An analysis of the output distribution of the discriminator is performed, as shown in Fig. 1.9, observing a 2.0σ significance for a 125 GeV signal hypothesis.

The h $\rightarrow$ bb̄ decay should be viable also in the tt̄ associated production channel, probing the coupling of the Higgs boson to the top quark in a model-independent manner. Although the tt̄ production has a smaller predicted cross section (0.6% of the total produced Higgs bosons), it leaves a distinctive signature of at least one charged lepton (electron or muon) in the event. The main background is the QCD production of a $t\bar{t}$ pair and additional light- and heavy-flavour jets, which has a larger cross section than the signal and it is affected by sizable uncertainties. The high multiplicity of b jets in the event (two from the top quark decays, and other two from the Higgs boson decays) makes it difficult to reconstruct the Higgs candidate invariant mass. Searches with multivariate techniques, like the Matrix Element Method, are more suitable for the $t\bar{t}h$ and $h \to b\bar{b}$. The CMS analysis in this channel observes a modest excess, but compatible with the SM Higgs with a measured best–fit value of $\mu = 1.2^{+1.6}_{-1.5}$.

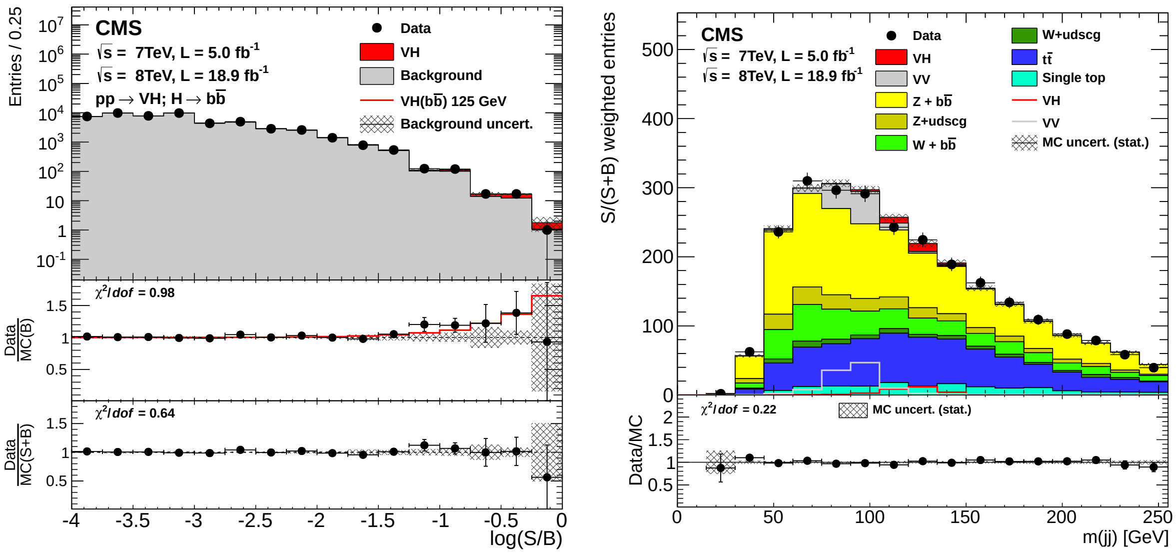

Figure 1.9: Left: combination of all $Vh \to b\bar{b}$ channels ($W \to e\nu$, $W \to \mu\nu$, $Z \to ee$, $Z \to \mu\mu$, and $Z \to \nu\nu$) into a single distribution. Events are sorted in bins of similar expected signal-to-background ratio, as given by the value of the output of their corresponding BDT discriminant. The two bottom insets show the ratio of the data to the background-only prediction (above) and to the predicted sum of background and SM Higgs boson signal with a mass of 125 GeV (below). Right: dijet invariant mass, combined for all channels and weighted by S/(S+B) in each channel and boost region.

1.6.6 COMBINATION AND PROPERTIES OF THE SM HIGGS BOSON

The results of the CMS Higgs analyses in the $h \to ZZ$, $\gamma\gamma$, WW, $b\bar{b}$, $\tau\tau$ final states, performed with the Run I data collected in 2011 at $\sqrt{s} = 7$ TeV and in 2012 at $\sqrt{s} = 8$ TeV, are combined together to measure the properties of the newly-discovered boson, and test the compatibility with the SM Higgs [34].

The Higgs boson mass is directly measured by reconstructing the invariant mass of its decay products in the $h \to ZZ \to 4\ell$ and $h \to \gamma\gamma$ channels, due to the excellent invariant mass resolution and the absence of neutrinos in the final state. The results of a fit to the data, without relying on the SM Higgs prediction for the expected signal yield, is reported in Fig. 1.10, together with the combination. The CMS measured mass of the Higgs boson is $m_h = 125.03^{+0.26}_{-0.27}(\text{stat})^{+0.13}_{-0.15}(\text{syst})$ GeV.

The compatibility of the new resonance with the standard model Higgs boson is quantified in terms of the best fit of the signal strength modifier $\mu = \sigma/\sigma_{SM}$. The 125 GeV observed excess is fully compatible with the SM Higgs hypothesis in all the considered channels, as shown in

Fig. 1.10, and the combined global best fit value is found to be $\mu = 1.00 \pm 0.09$ (stat)$^{+0.08}_{-0.07}$ (theo) $\pm$ 0.07 (syst). A slight discrepancy is observed in the ttH associated production and decays in final states with multileptons, where an excess of events (2.1σ from the SM Higgs expectation) is observed. The SM Higgs compatibility for different mass hypotheses is essentially zero.

Channel	$\mathcal{B}r$	Mass resolution	Expected	Observed	$\mu = \sigma/\sigma_{SM}$
$\mathrm{h} \to \mathrm{ZZ} \to 4\ell$	$1.25 \cdot 10^{-4}$	$1 - 2\%$	6.3σ	6.5σ	$1.00^{+0.32}_{-0.26}$
$\mathrm{h} \to \gamma\gamma$	$2.28 \cdot 10^{-3}$	$1 - 2\%$	5.1σ	5.6σ	$1.13^{+0.26}_{-0.23}$
$\mathrm{h} \to \mathrm{WW} \to 2\ell 2\nu$	$1.06 \cdot 10^{-2}$	20%	5.7σ	5.4σ	$0.83^{+0.22}_{-0.20}$
$\mathrm{h} \to \mathrm{b}\bar{\mathrm{b}}$	57.7%	10%	2.2σ	2.0σ	$0.91^{+0.29}_{-0.26}$
$\mathrm{h} \to \tau\tau$	6.32%	20%	4.1σ	3.8σ	$0.93^{+0.51}_{-0.48}$
Combination					1.00 ± 0.13

Table 1.2: Summary of the CMS SM Higgs boson measurements. For each considered channel, the total branching ratio and the mass resolution is reported, together with the expected and observed significance. The last column is the best fit value $\mu = \sigma/\sigma_{SM}$ [34].

The total width of the Higgs boson, Γ_h, can be constrained using its relative on-shell and off-shell production and decay rates to a pair of Z bosons with final states with leptons and neutrinos [35]. In the dominant gluon fusion production mode the off-shell production cross section is known to be sizable (about 8% of the total), due to an enhancement in the decay amplitude from the vicinity of the Z boson pair production threshold. The observed relative contributions of off-shell and on-shell events leads to an upper limit $\Gamma_\mathrm{h} < 22\,\mathrm{MeV}$ at a 95% CL, which is 5.4 times the expected value in the SM.

The study of the spin-parity of the Higgs boson [36] is performed using the $\mathrm{h} \to \mathrm{ZZ} \to 4\ell$, $\mathrm{h} \to \mathrm{WW} \to 2\ell 2\nu$, and $\mathrm{h} \to \gamma\gamma$ decay modes, testing the observed distribution of sensitive production and decay angular variables with the various spin (0, 1, 2) and parity (scalar or pseudoscalar) signal hypothesis. While the first two channels lead to tighter constraints on the spin-parity, the latter allows tighter constraints in the spin-two case. The spin-one hypotheses are excluded at a greater than 99.999% CL, and the spin-two boson are excluded at a 99% CL.

All observations are consistent with the expectations for a scalar SM-like Higgs boson. Under the assumption that C-parity or CP-parity are conserved, our measurements require the quantum numbers of the new state to be $J^{PC} = 0^{++}$.

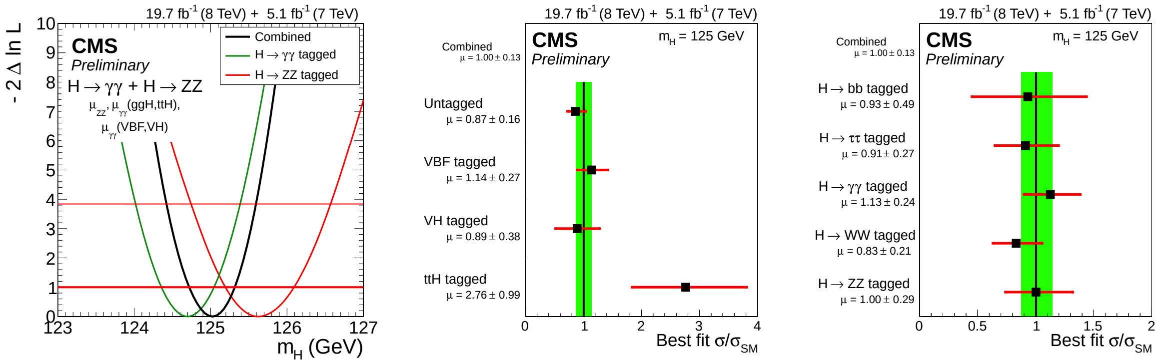

Figure 1.10: Test statistics scan vs hypothesized Higgs boson mass for the $\gamma\gamma$ and 4ℓ final states separately and for their combination (left); values of the best-fit $\mu = \sigma/\sigma_{SM}$ for the combination and separately by production mechanism (center) and decay mode (right) [34].

Chapter 2

Beyond standard model extensions of the Higgs sector

Despite the remarkable success of the SM, a consistent description of the fundamental laws of Nature is far to be completed. There are strong theoretical and experimental motivations suggesting that SM is not the ultimate theory, since it does not integrate gravitational forces, and does not foresee neutrino masses, does not provide a dark matter candidate and does not explain the hierarchy problem. Several theories have been proposed to overcome these issues, and many posit the presence of additional particles, including more than one Higgs bosons. In this chapter, the motivations to extend the standard model are summarized, together with the description of the predictions of the most accredited theories extending the SM, with the focus on the phenomenology of the Higgs sector and on the searches at leptonic and hadronic colliders.

2.1 ISSUES OF THE STANDARD MODEL

2.1.1 DARK MATTER AND DARK ENERGY

One of the most striking evidences of new physics beyond the standard model (BSM) comes from astrophysics observations, suggesting that the known types of matter in the universe, the baryonic matter and the electromagnetic radiation, amount only to about 5% of the entire universe's content. According to the prevailing cosmological models, the remaining 95% of the universe is made of two different constituents. The first is called *Dark Matter* (DM), and accounts for about 26% of the universe mass. The fraction is derived from studies of the anisotropy of the Cosmic Microwave Background radiation [37]. Other evidences are also provided by gravitational lensing observations, simulations of large-scale structures formation of the universe, and by observation of galaxy rotation curves and dynamics of galaxy clusters. The remaining 69% of the mass is believed to be constituted by *Dark Energy*, an hypothetical form of energy contributing to the acceleration of the universe's expansion, which was discovered studying the supernovae. The dominant theoretical explanation for Dark Matter is the Weakly Interacting Massive Particle (WIMP), a massive and cold particle that interacts mainly through gravitational forces, and to a lesser extent with weak, electromagnetic, and strong interactions. The estimated fraction of DM in the universe is surprisingly coherent with the hypothesis of a stable particle with a mass at the electroweak scale ($m_\chi \sim 100\,\text{GeV}$) and cross sections $\sigma_\chi \sim 10\,\text{fb}$ at the electroweak scale, copiously produced in the early phases of the universe [38].

2.1.2 NEUTRINO MASSES

The SM postulates the presence of massless, left-handed neutrinos. However, experimental measurements of neutrino flavour oscillations [39, 40] implies that the mass, albeit very small, is not null. The SM procedure of the mass terms introduction via gauge-invariant Yukawa interactions with the Higgs field should in principle work also for neutrinos, but the very small neutrino masses as compared to the lepton and quark masses require an explanation.

The right-handed counterpart of the observed left-handed neutrinos, required by the Yukawa term $M\bar{\nu}_R\nu_R^\dagger$, would be a singlet of all gauge groups, and hence called the "sterile" neutrino, and should not compromise the renormalizability and the gauge invariance of the theory. The mass M would be the only dimensional parameter that enter the SM Lagrangian apart from the Higgs mass, and it can be related to some new physics scale such as M_{GUT}. The diagonalization of the neutrino mass matrix leads to very small mass eigenstates, $m_\nu \sim M_{EW}/M$. Setting $M \sim M_{GUT} \sim 10^{15}\,\text{GeV}$ and $M_{EW} \sim 100\,\text{GeV}$, the expected neutrino masses should be at $m_\nu \simeq 0.01$ eV, which roughly corresponds to the actual experimental constraints on the neutrino masses. This is the simplest implementation of the so-called *See-Saw* mechanism [41], used for the explanation of the neutrino masses.

2.1.3 NATURALNESS SCALE OF THE SM

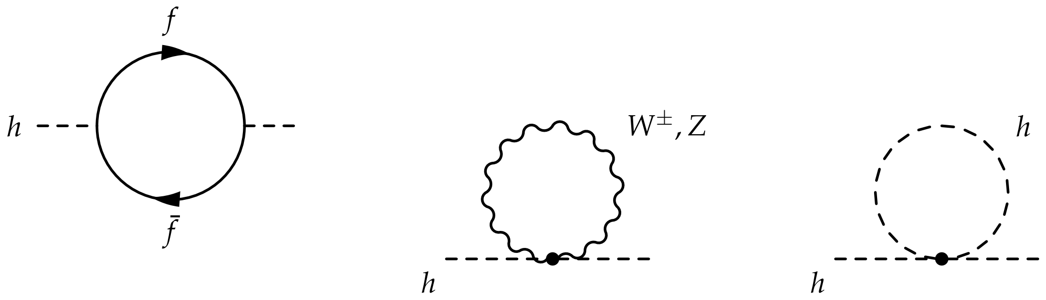

Figure 2.1: 1-loop radiative corrections to the Higgs boson mass.

The radiative contributions to the Higgs mass of standard model particles via 1-loop diagrams (Fig. 2.1) are quadratically divergent. Assuming that SM is a reliable theory up to the energy scale Λ, and assuming that the loop diagrams are valid up to the same scale, the radiative corrections to the Higgs mass are:

$$\delta m_{\mathrm{h}}^2 \approx \frac{3\Lambda^2}{8\pi^2 v^2}(2m_{\mathrm{W}}^2 + m_{\mathrm{Z}}^2 + m_{\mathrm{h}}^2 - 4m_t^2) \approx -\frac{\Lambda^2}{25}$$

The physical measurable Higgs boson mass m_{h} requires a cancellation between the bare Higgs mass $m_{h,bare}^2$ and the correction δm_{h}^2, the latter being much larger than m_{h}^2, depending on Λ. If the SM is correct up to $\Lambda \approx 5\,\text{TeV}$, the difference between the two is about two orders of magnitude. Regardless on the Λ scale, the quadratic dependence of the divergence requires a unnatural fine-tuning of the SM parameters at higher energy scales. This problem, known as *the hierarchy problem*, is one of the reasons to expect BSM new physics at the TeV scale.

2.1.4 VACUUM ENERGY AND STABILITY

To account for the Dark Energy contribution, an uniform (or very slowly varying) energy density is introduced in cosmological models by means of the Cosmological constant Λ. Associating the cosmological vacuum energy from Λ measurements from the Cosmic Microwave Background anisotropy [37] with the QFT vacuum [42], the latter is many orders of magnitude larger than the first. The large discrepancy between the two of about $\sim 10^{120}$ makes this problem the worst fine-tuning problem in modern physics.

Another cosmological problem, associated with the ground state of the Higgs potential is the fact that a $m_{\mathrm{h}} \approx 125\,\mathrm{GeV}$ Higgs boson violates the so-called stability bound [43]. This originates from calculations of the renormalization group running of the Higgs self-coupling constant λ, that becomes negative at higher scales; the effect is that the Higgs potential becomes unstable at that scale, making the theory inconsistent. The current measurements strongly favor a meta-stability region, as can be seen in Fig. 2.2.

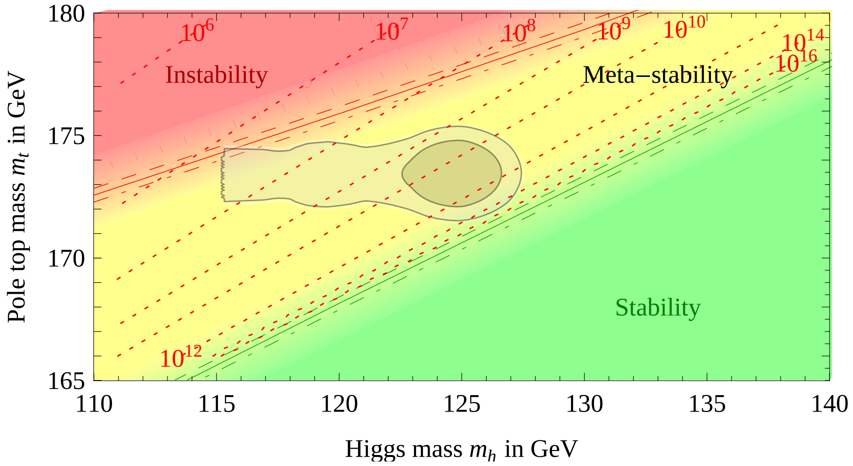

Figure 2.2: Stability, instability and meta-stability regions in the m_{h} vs. m_t plane. Contour lines denote the current SM measurements, while diagonal lines show the scale at which the theory becomes inconsistent. The three boundaries lines corresponds to the constraints on the α_S, and the grading of the colours indicates the size of the theoretical errors [43].

2.2 THE TWO HIGGS DOUBLET MODEL

The Two-Higgs Doublet Model (2HDM) is the simplest extension of the SM Higgs sector with two complex doublets of scalar fields [44]. The 2HDM is particularly attractive because it predicts new phenomena [45], such additional physical Higgs bosons, and is also a minimal extension, adding few new arbitrary parameters. It satisfies the theoretical constraints $\rho \simeq 1$ without any adjustment of the model parameters [46] and the absence of FCNC at tree level, if appropriate Higgs-fermion couplings are chosen.

Two doublets Φ_1 and Φ_2 of $SU(2)_L$ are introduced, made by complex scalar fields with hypercharge 1:

$$\Phi_1 = \begin{pmatrix} \phi_1^+ \\ \phi_1^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \qquad \Phi_2 = \begin{pmatrix} \phi_2^+ \\ \phi_2^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_5 + i\phi_6 \\ \phi_7 + i\phi_8 \end{pmatrix}$$

The most general Higgs potential $V(\Phi_1, \Phi_2)$ for the two-doublet model is:

$$
\begin{aligned}
V = \; & \tfrac{\lambda_1}{2}(\Phi_1^\dagger\Phi_1)^2 + \tfrac{\lambda_2}{2}(\Phi_2^\dagger\Phi_2)^2 + \lambda_3(\Phi_1^\dagger\Phi_1)(\Phi_2^\dagger\Phi_2) + \lambda_4(\Phi_1^\dagger\Phi_2)(\Phi_2^\dagger\Phi_1) \\
& + \left[\tfrac{\lambda_5}{2}(\Phi_1^\dagger\Phi_2)^2 + \lambda_6(\Phi_1^\dagger\Phi_1)(\Phi_1^\dagger\Phi_2) + \lambda_7(\Phi_2^\dagger\Phi_2)(\Phi_1^\dagger\Phi_2) + h.c.\right] \\
& -\tfrac{1}{2}\left[m_{11}^2(\Phi_1^\dagger\Phi_1) + m_{22}^2(\Phi_2^\dagger\Phi_2) + m_{12}^2(\Phi_1^\dagger\Phi_2) + m_{12}^{2*}(\Phi_2^\dagger\Phi_1)\right]
\end{aligned}
$$

The free parameters λ_1, λ_2, λ_3, λ_4, m_{11}, m_{22} are real, and λ_5, λ_6, λ_7, m_{12} are generally complex. The general 2HDM has thus a total of 14 free parameters. To achieve FCNC suppression the potential must be invariant under the following discrete symmetries:

$$\Phi_1 \to -\Phi_1,\ \Phi_2 \to \Phi_2 \qquad \text{or} \qquad \Phi_1 \to \Phi_1,\ \Phi_2 \to -\Phi_2$$

After the symmetry breaking, the scalar fields $\Phi_{1,2}$ assume their corresponding vacuum expectation values (v.e.v.) $\langle\Phi_1\rangle$ and $\langle\Phi_2\rangle$, which correspond to the value at the global minimum of the Higgs potential. If the previous conditions are verified, the global isospin symmetry can be used to write the vacuum expectation values as:

$$\langle\Phi_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v_1 \end{pmatrix},\ \langle\Phi_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v_2 e^{-i\xi} \end{pmatrix}$$

with v_1 and v_2 real positive numbers, and $e^{-i\xi}$ is the complex phase between the expectation values of the two doublets, and describes the CP-violation in the 2HDM Higgs sector.

As a result of the Higgs mechanism in the 2HDM, three degrees of freedom are absorbed by the longitudinal polarization of the $W^\pm$ and Z bosons when they become massive. The remaining five degrees of freedom lead to the existence of five physical Higgs bosons:

- two charged bosons, $H^\pm$, with quantum numbers $J^P = 0^+$
- two neutral bosons, h and H, with quantum numbers $J^{CP} = 0^{++}$
- one neutral pseudoscalar bosons A, with quantum numbers $J^{CP} = 0^{+-}$

The masses of the bosons depend on $\lambda_{1..7}$. Instead of only one λ parameter in the SM, the parameters of the 2HDM $\lambda_{1..7}$ can be expressed by:

$$m_{\mathrm{H}^\pm}, \qquad m_{\mathrm{H}}, \qquad m_{\mathrm{h}}, \qquad m_{\mathrm{A}}, \qquad \tan\beta, \qquad \alpha$$

assuming no CP-violation with $\xi = 0$. In the 2HDM, the ratio of the vacuum expectation values of the two doublets is defined as the parameter:

$$\tan\beta = \frac{v_2}{v_1}\ .$$

The mixing between the two Higgs doublets is not zero, and it is expressed by the angle α.

With five Higgs bosons there is a large number of possible couplings of the bosons to the particles of the model. A choice of certain couplings has experimental consequences. On the base of the C and P quantum numbers of the bosons, the Higgs-vector fields couplings can be established. Table 2.1 summarizes the couplings that are either proportional to $\cos(\alpha-\beta)$, $\sin(\alpha-\beta)$ or angle-independent.

$\cos\alpha-\beta$	$\sin\alpha-\beta$	Angle-independent
$\mathrm{HW^+W^-}$	$\mathrm{hW^+W^-}$	
HZZ	hZZ	
ZAh	ZAH	$\mathrm{ZH^+H^-}$, $\gamma\mathrm{H^+H^-}$
$\mathrm{W^\pm H^\mp h}$	$\mathrm{W^\pm H^\mp H}$	$\mathrm{W^\pm H^\mp A}$
$\mathrm{ZW^\pm H^\mp h}$	$\mathrm{ZW^\pm H^\mp H}$	$\mathrm{ZW^\pm H^\mp A}$
$\gamma\mathrm{W^\pm H^\mp h}$	$\gamma\mathrm{W^\pm H^\mp H}$	$\gamma\mathrm{W^\pm H^\mp A}$
		$VV\phi\phi$, $VV\mathrm{AA}$, $VV\mathrm{H^+H^-}$

Table 2.1: Couplings between Higgs bosons and vector fields. Here $\phi = \mathrm{h}, \mathrm{H}$ and $VV = \mathrm{W^+W^-}, \mathrm{ZZ}, \mathrm{Z}\gamma, \gamma\gamma$.

The Higgs bosons h and H couplings to vector fields (collectively denoted as $V = \mathrm{W}^\pm, Z$) in the 2HDM are different than the SM case:

$$\frac{c_v(VV\mathrm{h})}{c_v(VV\mathrm{h}_{SM})} = \sin(\beta-\alpha) \qquad \frac{c_v(VV\mathrm{H})}{c_v(VV\mathrm{h}_{SM})} = \cos(\beta-\alpha)$$

In the SM, CP-violation is possible thanks to the existence of a the complex phase of the CKM mixing matrix elements, but this source is not sufficient to explain the asymmetry between matter and antimatter in the universe [47]. The neutrino mixing PMNS matrix could be an additional source of CP-violation, analogously to the CKM matrix in the quark sector [48]. Beyond the standard model theories like the 2HDM give new potential sources of CP asymmetry and make possible the baryogenesis at the scale of the weak interaction. In the Higgs sector of the SM, the ρ parameter, defined as $\rho = m_\mathrm{W}^2/(m_\mathrm{Z}^2\cos^2\theta_W)$ is very close to 1 [49, 50], and this requirement has to be satisfied by extensions of the SM. A theoretical constraint on the Higgs sector comes from experimental limits on the Flavour-Changing Neutral Currents (FCNC). In the minimal Higgs sector FCNC at tree-level are not present, and this requirement has to be satisfied also in non-minimal Higgs models. Requiring all fermions of a given electric charge to couple only to one single Higgs doublet is the solution of the FCNC problem [51]. Other conditions to be satisfied are explained in Ref. [52].

Assuming no CP-violation in the extended Higgs sector, four different 2HDM declinations arise from the possible couplings of the Higgs bosons to fermions:

Type-I : all fermions couple only to the Higgs fields in the Φ_2 doublet, and Φ_1 is decoupled from fermions

Type-II : *up*-type quarks couple to the Φ_2 doublet, *down*-type quarks and charged leptons couples to the Higgs field in the Φ_1 doublet

Type-III : *up*-type quarks and charged leptons couple to the Φ_2 doublet, *down*-type quarks couple to Φ_1

Type-IV : all quarks couple to the Higgs field in the Φ_2 doublet, and charged leptons couple to Φ_1

There are no Higgs bosons couplings to neutrinos; as in the SM, neutrinos remain massless. Couplings with charged fermions receive enhancements of suppression depending on the model parameters and the type of 2HDM model. Table 2.2 reports these factors with respect the SM case for the Type-II 2HDM.

	$u\bar{u}, c\bar{c}, t\bar{t}$	$d\bar{d}, s\bar{s}, b\bar{b}, e^+e^-, \mu^+\mu^-, \tau^+\tau^-$
$c_f(\mathrm{h}f\bar{f})$	$\cos\alpha/\sin\beta$	$-\sin\alpha/\cos\beta$
$c_f(\mathrm{H}f\bar{f})$	$\sin\alpha/\sin\beta$	$\cos\alpha/\cos\beta$
$c_f(\mathrm{A}f\bar{f})$	$-i\gamma_5/\tan\beta$	$-i\gamma_5\tan\beta$

Table 2.2: Suppression or enhancements in the couplings between Higgs bosons and charged fermions with respect to the SM case.

2.3 THE HIGGS SECTOR OF THE MINIMAL SUPERSYMMETRIC STANDARD MODEL

Supersymmetric (SUSY) theories [53] are among the most favorable extensions of the SM, and are based on the introduction of an extra symmetry between fermions and bosons, implying that every boson must have a corresponding supersymmetric fermion counterpart and vice-versa. The SM does not respect this symmetry, since there are no fermionic and bosonic fields with the same quantum numbers except for the spin. This implies that SUSY must be a symmetry broken at the Λ_{SUSY} energy scale, below which only some particles can be observed. The boson-fermion symmetry is attractive even at the Λ_{SUSY} scale because it solves the hierarchy problem: the contribution of every boson loop would be canceled by the corresponding fermion loop, and vice-versa. Certain models of SUSY also provide an ideal candidate for the Dark Matter, predicting the presence of stable, massive particles that do not interact with ordinary matter. SUSY theories have several declinations, given the high number of assumptions that can be made. Among the most immediate extensions of the SM involving SUSY, there is the Minimal Supersymmetric Standard Model (MSSM). In the MSSM Higgs sector, the supersymmetric partner of the Higgs boson, called *higgsino*, contributes to fermionic loops that must be canceled to avoid inconsistencies in the theory. For this reason, in the MSSM two oppositely-charged higgsinos are posited, and consequently two Higgs doublets with opposite hypercharges are introduced:

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

Under these assumptions, the MSSM Higgs sector is a special case of the general Two Higgs Doublet Model. It predicts, after symmetry breaking, five Higgs bosons of which three are neutral, denoted as h, H, A, and two charged, H^+ and H^-. The symmetry considerations require the Yukawa interactions between the Higgs doublets and the fermions to have the 2HDM Type-II structure. However, the λ_i parameters of the Higgs potential are not free in the MSSM, but are fixed by gauge interaction couplings. The MSSM Higgs potential is the following:

$$V = \frac{g^2 + g'^2}{8}(\Phi_1^\dagger\Phi_1 - \Phi_2^\dagger\Phi_2)^2 + \frac{g^2}{2}(\Phi_1^\dagger\Phi_2)^2 + m_1^2\Phi_1^\dagger\Phi_1 + m_2^2\Phi_2^\dagger\Phi_2 - 2m_3^2\Re(\Phi_1\Phi_2)$$

where the parameters m_i^2 $(i = 1, 2, 3)$ are also fixed at Λ_{SUSY}, but their values are such that the potential has only a trivial minimum. The model still does not allow CP-violation in the Higgs sector.

The general MSSM contains 126 free parameters, making the study of the full parameter space unfeasible. Experimental results from collider experiments, cosmology and symmetry requirements can substantially reduce the number of effective parameters, but the remaining ones

($\sim$ 10) are still too many to be tested independently. A number of *benchmark* points, highlighting various phenomenological possibilities, are selected by assigning fixed values to the free parameters of the theory. A common choice is to fix the masses of the squarks M_{SUSY}, the gluino mass M_1, the gaugino mass parameter M_2, the higgsino mass parameter μ, the stop and sbottom trilinear couplings A_t and A_b, the stop mixing parameter X_t. This last one is important because of the large top-Yukawa corrections that the MSSM Higgs sector receives. The values $X_t = 0$ and $X_t = \sqrt{6}M_s$ are referred to as “no mixing” and “maximal mixing”, because they extremize the loop corrections. The h-H mixing angle α and the masses of the CP-even Higgs bosons are determined at tree-level as:

$$\alpha = \frac{1}{2}\tan^{-1}\left(\tan 2\beta \frac{m_{\rm A}^2 + m_{\rm Z}^2}{m_{\rm A}^2 - m_{\rm Z}^2}\right), \qquad -\frac{\pi}{2} \leq \alpha \leq 0$$

$$m_{\rm h,H}^2 = \frac{1}{2}\left(m_{\rm A}^2 + m_{\rm Z}^2 \pm \sqrt{(m_{\rm A}^{+} m_{\rm Z}^2)^2 - 4m_{\rm Z}^2 m_{\rm A}^2 \cos^2(2\beta)}\right)$$

The charged Higgs mass is rather more simple:

$$m_{\rm H^\pm}^2 = m_{\rm A}^2 + m_{\rm W}^2$$

The only remaining free parameters are the mass of the CP-odd Higgs $m_{\rm A}$, and the ratio of the two Higgs doublets vacuum expectation values $\tan\beta$. Benchmark scenarios can be defined for specific purposes. At tree-level, the light Higgs h is predicted to have a mass $m_{\rm h} \lesssim m_{\rm Z}$, which has been excluded by the LEP and Tevatron experiments [19]. However, a specific benchmark scenario, called m_h^{max}, has been proposed to be used for LHC searches [54]. Maximizing the radiative corrections to the Higgs mass from the stop and sbottom, the mass bounds are then pushed up to $m_{\rm h} \lesssim 135\,\text{GeV}$. A complete set of benchmark scenarios [55] used at LHC before and after the discovery of the $m_{\rm h} = 125\,\text{GeV}$ Higgs boson, are summarized in Tab. 2.3.

Parameter	m_h^{max}	m_h^{mod+}	m_h^{mod-}	light stop	light stau	tau-phobic	low-$m_{\rm H}$
M_{SUSY}	1000	1000	1000	500	1000	1500	1500
X_t^{OS}/M_{SUSY}	2.0	1.5	-1.9	2.0	1.6	2.45	2.45
X_t^{MS}/M_{SUSY}	$\sqrt{6}$	1.6	-2.2	2.2	1.7	2.9	2.9
A_t				$= X_t + \mu\cot\beta$			
A_b	$= A_t$	$= A_t$	$= A_t$	$= A_t$	$= A_t$	$= A_t$	$= A_t$
A_τ	$= A_t$	$= A_t$	$= A_t$	$= A_t$	0	0	$= A_t$
μ	200	200	200	350	500	2000	varied
M_1				fixed by GUT relation to M_2			
M_2	200	200	200	350	200	200	200

Table 2.3: MSSM parameters in different benchmark scenarios [55]. Dimensional quantities are expressed in GeV. Other parameters are $m_{\tilde{g}} = 1500\,\text{GeV}$, $m_{\tilde{q}} = 1500\,\text{GeV}$, $m_{\tilde{l}} = 500\,\text{GeV}$, $A_{f\neq t,b,\tau} = 0$.

2.4 THE PHENOMENOLOGY OF THE EXTENDED HIGGS SECTOR

In Beyond Standard Model theories the phenomenology of the Higgs sector can be significantly different than the one of the SM. Additional free parameters are introduced, modifying the couplings of the Higgs bosons with the other particles and consequently the production cross sections and branching fractions. In general 2HDM and MSSM, cross sections and decays are strongly dependent on the model parameters.

A relevant case, where cross sections can be considerably enhanced with respect the SM case, is provided by Type-II 2HDM and MSSM models. At high $\tan\beta$ values, $\tan\beta \gtrsim 10$, either h or H is SM–like and its couplings to other particles are the same as those of the SM Higgs boson, while the other CP-even state behaves as the CP–odd A, having the same couplings and almost the same mass. The collective state of the two bosons is denoted as ϕ, leading to an approximate doubling of the production cross section.

The couplings of the neutral Higgs bosons to down-type quarks and charged leptons are enhanced by a factor of $\tan\beta$ and the corresponding coupling to the up-type quarks and leptons is suppressed. This means that decay modes to bottom quarks and tau leptons are of specific interest, with a predicted decay rate of around 90% to bottom quarks and 10% to tau leptons. The suppression of the coupling to the top quark implies that the gluon-gluon fusion production rate decreases, as well as the associated production with top pairs. On the other hand, the enhancement of the coupling to down-type fermions implies that the bottom contribution becomes non-negligible, and the associated production mode reaches and overtakes the gluon-gluon fusion cross section if $\tan\beta \gtrsim 5$. At even higher values of $\tan\beta$, the b quarks associated production cross section becomes so large that could be observed at hadron colliders such as the Tevatron or LHC. In these scenarios, the VBF and VH production of Higgs bosons and decays to a vector bosons pair are unimportant due to the relatively small cross section; additionally, in the case of the pseudoscalar A boson, these modes are forbidden due to CP-violation in the VVA vertex.

Higgs bosons decays are generally similar to the SM case for the lightest scalar. The $\phi \to \mathrm{b\bar{b}}$ and to a lesser extent $\phi \to \tau\tau$ decays are dominant at low mass, while $\phi \to t\bar{t}$ becomes particularly important if kinematically allowed ($m_\phi \gtrsim 2 \cdot m_t$). Couplings with vector bosons can be suppressed, and these have also effects on $\phi \to \gamma\gamma$ and $\phi \to \mathrm{Z}\gamma$ decay rates. Furthermore, $\mathrm{A} \to \mathrm{WW}, \mathrm{ZZ}$ decays are forbidden. However, several new possible decay modes become available for the heavy neutral Higgs bosons involving the light Higgs boson h. The heavy CP-even H can have a sizable branching fraction into pairs of h, while the pseudoscalar A can decay into a Z boson and a light Higgs. These decay modes depend on the mass of the heavy bosons, being relevant above the kinematic threshold and below twice the top mass, where the $\phi \to t\bar{t}$ dominates; they can also be suppressed in certain regions of the parameter space, such as with high $\tan\beta$ in MSSM.

The phenomenology of the charged Higgs bosons is simpler. The main $\mathrm{H}^\pm$ production mode at hadron colliders is through top quark decays $t \to b\mathrm{H}^\pm$ for charged Higgs boson masses smaller than the top quark mass. The cross section for production from single top quark events is much smaller, and often not considered. The decay mode largely depends on the mass, being $\mathrm{H}^\pm \to tb$ if kinematically accessible, and $\mathrm{H}^\pm \to \tau\nu$ if lighter. There is a rare decay $\mathrm{H}^\pm \to \mathrm{W}^\pm\mathrm{h}$, suppressed at high $\tan\beta$.

2.5 BSM HIGGS BOSON SEARCHES LEP

The LEP experimental program was focused not only on the SM scalar, but it included also searches for signatures of an extended Higgs sector. As in the SM case two production modes are possible, depending on the mass of the boson. If the Higgs is light enough ($m_\Phi \lesssim 115\,\mathrm{GeV}$), pair productions in the $e^+e^- \to \Phi_1\Phi_2$ are possible, otherwise the Higgsstrahlung $e^+e^- \to \mathrm{Z}\Phi_1$ is viable up to slightly higher masses. Since the phase space probed by LEP was limited to light scalar bosons, the Higgs bosons were expected to decay into pairs of b quarks ($\approx$ 90%) or τ leptons ($\approx$ 10%), and all the possible combinations of these two decay modes were taken into account. Also the presence of more than one light Higgs with different masses, $m_{\Phi 2} > m_{\Phi 1}$, can involve cascade decays such as $\Phi_2 \rightarrow \Phi_1\Phi_1$ or $\Phi_2 \rightarrow Z\Phi_1$ if kinematically allowed. The final states measured in the detector can lead to two of four b jets in the event, or either two b jets and two τ or four τ. If the Z boson is present, it can be detected through the missing energy from its decay into neutrinos.

The combination of the ALEPH, DELPHI, L3, OPAL experimental results of searches for an extended Higgs sector have been published as a 95% CL limit on the cross section [56]. Additional experimental constraints were applied from $\Delta\Gamma_Z$, decay-mode independent $Z\Phi_1$ searches, and searches for the Yukawa processes which contribute mostly for very light Φ_1 masses. The combined LEP data interpretations were given for different MSSM benchmark scenarios [55].

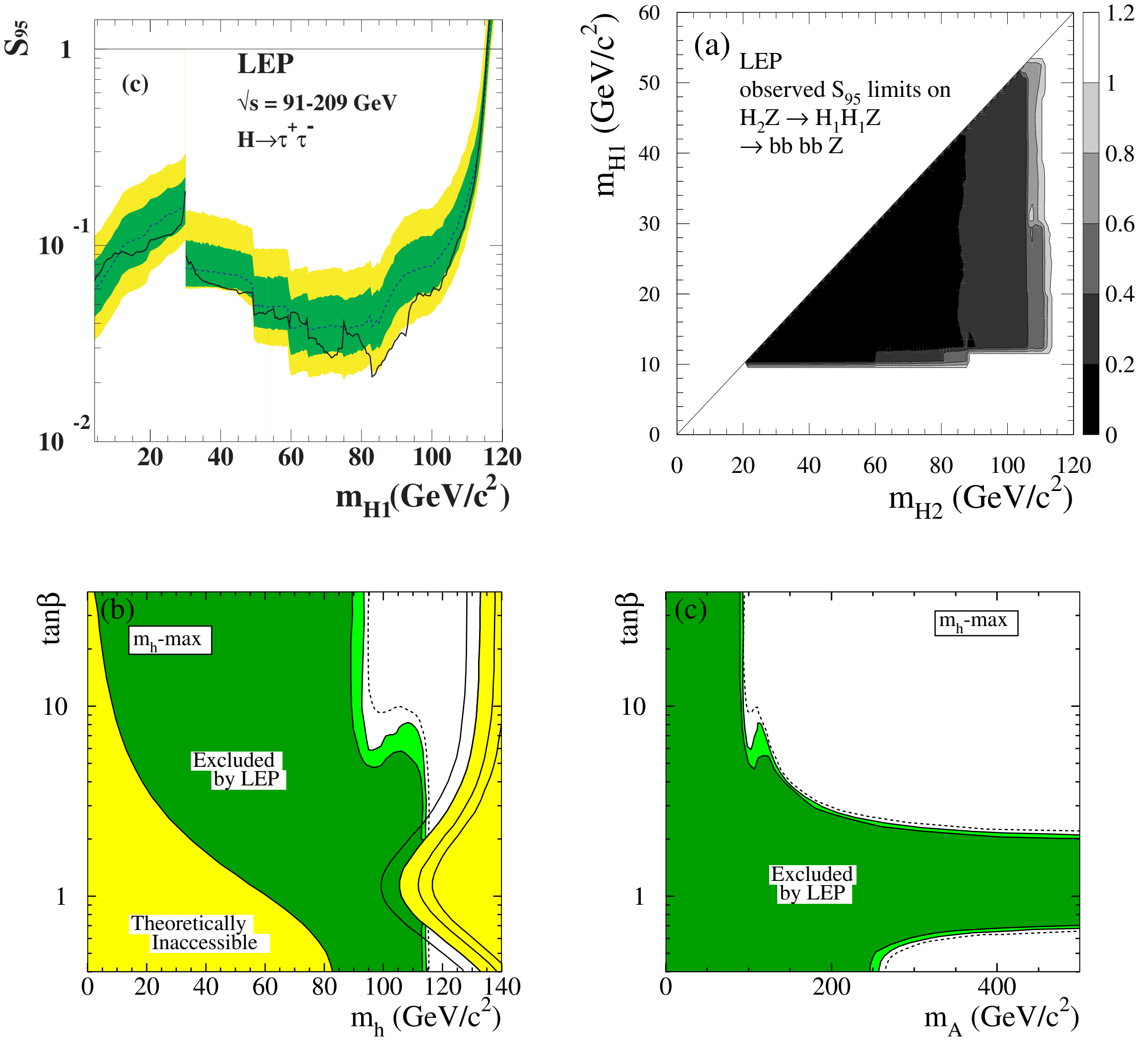

Figure 2.3: Top: the 95% CL upper bounds (S_{95}), for the Higgsstrahlung process $\Phi_1 \rightarrow \tau\tau$, as a function of the Higgs boson mass (left) and upper bounds for the Higgsstrahlung cascade process $e^+e^- \rightarrow Z\Phi_2$, with $\Phi_2 \rightarrow \Phi_1\Phi_1$ and $\Phi_1 \rightarrow b\bar{b}$ (right). Bottom: exclusions at 95% CL (medium-grey or light-green) and the 99.7% CL (dark-grey or dark-green), in the case of the CP-conserving m_h^{max} benchmark scenario. The figure shows the theoretically inaccessible domains (light-grey or yellow) and the regions excluded by this search in the (m_h, $\tan\beta$) and (m_A, $\tan\beta$) parameter space.

The exclusion limits are presented in different declinations of the MSSM parameter space, for both CP-violating and CP-conserving scenarios. In the latter case, the exclusions for the m_h^{max} benchmark scenario are shown in Fig 2.3. In the region with $\tan\beta$ less than about five, the exclusion is provided mainly by the Higgsstrahlung process, giving a lower bound of about 114 GeV for m_h. At high $\tan\beta$, the pair production process is most useful, providing limits in the vicinity of 93 GeV for both m_h and m_A. For m_h in the vicinity of 100 GeV, a deviation between the expected and the experimental exclusions is observed. This deviation, which is also present in other CP-conserving scenarios, is due to the excess in the Higgsstrahlung channel visible also in the SM search. The data also exclude certain domains of $\tan\beta$. This is illustrated in the (m_A, $\tan\beta$) projection (Fig. 2.3). In the CP-conserving m_h^{max} scenario, values of $\tan\beta$ between 0.7 and 2.0 are excluded, but this range depends considerably on the assumed top quark mass and M_{SUSY}.

2.6 BSM HIGGS BOSON SEARCHES AT TEVATRON

The current searches at the Tevatron were not sensitive to a SM-like Higgs in the mass region allowed by the LEP exclusion bounds [19, 25], except for a small mass range around twice the W mass. On the other hand, scenarios with enhanced Higgs boson production cross sections could be probed with the accumulated luminosity. Enhanced production cross sections can occur in particular for low m_A in combination with large $\tan\beta$ due to the enhanced couplings of the Higgs bosons to down-type fermions.

The two dominant production mechanisms of neutral MSSM Higgs bosons at at high $\tan\beta$ are gluon-gluon fusion and $\mathrm{b\bar{b}}$ fusion, and the leading decay modes are $\Phi \to \mathrm{b\bar{b}}$ ($\sim 90\%$) and $\Phi \to \tau\tau$ ($\sim 10\%$). The presence of spectator b quarks in the final state significantly decreases the Z boson background, making the $\Phi \to \tau\tau$ decay channel competitive against the $\Phi \to \mathrm{b\bar{b}}$ decay, which has a larger branching ratio but suffers from a large multijet background and is more sensitive to MSSM parameters. The LEP experiments have excluded $m_\mathrm{A} < 93$ GeV [19], and higher-mass A for small $\tan\beta$. Searches at the Tevatron collider were complementary, providing sensitivity in the large $\tan\beta$ region.

Searches for neutral Higgs bosons decaying into τ lepton pairs and produced in association with b quarks, $b\Phi \to b\tau\tau$, have been performed by the D0 Collaboration [57] with the full Tevatron luminosity in $\tau_\mu\tau_h$ final states, where τ_h denotes hadronically-decaying τ. The CDF search for the same process [58] used a sample corresponding to an integrated luminosity of 1.8 fb^{-1} of Run I and Run II data and studied three final states, adding the $\tau_e\tau_h$ and $\tau_e\tau_\mu$ channels. No significant excess in signal over background has been observed and thus limits on the production cross section for neutral Higgs boson times the branching fraction into τ leptons are given for neutral Higgs bosons in the range of 90 to 300 GeV, and interpreted in the MSSM benchmark scenario (Fig. 2.4).

A MSSM signal was sought also in the $b\Phi \to b\mathrm{b\bar{b}}$ process by the CDF [59] and D0 [60] experiments. The analyses were challenging due to the intrinsic difficulty of modeling the multijet background, but offered a good sensitivity due to the high branching ratio of the neutral Higgs bosons into pairs of b quarks. Requiring at least one additional b quark allowed to discriminate the copious multi-b QCD yield, exploiting the relatively high cross section of this kind of electroweak process at high $\tan\beta$.

The CDF analysis, limited to a data sample of 2.6 fb^{-1}, showed a modest and broad excess in the 150 GeV mass region. The significance of the excess is 2.8 standard deviations, corresponding to a background p-value in this region of 0.23%, which becomes 2.5% taking into account the

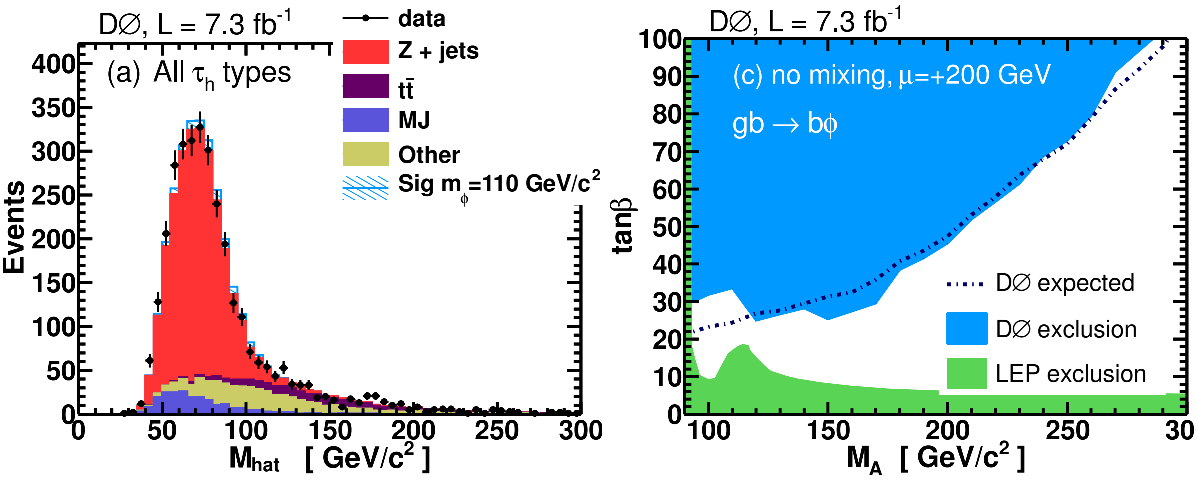

Figure 2.4: Left: Visible $\tau\tau$ invariant mass M_{vis} spectrum in the D0 search for the $\Phi \rightarrow \tau\tau$ process, combining the considered τ decay modes. Right: combined 95% CL upper bounds for the D0 search interpreted in the MSSM (m_A, $\tan\beta$) parameter space in the m_h^{max} benchmark scenario.

trials factor for the number of mass points searched. If this excess were to be attributed to the production of a narrow resonance in association with a b jet with kinematics characteristic of Higgs boson production, it would correspond to a production cross section times branching ratio of about 15 pb, as shown in Fig. 2.5. In the modified MSSM m_h^{max} scenario with negative $\mu = -200\,\text{GeV}$, the analysis allows exclusion of $\tan\beta$ values greater than 40 for $m_A = 90\,\text{GeV}$ and 90-140 for the mass range $110 - 170\,\text{GeV}$.

The analogous D0 analysis in the same final state used a larger dataset corresponding to an integrated luminosity of 5.2 fb^{-1} and a different method to estimate the background expectation. The data show a deviation from the expectation around 120 GeV (Fig. 2.5), corresponding to a 2.5σ excess. Limits on $\tan\beta$ as a function of m_A are derived for a m_h^{max} scenario with $\mu = -200\,\text{GeV}$; weaker limits are obtained for $\mu > 0$ scenarios, due to the decreased cross sections and branching ratios.

2.7 BSM HIGGS SEARCHES AT CMS

At LHC, the good τ reconstruction and identification performance, together with the relatively small SM backgrounds, makes the MSSM $\Phi \rightarrow \tau\tau$ channel the most sensitive to signatures of an extended Higgs sector. The CMS search [61] is very similar to the analysis targeting the same standard model final state, already described in Sec. 1.6.4, and the results presented here are based on the full 7 and 8 TeV datasets. The neutral MSSM Higgs bosons are expected to be predominantly produced either in gluon-gluon fusion or in b quark associated production. For larger values of $\tan\beta$ the latter mode dominates, and the branching fraction $\mathcal{B}r(\Phi \rightarrow \tau\tau)$ is also enhanced. An additional categorization, with either 0 or ≥ 1 b tagged jets, is defined to enhance the sensitivity to the two production modes. Model-independent upper limits at 95% CL are derived on the resonance production rate $\sigma(gg\Phi)B(\tau\tau)$ and $\sigma(bb\Phi)\mathcal{B}r(\tau\tau)$ independently (Fig. 2.6), and interpreted as constraints on the MSSM parameters in the m_A-$\tan\beta$ plane assuming the other MSSM parameters to be defined by the m_h^{max} benchmark scenario. In all cases, the signal hypothesis is tested against a background plus SM-Higgs hypothesis, taking into account the Higgs boson at $m_h = 125$ GeV.

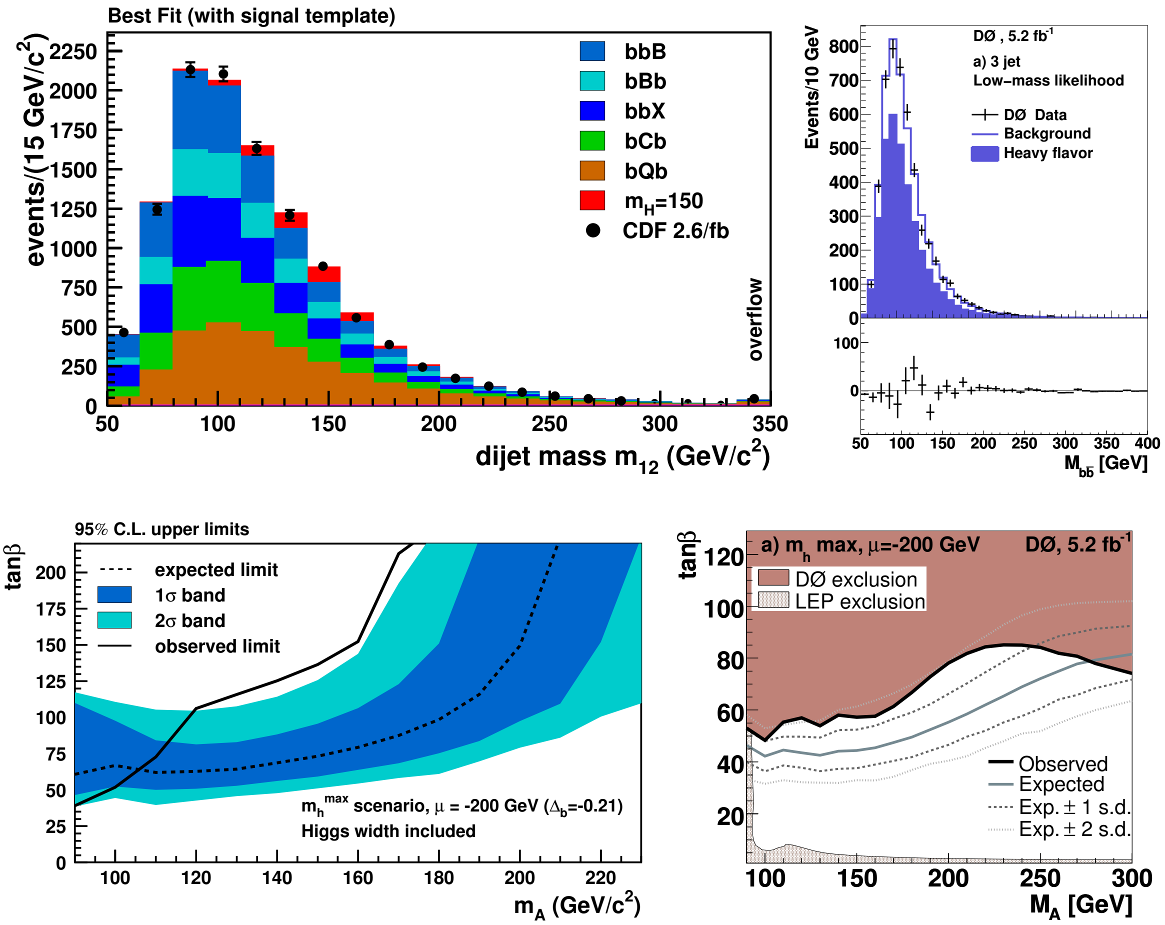

Figure 2.5: Top: dijet invariant mass spectrum in the CDF (left) and D0 (right) searches in the $b\Phi \to bb\bar{b}$ channel. Bottom: 95% CL limits interpretation in the MSSM parameter space and m_h^{max} benchmark scenario at CDF (left) and D0 (right).

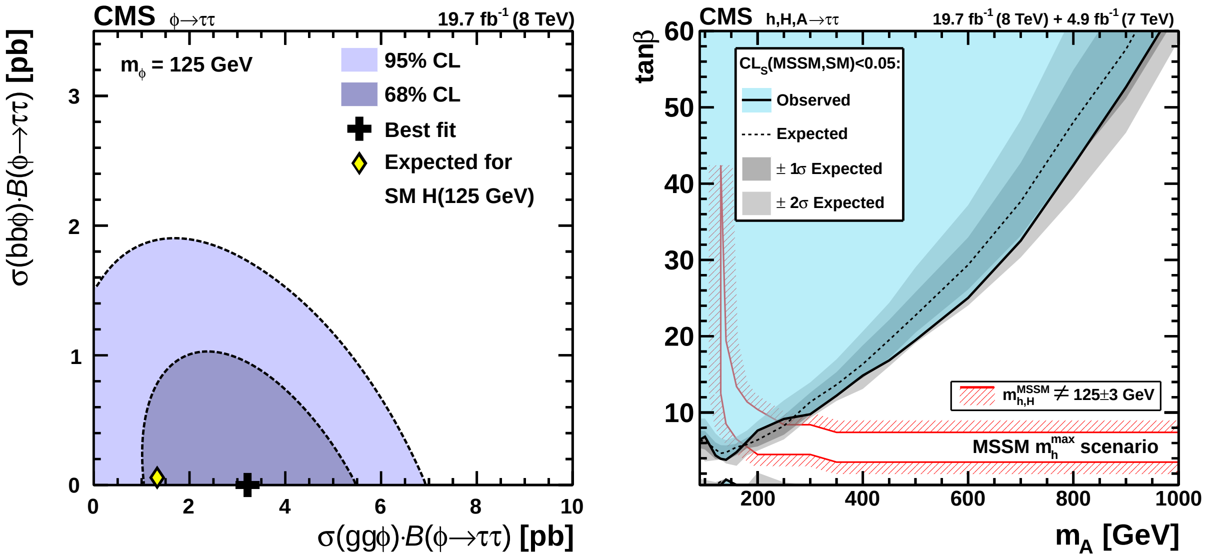

Figure 2.6: Left: likelihood contours of $\sigma(gg\Phi)B(\tau\tau)$ versus $\sigma(bb\Phi)B(\tau\tau)$, with the best fit value (cross) and the SM expectation for a 125 GeV Higgs (diamond). Right: limit in the MSSM m_A-$\tan\beta$ plane in the m_h^{max} scenario. The red band is the space allowed for $m_h = 125\,\text{GeV}$ [61].

Chapter 3

The LHC accelerator and the CMS experiment

The Run I collisions data provided by the LHC accelerator offer an unique possibility to investigate new physics phenomena at an energy scale never reached before. The excellent performance of the accelerator and the detector commissioning allowed physicists to perform relevant measurements and discover the Higgs boson within the very first years of operation. This chapter briefly describes the LHC accelarator and the CMS experimental apparatus and its subdetectors, explaining how particles produced in the collisions are detected, identified and reconstructed. This thesis consists in the analysis of 2011 and 2012 collision data collected in a center-of-mass energy of 7 and 8 TeV by the CMS experiment.

3.1 THE LARGE HADRON COLLIDER

The Large Hadron Collider (LHC) [62] is a proton-proton and heavy-ion collider operating at CERN since 2009 in the 27 km ring tunnel previously used by the LEP electron-positron collider [63]. The choice of a non-elementary particle as the proton for the beams instead of electrons-positrons is dictated by the necessity to increase the energy of the collisions, being LEP II limited by the energy loss due to synchrotron radiation.

The LHC injection chain is composed by several accelerators. The proton beam starts from LINAC, a small linear accelerator, where its energy firstly reaches 50 MeV. It then passes through a booster, then in the PS where it is accelerated up to 25 GeV, and then in the SPS up to 450 GeV. The beam is finally injected in the LHC ring, where it is accelerated up to 3.5 TeV in the 2010-2011 period and to 4 TeV in 2012. Starting from 2015 collisions with a center-of-mass energy of $\sqrt{s} = 13$ TeV close to the 14 TeV design value should take place. The LHC ring and the acceleration chain are sketched in Fig. 3.1.

One critical aspect of accelerating protons to such high energies is the magnetic field, obtained employing about 9300 superconducting magnets along the ring, operating at a temperature of 1.9 K to provide a maximum magnetic field of 8.3 T. The accelerating power is given by radiofrequency cavities, giving a 485 keV energy gain per revolution, after accounting for the 7 keV loss per turn due to synchrotron radiation.

The high luminosity of the LHC beams is obtained with a high frequency bunch crossing and a

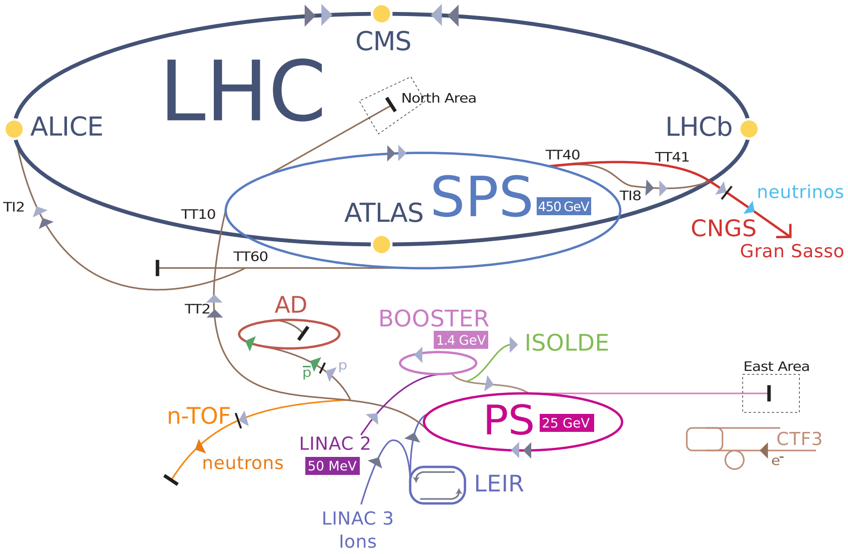

Figure 3.1: Schematic view of the CERN accelerator complex.

high density of protons per bunch. Each proton beam consists of 1380 bunches separated by at least 50 ns, and each bunch contains about 10^{11} protons, while the transverse dimension of the beam, fully squeezed at the collision points, is around 16μm. The usage of two proton beams instead of having one composed by anti-protons allow an easy and fast population of both beams with a high number of particles. Since the two partons involved in the interaction have unknown momentum, the total effective collision energy of the hard interaction is unknown, because the proton remnants are scattered at small angles and escape undetected in the beam pipe. As a consequence, it is not possible to precisely measure the total and missing energy of the event, but only their transverse components, i.e. their projections in the plane transverse to the beams.

The LHC beams collide in four different interaction points, where four experiments take place. The *Compact Muon Solenoid* (CMS) and the *A Toroidal LHC Apparatus* (ATLAS) are multipurpose detectors, designed around the experimental requirements to provide sensitivity to the Higgs boson, supersymmetry, extra dimensions, and heavy resonances. The *LHC beauty* experiment (LHCb) is focused on the physics of the quark b, while *A Large Ion Collider Experiment* (ALICE) is intended to study the quark-gluon plasma produced in heavy ion collisions.

3.1.1 LHC COMMISSIONING AND RUN I PERFORMANCE

In November 2009, LHC provided the first collisions at injection energy, 450 GeV per beam, increased soon afterwards to 1.18 TeV. In Spring 2010 the beam energy reached 3.5 TeV, allowing the experiments to collect up to 36 pb^{-1} of data by end of 2010. With the progressive increase in the instantaneous luminosity (Fig. 3.2), whose peak values reached $3.5 \cdot 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, 4.8 fb^{-1} were delivered by end of 2011. In 2012 the beam energy increased to 4 TeV per beam, with peak instantaneous luminosities up to $6.5 \cdot 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ allowing the experiments to collect about 5 fb^{-1} by the end of summer, and 20 fb^{-1} at the end of the year. In 2011 and 2012 the LHC operated with a 50 ns bunch spacing and about 1380 colliding bunches, and since the cross section for inelastic collisions is approximately 70 and 75 mb [64], respectively, the expected number of interactions in the same bunch crossing (pileup) for the 2011 and 2012 instantaneous luminosities is about 15 and 21 respectively.

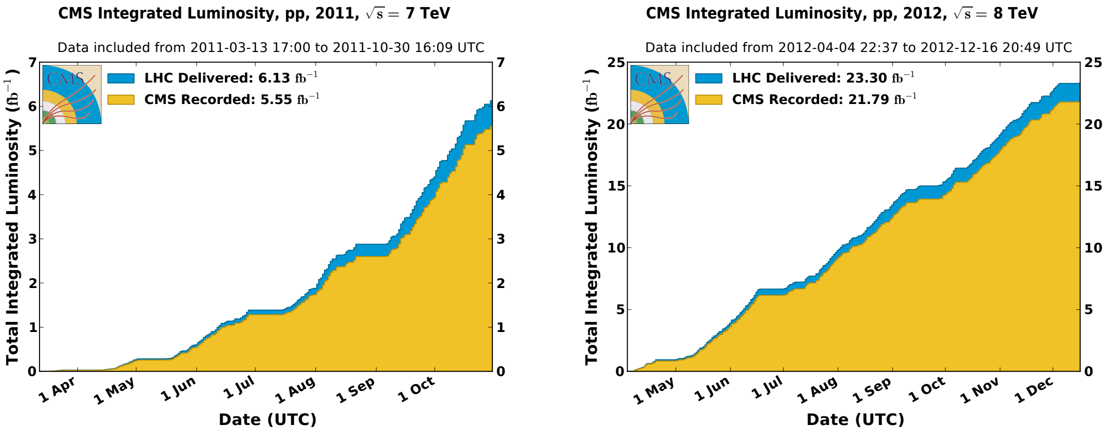

Figure 3.2: Cumulative luminosity versus day delivered to (blue), and recorded by CMS (orange) during stable beams and for pp collisions at 7 TeV center-of-mass energy in 2011 (left) and 8 TeV in 2012 (right).

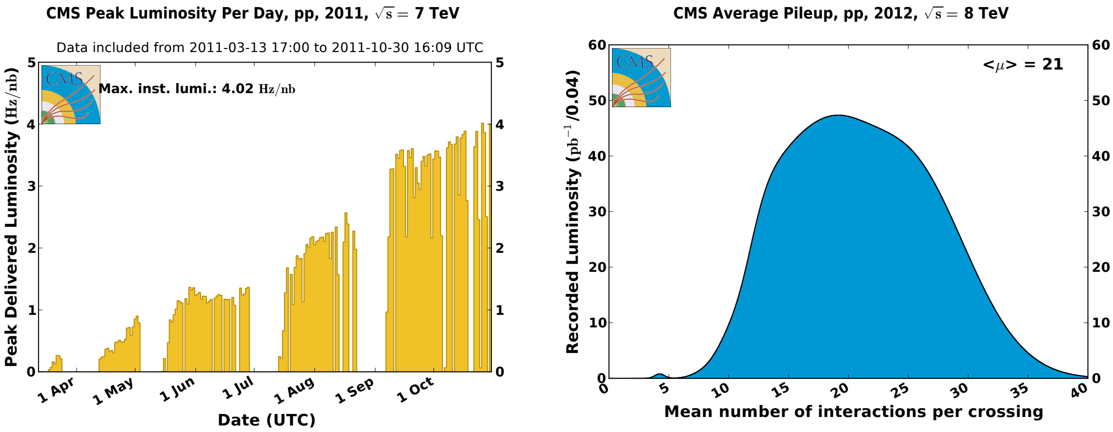

Figure 3.3: Left: peak delivered luminosity per day during the 2011 pp run. Right: mean number of interactions per bunch crossing in 2012.

3.2 THE COMPACT MUON SOLENOID

The Compact Muon Solenoid (CMS) [65] experiment, sketched in Fig. 3.4, is a 21.6 meters long, 15 m wide, general purpose, hermetic detector, designed around its superconducting magnet [66]. The main purposes of its design are a precise measurements of the muon momentum in the muon system, an excellent energy determination for electron and photon energy in the electromagnetic calorimeter, and a state-of-the-art tracking system to measure the transverse momentum and impact parameter of charged particles.

The CMS magnet is the largest of its type ever constructed and allows the tracker (pixel and strips) detectors, the electromagnetic (ECAL) and hadronic (HCAL) calorimeters to be placed inside the superconducting coil. The high current provides an homogeneous 3.8 T magnetic field of the strength [67]. The field is closed by a iron yoke which increases the field homogeneity reducing the stray field by returning the magnetic flux of the solenoid. In addition, the yoke is instrumented with four layers of muon stations, which provide for a measurement of the muon momentum independent of the inner tracking system. ATLAS, in contrast, chose to place only the tracker is inside its 2 T solenoid, and their calorimeters and the muon system are located outside of the magnet in two additional toroidal magnetic fields.

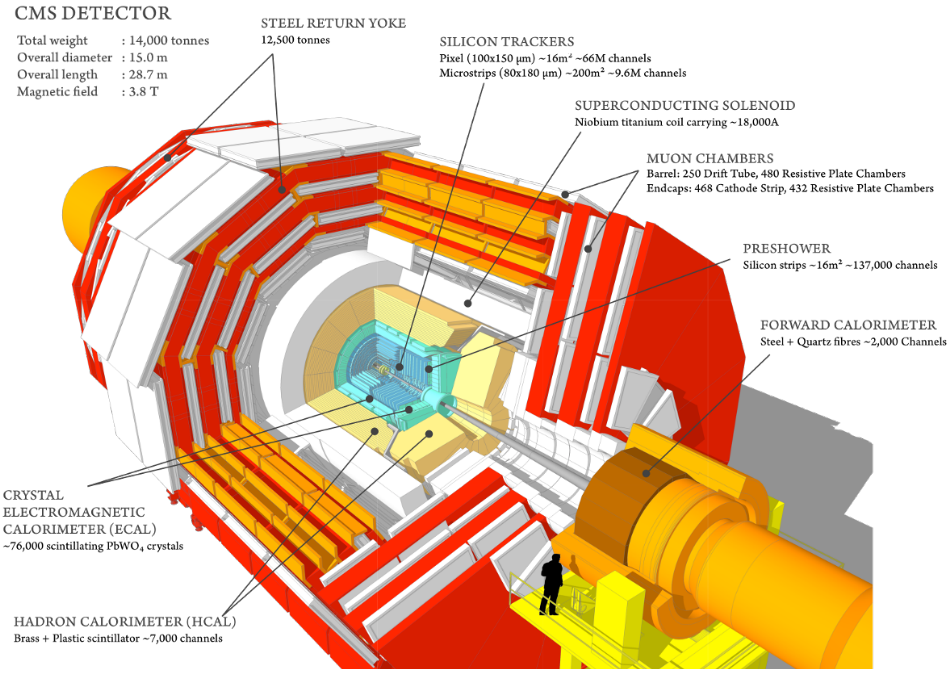

Figure 3.4: The Compact Muon Solenoid experiment.

The CMS coordinate system is oriented such that the x-axis points to the center of the LHC ring, the y-axis points vertically upward and the z-axis is in the direction of the beam towards the Jura mountains, with the origin of the axes located in the interaction point. The azimuthal angle φ is measured from the x-axis in the xy plane, and the radial coordinate in this plane is denoted as r. The polar angle θ is defined in the rz plane, and the pseudorapidity η is defined as $\eta = -\log\tan(\theta/2)$. This leads to an η of 0 for particles moving perpendicular to the beam direction and the beam direction itself has a pseudo-rapidity of $\pm\infty$ in the z direction. A quantity similar to the pseudorapidity is the rapidity y, defined as $y = \frac{1}{2}\log\left(\frac{E+p_z}{E-p_z}\right)$. For massless particles, rapidity and pseudorapidity are equivalent. The momentum component transverse to the beam direction, denoted p_{T}, is computed from the x-and y-components, and similarly the transverse energy is defined as $E_{\mathrm{T}} = E\sin\theta$. Both $\Delta\eta$ and $\Delta\varphi$ of two particles are independent of Lorentz boosts, therefore the distance between two particles can be measured in a third Lorentz-invariant variable, called ΔR and defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\varphi^2}$.

In the following a brief description of the main features of the CMS detector is reported, while a detailed description can be found in Ref. [65].

3.2.1 TRIGGER AND DATA ACQUISITION

At the nominal bunch crossing interval of 25 ns, the pp collisions rate provided by LHC is 40 MHz. Since the average disk space for a single event is of the order of few MBytes, the rate at which events can be recorded on tape is technically limited to few hundreds per second. Due to these limitations, events have to pass selections during data taking (*online*), drastically reducing the storage rate maintaining a high efficiency on the potentially interesting events [68]. The system dedicated to this task is the *trigger*, while the data acquisition system (DAQ) is dedicated to transfer data from subdetectors up to the storage systems. Selections made by a given trigger logic are identified by the *trigger path*, a string contains the name of the object, followed by a number which indicates the threshold, and a prefix to indicate how many objects are required.

The CMS trigger is implemented at two levels:

Level-1 trigger (L1): has to reduce the rate from 40 MHz (with 30 events per bunch crossing) to $\approx$ 100 kHz in less than 3 μs. The L1 trigger, consisting of customized processors, involves the calorimetric measurements and the muon system without looking at the tracker. The trigger decision is based on the so-called "trigger primitive", that is the presence and the number of objects like electrons, photons, muons, jets and τ-jets, and $E_{\mathrm{T}}^{\mathrm{miss}}$ with a transverse energy or p_{T} above a given threshold.

High level trigger (HLT): based on a based on a computer farm, consists in a multi-step procedure, reducing the event rate from $\sim$ 100 kHz to about 300 Hz before data storage. The first step, called L2, consists in the identification and measurements of particle candidates and global variables using only the information coming from calorimeters and muon system. In the next step, denoted as L2.5, information from the pixel detector can be exploited for track reconstruction and primary vertex identification. The last step, known as L3, runs the same algorithms employed by the offline reconstruction. Since at trigger level computing time is more critical than reconstruction accuracy, the algorithms are modified in order to be faster, even with a slightly lower precision. In order to meet the timing requirements given by the L1 input rate, events can be discarded before being fully reconstructed, as soon as the available information is enough to take the decision, or reconstruction can be limited only to a restricted region of the detector, identified by the L1 trigger object.

3.2.2 THE GLOBAL EVENT RECONSTRUCTION AT CMS

The global event reconstruction at CMS relies on the particle flow (PF) [69, 70] algorithm, which reconstructs and identifies each individual particle with an optimized combination of information from the various elements of the CMS detector. The energy of photons is directly obtained from the ECAL measurement, while the energy of electrons is determined from a combination of the electron momentum at the primary interaction vertex as determined by the tracker, the energy of the corresponding ECAL cluster, and the energy sum of all bremsstrahlung photons spatially compatible with originating from the electron track. The momentum of muons is obtained from the curvature of the corresponding track. The energy of charged hadrons is determined from a combination of their momentum measured in the tracker and the matching ECAL and HCAL energy deposits, and the energy of neutral hadrons is obtained from the corresponding ECAL and HCAL energy.

The core of the particle flow reconstruction technique is the algorithm used to link the signals of the individual subdetectors. The association criteria is purely geometrical. The aim is a full reconstruction of each single particle, while reducing any possible double counting from different detectors. Given the granularity of the calorimeters, energy deposits and tracks are linked if their trajectory intersects one of the calorimetric cells, and likewise clusters in the ECAL pre-shower, ECAL and HCAL are linked if the cluster position is compatible in the η, φ plane between the extrapolated track position and the cluster position.

3.2.3 TRACKER

The tracker [71] is the CMS innermost subdetector. It is placed inside the coil of the 3.8 T superconductive solenoid and it is designed to reconstruct the interaction primary and secondary vertices and to measure the momentum of charged particles. Charged hadrons are important to reconstruct hadronic jets, and are also useful to select isolated leptons or photons. The tracker performance is also crucial for the reconstruction of secondary vertices used to identify b jets, and to reject tracks from additional interactions. The coverage extends up to $|\eta| = 2.5$ with a surface of active silicon of 210 m^2. The CMS tracker is composed of two different subdetectors:

The *pixel tracker* detector is composed of approximately 66 million pixel cells, with size $100 \times 150\,\mu\mathrm{m}^2$, that allow a fine 3D vertex reconstruction. Pixels are used to ensure low cell occupancy (10^{-4} per pixel and pp collision) and a spatial resolution of about $10\,\mu$m in the $r\varphi$ plane and $15\,\mu$m in the z coordinate.

The *strip tracker* detector is divided in the barrel and endcaps regions. The barrel is divided into two parts, the Tracker Inner Barrel (TIB) and the Tracker Outer Barrel (TOB). The former is composed of four layers of silicon sensors with a thickness of $320\,\mu$m and of strip pitches varying from 80 to $120\,\mu$m. The TOB is made of six layers. In this subdetector thicker silicon sensors ($500\,\mu$m) are employed, while the strip pitch varies from 120 to $180\,\mu$m. The endcap region ($|\eta| > 1.6$) is covered by the Tracker Inner Disks (TID) and the Tracker End Cap (TEC). All four regions (TIB, TID, TOB, TEC) are instrumented with both single-sided and double-sided microstrip modules. The strips are oriented along the z direction in the barrel and along the r coordinate in the endcaps. The microstrip detector is designed to provide a spatial resolution of about $20 - 50\,\mu$m in the $r\varphi$ plane and about $200 - 500\,\mu$m along z.

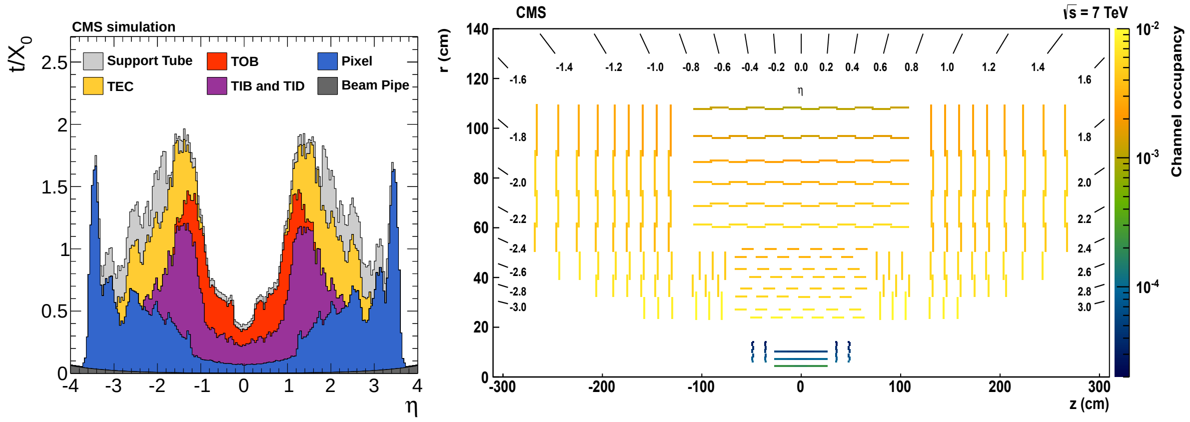

Figure 3.5: Left: channel occupancy in events taken with an average of nine pp interactions per beam crossing and an unbiased trigger. Right: Radiation length t/X_0 of the tracker material as a function of pseudorapidity η [72].

The material budget of the CMS tracker in units of radiation lengths (X_0) as estimated from simulation is shown in Fig. 3.5 [72]. It adds up to less than half a radiation length in the center of the barrel, increasing to a maximum of about 1.8 X_0 in the barrel-endcap transition. The high granularity of the pixel detector ensures that the channel occupancy of the inner pixel layer is much lower than that of the outer strip layer (Fig. 3.5).

3.2.3.1 TRACK RECONSTRUCTION

Trajectories of charged particles are reconstructed in CMS through multiple iterations of the Combined Track Finder (CTF) algorithm based on the *combinatorial Kalman Filter* [73] starting from the list of hits recorded by the tracker. The CTF algorithm consists of four different steps, and after each iteration hits associated to the high quality tracks candidates are removed from the input list.

The first step is *seed generation*, which provides an estimate of the helix parameters by using only pairs or triplets of hits compatible with the hypothesis of a track coming from the interaction region. Recognition begins with trajectory seeds created in the inner region of the tracker, producing 3D spatial measurements and facilitating reconstruction of low-momentum tracks deflected by the magnetic field before reaching the outer part of the tracker. In the *track finding* step, seed trajectories are extrapolated along the expected flight path of a charged particle. It builds track candidates by adding one compatible hit from the next detector layer updating the parameters. Once the track is completed another search is performed backward starting from the outermost hit to improve the hit collection efficiency. Once all the track candidate hits are collected, the trajectory is refitted in the *track fitting* step using the Kalman Filter and smoother methods to determine the most accurate estimate of the helix parameters. The last iteration, called *track filter*, maximizes the efficiency rejecting of fake tracks by applying some quality requirements. After all steps of iterative tracking have been performed, the output collections of the different steps are merged. Particles reconstructed by multiple steps are removed at the merging step by selecting only the best track among those that share a large fraction of hits.

The performances of the track reconstruction algorithm are evaluated by using simulated events and verified with collision data [72]. The average track-reconstruction efficiency for promptly-produced charged particles with $p_{\mathrm{T}} > 0.9\,\mathrm{GeV}$ is 94% in the barrel and 85% in the endcaps; the inefficiency is caused mainly by hadrons that undergo nuclear interactions in the tracker material. For isolated muons with $1 < p_{\mathrm{T}} < 100\,\mathrm{GeV}$ the tracking efficiency is $> 99\%$ [72]. For transverse momenta above a few GeV, the resolution on the impact parameter is mostly independent from η and approaches $\approx 30\,\mu\mathrm{m}$.

3.2.3.2 VERTICES RECONSTRUCTION

The position of the vertices resulting from the pp interactions in the same buch crossing (primary vertices) and those originated by heavy flavours and long-lived particles (secondary vertices), together with their corresponding uncertainty, is determined through a multistep procedure [74], starting from the reconstructed tracks in the event. In the first step, tracks originating from the primary interaction region are clustered using a *deterministic annealing algorithm* [75] based on the impact parameter along the z coordinate, followed by the application of an *adaptive vertex fitter* [76]. The vertex-finding algorithms can be very different depending on the physics case, whether the vertex to be reconstructed is primary or secondary, or specifically for exclusive decays. Vertex fitting involves determining the best estimate of the vertex parameters (position and covariance matrix) for a given set of tracks, with indicators of the fit quality, like the total χ^2, the number of degrees of freedom, or track weights. The primary vertex originating the hard scattering is chosen as the vertex with highest sum of p_{T}^2 of the associated tracks. The resolution depends on the event topology and is typically between $10 \sim 40\,\mu\mathrm{m}$ in the transverse plane and $15 \sim 50\,\mu\mathrm{m}$ in the z-direction.

3.2.4 ELECTROMAGNETIC CALORIMETER

The Electromagnetic Calorimeter (ECAL) measures the energy of electrons and photons, and is placed immediately outside the tracker, but still inside the solenoid. ECAL covers a fiducial region up to $|\eta| < 2.5$, and is made of 74 848 lead-tungstate ($PbWO_4$) crystals, a transparent material denser (8.3 g/cm^3) than iron, with a radiation length X_0 of 0.89 cm, a Molière radius R_M of 2.19 cm and with fast response (80% of light emitted within 25 ns). The compact dimensions of the crystals, while retaining the possibility to fully contain the expansion of the electromagnetic shower, makes ECAL particularly suitable to be placed inside the magnetic coil. The barrel crystals have a front face area of $2.2 \times 2.2\,\mathrm{cm}^2$, a length of 23 cm ($25.8X_0$) and they are positioned at $r = 1.29$ m. The crystals in the endcaps have a $2.47 \times 2.47\,\mathrm{cm}^2$ front face, a 22 cm length and they are positioned at $z = \pm 3.17$ m. For triggering purposes ECAL crystals are grouped together into 5×5 trigger towers. A pre-shower device is placed in front of the endcaps, made of two disks of lead absorber at $2X_0$ and $3X_0$, and of two planes of silicon strip detectors. It allows the rejection of photon pairs from π_0 decays and improves the estimation of the photon direction, enhancing the two-photon invariant mass resolution. The ECAL energy resolution can be parametrized by three different contributions:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c$$

where the first term is statistical in nature and contains fluctuation in showering and in the amplification through photodiodes ($a = 1.8\%$), the second considers electronic noise and pileup ($b = 4\%$) and the last term is related to calibration ($c = 0.5\%$).

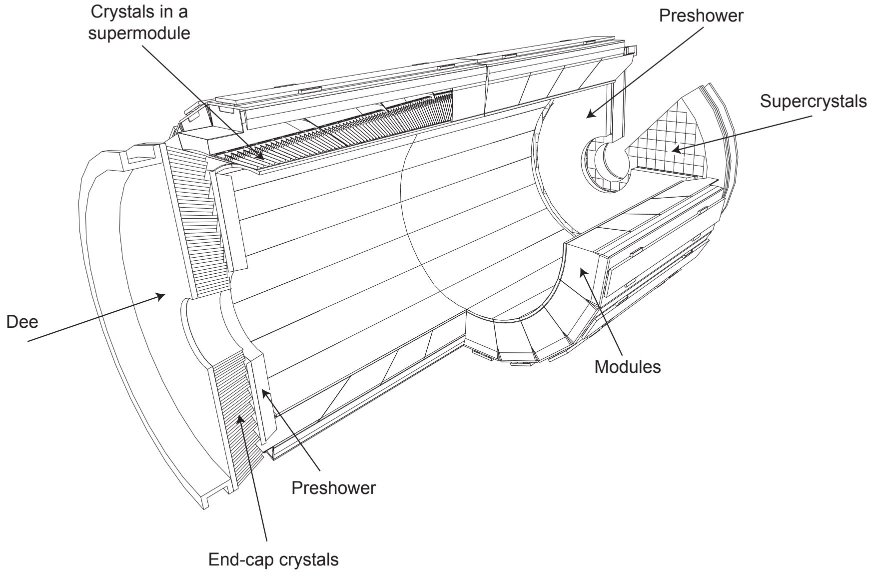

Figure 3.6: Illustration of the CMS electromagnetic calorimeter (ECAL).

3.2.5 ELECTRON IDENTIFICATION AND RECONSTRUCTION

Electrons are identified and reconstructed [77] by using the good track resolution of the silicon tracker and the high granularity of the electromagnetic calorimeter. Since electrons are stopped inside the ECAL, they are affected by the jet-induced background. Two complementary approaches are employed to reconstruct electrons:

Tracker seeding: the track is fitted starting from the tracker hits, and afterwards the calorimeter information is added; this method gives the best results for low-p_T electrons inside jets.

ECAL-driven seeding: the first step is to reconstruct a group of clusters of energy deposits in the calorimeter (SuperCluster). The information about the electron cluster characteristic width in η and φ, the latter depending also on the track bending due to the magnetic field, are taken into account. The supercluster position is then used to make a match to tracker seeds, and finally a global fit is performed, using an appropriate modeling of energy loss in the tracker material, which is not negligible for electrons. This method is the most adequate for electrons in a medium or high p_T range.

The bremsstrahlung energy loss distribution of electrons propagating in matter is highly non-Gaussian. Because the Kalman filter relies solely on Gaussian probability density functions, it is not necessarily the optimal reconstruction algorithm for electron tracks. A *Gaussian-sum filter* (GSF) algorithm for electron reconstruction in the CMS tracker has therefore been developed and implemented [78]. The bremsstrahlung energy loss distribution due to photons emitted in a direction tangent to the electron trajectory in the tracker material is modeled by a sum of Gaussians rather than by a single Gaussian, improving the resolution on reconstructed momentum with respect to the standard Kalman filter.

The identification of electrons relies on a Boosted Decision Tree (BDT) multivariate technique [77] that combines observables sensitive to the amount of bremsstrahlung along the electron trajectory, the geometrical and momentum matching between the electron trajectory and associated clusters, as well as shower-shape observables. The multivariate identification is trained using a Higgs boson Monte Carlo (MC) sample for the signal and a W + 1-fake electron data sample for background. This allows to train the BDT directly on a data control sample for a fake-enriched sample which is poorly described by simulation. The MVA selection reduces fakes by about half with respect to the cut-based working points while retaining the same identification efficiency, as seen in Fig. 3.7.

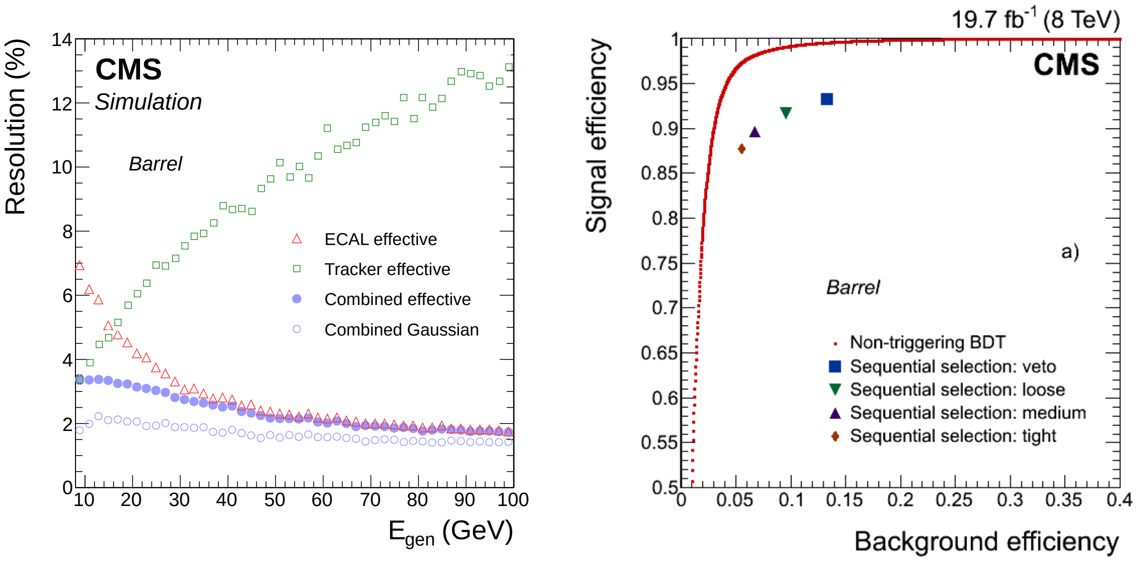

Figure 3.7: Left: electron resolution as a function of the measured electron p_T. Right: electron identification efficiency vs. misidentification efficiency for the multivariate and cut-based selections. Both refer to electrons in the barrel region.

The electron energy measurement [77] is obtained starting from the raw energy collected in ECAL superclusters, corrected for the imperfect containment of the clustering algorithm, losses due to bremsstrahlung and interactions upstream the ECAL and leakage arising from showers near gaps between crystals or between ECAL modules or supermodules. The technique used is a multivariate regression trained on a simulated $Z \rightarrow e^+e^-$ sample with different set of variables depending on whether the electron is detected in the ECAL barrel or endcaps. The pileup contribution is taken into account by including the average energy density ρ in the event. With respect to traditional supercluster corrections, this technique improves the electron resolution by about 10%.

Remaining small residual discrepancies between data and simulation are corrected by factorizing the crystals transparency loss effect as a function of the run period, the electron pseudo-

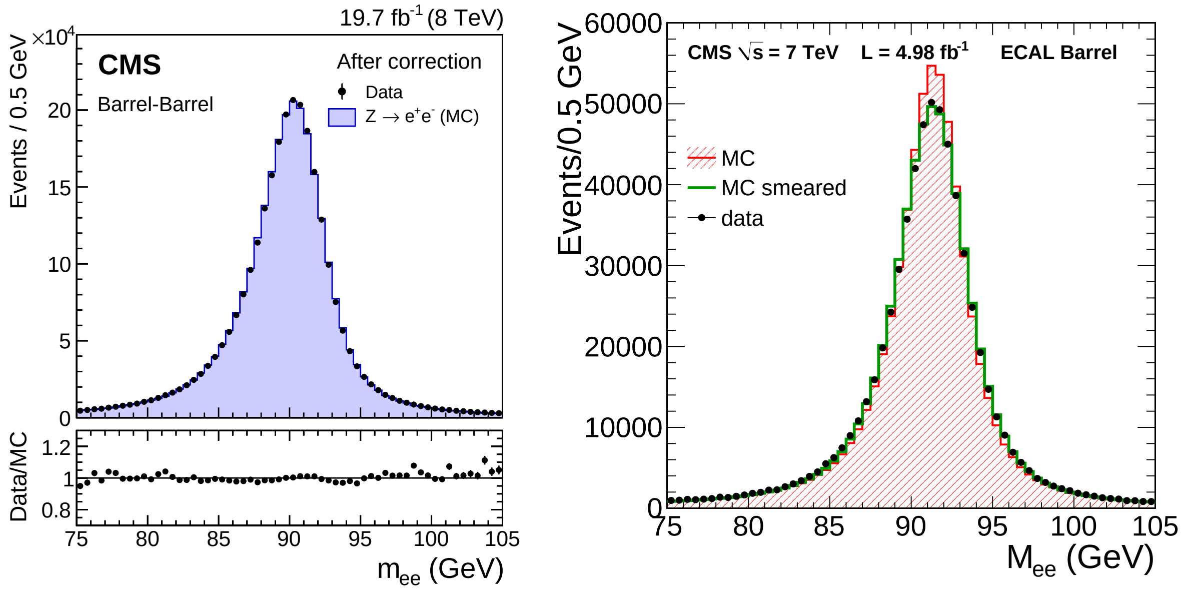

Figure 3.8: Effect of the electron energy regression (left) and smearing corrections (right) on the $Z \to e^+e^-$ peak for electrons in the barrel region.

rapidity, and the bremsstrahlung pattern. Correction factors are below 0.5% and ≈ 1.5% in the barrel and endcap regions. In addition, the energy in the simulation is modified applying a random Gaussian multiplicative factor to match the electron resolution measured in data.

3.2.6 HADRONIC CALORIMETER

The hadronic calorimeter (HCAL) [79] is placed just outside ECAL and inside the magnet coil. The CMS HCAL is a sampling calorimeter, whose active elements are plastic scintillators interleaved with brass absorber plates and read out by wavelength-shifting fibers. Brass has been chosen as absorber material for its short interaction length and because it is non-magnetic. The thickness of the absorber layers is between 60 mm in the barrel and 80 mm in the endcaps. The barrel ranges between 5.46 interaction lengths at $\eta \approx 0$ to 10.8 at $\eta \approx 1.3$, while the endcaps have an average of 11 interaction lengths. The photodetection readout is based on multi-channel hybrid photodiodes, able to operate in a high magnetic field, that give an amplified response proportional to the original signal for a large range of particle energies. The HCAL energy resolution is parametrized as:

$$\frac{\sigma}{E} \approx \frac{a}{\sqrt{E}} \oplus b$$

where $a \approx 65\%$ in the barrel, ≈ 85% in the endcaps and ≈ 100% in the hadron forward calorimeter, while b is a flat 5% over the full η range.

HCAL plays also an important role in the reconstruction of missing energy. The design is strongly influenced by this goal, hence an important requirement is a high hermeticity. Such features require the detector to cover the biggest possible portion of the solid angle. For this reason, the barrel and endcaps are complemented by a hadron forward calorimeter, which is placed outside the magnet return yokes, at 11 m from the interaction point, with a total coverage of $3 < |\eta| < 5.3$. Moreover, an outer hadronic calorimeter is placed in the first muon absorber layer in order to enhance the containment of high-energy jets in the central region of the detector.

3.2.7 JET RECONSTRUCTION

Most of the processes of interest at the LHC contain quarks or gluons in the final state. While partons are not directly observable they manifest themselves through hadronization to stable particles which can be detected in tracking chambers and calorimeters. Perturbative theory and the hadronization model describe the interaction between constituent partons of the protons and the subsequent showering into stable particles.

CMS reconstructs jets starting from all the PF candidates in the event (hence called *PFJets*), namely muons, electrons, photons, charged hadrons and neutral hadrons. Particle flow charged hadrons candidates are identified as tracks in the inner tracker, and linked to calorimetric deposits if the particle p_T is large enough ($> 750\,\mathrm{MeV}$). If the tracker and calorimeter measurements are compatible, the best energy determination is obtained as a combination of the two, after accounting for non-linearities and zero suppression effects. If instead an excess of calorimetric energy deposition is found with respect to the momentum of the associated track, or calorimetric deposits are not linked to any tracks, the energy is identified as coming from a photon or a neutral hadron. If the track momentum exceeds significantly the measured calorimetric energy, the particle is considered a minimum ionizing particle and is identified as a muon after satisfying very loose identification criteria. In this case, linking between tracks in the tracker and a muon tracks in the muon system is established as well. Linked objects are processed and reconstructed as particle candidates; composite objects, like jets, are built starting from the list of particle candidates in the event.

Particle candidates are merged together into jets by clustering algorithms, implemented in the *FastJet* software [80]. The most widely used in CMS is the *anti-kT* [81, 82], with a radius parameter $R = 0.5$. The *anti-kT* is a sequential recombination jet algorithm, that merges pairs of particle candidates in order of increasing relative transverse momentum into jets, until a stopping requirement is achieved, typically when the distance between adjacent jets is greater than some value. The distance between two candidates i and j is defined as:

$$d_{ij} = min(p_{Ti}^{2k}, p_{Tj}^{2k}) \frac{\Delta R_{ij}^2}{R^2}$$

with ΔR_{ij} being the distance in the $y - \varphi$ plane. In addition to the radius parameter R, the parameter k governs the relative power of the energy versus geometrical (ΔR_{ij}) scales. The $k = 1$ choice corresponds the inclusive-kT algorithm, while the case of $k = 0$ is the inclusive *Cambridge/Aachen* algorithm [83]. The *anti-kT* algorithm considers $k = -1$. With this choice, the distance d_{ij} between a soft and a hard particle is dominated by the hard-particle p_T. Instead, two soft particles with a similar separation ΔR_{ij} would have a larger distance d_{ij}. As a consequence, soft particles will tend to cluster with hard ones before they cluster among themselves. If a hard particle has no hard neighbors within a distance $2R$, then it will simply accumulate all the soft particles within a circle of radius R, resulting in a conical jet. The key feature is that the soft particles do not modify the shape of the jet, unlike hard particles, making the jet boundary resilient with respect to soft radiation, but flexible to hard radiation. This feature is known as *infrared safety*. The anti-kT is also *collinear safe*, meaning that splitting an hard candidate in two or more collinear candidates with softer p_T does not change the result of the jet clustering sequence.

In order to reduce reconstruction biases due to pileup interactions, particle candidates whose impact parameter is not compatible with the primary vertex are removed from the clustered jets. This procedure is known as *charged hadron subtraction* (CHS). The jet momentum is determined as the vectorial sum of all particle momenta in this jet (*raw jets*), and at this stage the reconstructed jet energy has a bias caused by different reasons, the most important of which are the non-linear response of the calorimeters, the detector segmentation, the presence of material in front of calorimeters, electronic noise, and noise due to additional interactions from the same bunch crossing. Several levels of corrections, sketched in Fig. 3.9, are applied to the raw jets in order to obtain the energy value that is closer to the true energy of the initial parton [84]:

L1 Offset: the pileup and electronic noise effects are removed. This correction can be estimated using events collected by a random trigger, without any preconditions except a beam crossing, referred as *zero bias* events. The offset contribution from pileup is estimated by the FastJet method which relies on the definition of a jet area [85] from which a median energy density (ρ, in GeV/Area) per event can be defined. The correction subtracted to the jet p_{T} equals to ρ times the jet area. FastJet has the advantage of being able to remove the out-of-time pileup component, but has the disadvantage of subtracting the underlying event contribution as well.

L2 Relative (η): the variation in jet response with η is flattened. The unbalance between the jets transverse momentum that is observed on average, is due to the variation of the jet response across the detector versus η.

L3 Absolute (p_{T}): the calorimetric energy response varies as a function of the jet p_{T}. The absolute correction removes these variations and makes the response equal to unity. This correction is obtained from simulation using the Monte Carlo truth information.

L2L3 Residual: differences between data and simulation after L2 and L3 corrections are removed by applying a specific calibration to Monte Carlo events. Residual corrections are extracted from data using the transverse momentum balance in γ+jets and Z +jets events [84].

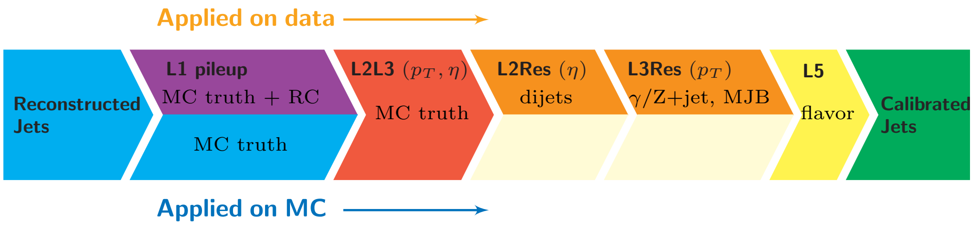

Figure 3.9: Summary of the CMS jet energy correction steps.

The overall uncertainty on the jet energy scale for jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ is smaller than 3% in the barrel and 5% up to $|\eta| \leq 4.5$. Figure 3.10 shows the jet energy correction (JEC) uncertainties as a function of the jet p_{T}. Since measurements on data show that the jet energy resolution (JER) is not the same in data and MC, corrections are applied in simulation, smearing by about 5% jets in the barrel region, and up to 15% jets in the endcaps [84]. The jet energy resolution, shown in Fig. 3.11, amounts typically to 15% at 10 GeV, 8% at 100 GeV, and 4% at 1 TeV, to be compared to about 40%, 12%, and 5% obtained when the calorimeters alone are used for jet clustering [84]. Jets are also requested to pass loose identification criteria, in order to reject fake jets due to calorimeter noise, with more than 99% efficiency for true jets.

3.2.8 B JETS TAGGING

Algorithms for b-jet tagging allow to distinguish jets originating from heavy flavours (B hadrons) and lighter flavours (including D hadrons). The tagging algorithms rely on several distinct properties of B hadrons: a large proper lifetime ($\tau \sim 1.5$ *ps*, $c\tau \sim 450$ μm), a large mass, and relatively large semileptonic branching ratios. In about 40% of the cases, B hadron decays involve also muons or electrons, from both direct decays from the b quark and cascade decays

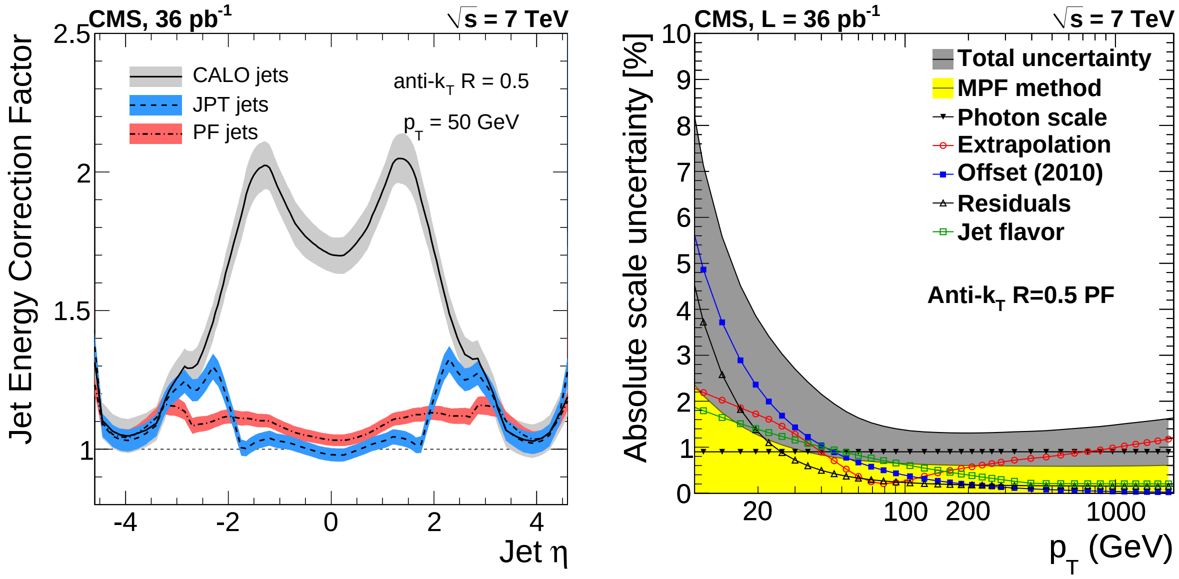

Figure 3.10: Jet Energy scale corrections for jets as a function of jet η for a typical jet $p_T = 50\,\text{GeV}$ (left) and jet energy correction uncertainties for each single source.

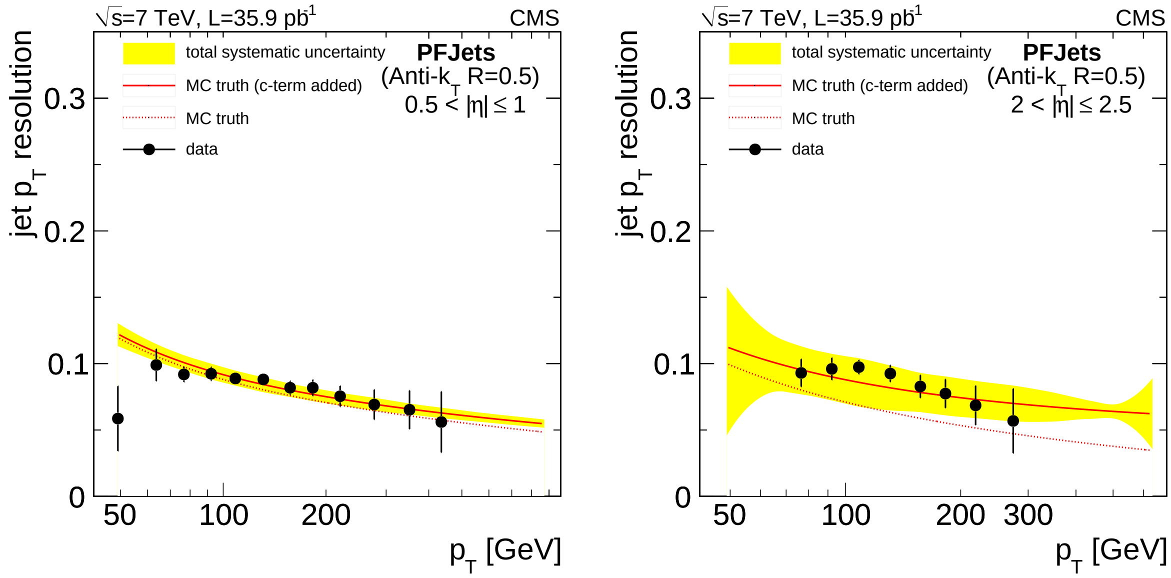

Figure 3.11: Jet Energy resolution for PFJets in two different η regions, $0 < |\eta| < 0.5$ (left) and $2 < |\eta| < 2.5$ (right).

from c quarks in D hadrons often produced from the decay of B hadrons.

The b-tagging algorithms rely upon the reconstruction of lower-level physics objects, mainly jets and tracks. Most of the B hadron properties are exploited using charged particle tracks because only tracking detectors offer the spatial resolution needed to detect the significant decay length of B hadrons. Efficient track reconstruction, and in particular precise spatial reconstruction close to the interaction point, is the key ingredient for almost all b-tagging algorithms. Tracks are required to pass minimal quality requirements: having a p_T has to be larger than 1 GeV, being reconstructed by at least 8 hits, with at least 2 of them in the pixel detector, and being considered associated to the primary vertex ($d_z < 2\,\text{mm}$); this removes tracks from pileup interactions and long-lived particles. Furthermore, the χ^2/ndf of the track fit should be < 10.

The physics motivation of the positively-signed track impact parameter requirement is that the track should be consistent with a particle decaying in a point with some positive displacement from the primary vertex along the direction of the jet (Fig. 3.12); on the contrary, a negative impact parameter corresponds to a decay point behind the primary vertex. Negative impact parameters arise mostly as a result of the finite detector resolution which results in mis-measured track parameters, and can be used to estimate the mistagging performance of jets that do not originate from b and c quarks.

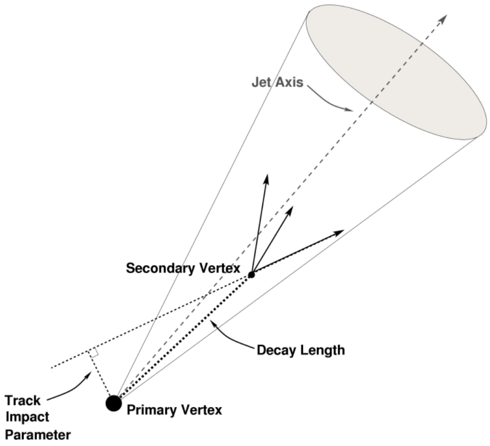

Figure 3.12: Sketch of the main characteristics of a b jet.

The association of tracks to jets is performed via a cone-based jet-track $\Delta R = \sqrt{\Delta\eta^2 + \Delta\varphi^2} < 0.3$ distance criterion. Several b-tag algorithms have been implemented by the CMS collaboration, each one with its own peculiarities. Some exploit the long B hadron lifetime by looking at displaced tracks (*track-based* algorithms) and the reconstruction of a secondary vertex within the jet (*vertex-based* algorithms), others its semi-leptonic decay mode. The best performance, expressed as b jet tagging efficiency versus c- or light-jet mistagging, is achieved by combined algorithms which collect in a single discriminator the information provided by all the possible variables (*combined algorithms*).

Track Counting: a simple b tagging method, based on the simple requirement of a minimum number of good-quality tracks with an impact parameter significance exceeding a given threshold. Tracks are ordered by decreasing signed impact parameter significance, which serves as the discriminator. The *high efficiency* algorithm (TCHE) uses the second track, while the *high purity* (TCHP) chooses the third.

Jet Probability: a more sophisticated track-based algorithm, also exploiting the long lifetime of B hadrons and based on track impact parameters. It entails computing the compatibility of a set of tracks with having come from the primary vertex. Tracks with negative impact parameter are used to extract a resolution function, which is used to calibrate the impact parameter significance distribution. The (signed) probability is flat between -1 and 1 for tracks coming from the primary vertex, and positive and concentrated near 0 for tracks with large impact parameter significance. A slight variation of the *JetProbability* is the *JetBProbability*, which takes into account only the four most displaced tracks (the average number of charged particles from b hadron decay).

Soft lepton: tags b jets by searching for electrons or muons from the semi-leptonic B hadron decay, which typically has a large momentum with respect to the jet axis and a large impact parameter. The b tag discriminator is the output of a neural net trained on four characteristic variables, the p_{T}^{rel} (the lepton p_{T} relative to the jet direction), the 3D impact parameter significance of the lepton track, the ratio between the lepton p_{T} and the jet energy, and the angular separation between the lepton and the jet axis. Although the efficiency of these lepton-based algorithms is limited by the intrinsic B$\rightarrow \ell\nu + X$ branching ratio, the information can be integrated in the more performing combined algorithms.

Secondary vertex: based on the reconstruction of the secondary decay vertex of the weakly decaying B hadron, using an adaptive vertex finder, and it uses variables related to it to calculate its b tag discriminator. Secondary vertices are required to have a distance from the primary vertex to the secondary vertex in the transverse plane between 100 μm and 2.5 cm and a significance larger than 3. Additionally, the invariant mass of charged particles associated to the vertex must not exceed 6.5 GeV, and the vertex must not be compatible with a K^0 decay. Two vertex-based algorithms are available, the *high efficiency* (SSVHE) and the *high purity* (SSVHP) which use primary vertices composed by at least two and three track, respectively.

Combined Secondary vertex: allows to overcome the intrinsic $50-65\%$ reconstruction efficiency of the secondary vertex that limits the vertex-based efficiency algorithms. This combined algorithm (CSV) merges the information about track impact parameters and secondary vertices, providing discrimination even when no secondary vertices are found. Jets are divided in three exclusive categories, depending on the presence of a reconstructed secondary vertex ("*vertex*"), or at least two tracks with impact parameter significance larger than 2 ("*pseudo-vertex*"), or none of the previous ("*no vertex*"), where the algorithm relies on track-based variables alone. A likelihood discriminator is built and trained separately on the three categories using the following variables, with a high discriminating power and low correlation:

- the number of tracks in the jet
- the number of tracks associated to the secondary vertex
- the secondary vertex mass
- the 2D flight distance significance (σ_{IP}/IP) of the secondary vertex
- the ratio of the transverse momentum of the tracks at the secondary vertex with respect to all tracks in the jet
- the η separation of the secondary vertex tracks with respect to the jet axis
- the 2D IP significance of the first track ordered by decreasing IP significance that raises the secondary vertex mass above the charm threshold
- the 3D signed IP significance for each track in the jet

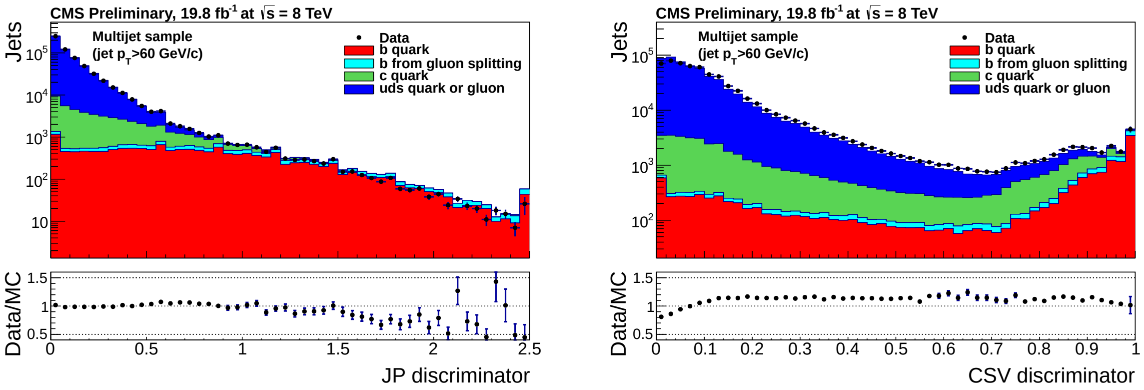

Figure 3.13: Reference distributions for the *JetProbability* (left) and *Combined Secondary Vertex* algorithms (right), for each flavour contribution.

The efficiency to tag b jets and the rate of misidentification of non-b jets depend on the cut on the algorithm discriminant. Three main working points are defined, the loose (L), medium (M), and tight (T) such that the misidentification probability for light-parton jets is close to 10%, 1%, and 0.1% respectively. For the CSV algorithm and the tight working point (CSVT) the efficiencies to tag b quarks, c quarks, and light quarks or gluons are approximately 50%, 6%, and 0.12%, respectively [86]. The performance of tagging algorithms are measured on independent samples of multijet events, exploiting the B hadron decays into muons, and $t\bar{t}$ events in both the lepton+jets and dilepton final state [86]. The probability of light-flavour quark and gluon jets being misidentified as b jets is evaluated with negative tagging algorithms, identical to the original ones, but using only on tracks with a negative impact parameter or secondary vertices with a negative decay length. Multiplicative scale factors to account for different tagging and mistagging efficiencies between data and simulation, for different jet flavours $i = b, c, l$, are defined as $SF_i = \varepsilon_i^{data}/\varepsilon_i^{sim}$. There are different ways to take scale factors into account. The simplest one is to consider a per-event weight given the number of tagged jets and their flavour. Other methods allow to modify the value of the tagger discriminator in simulation to match the efficiency measured in data [33].

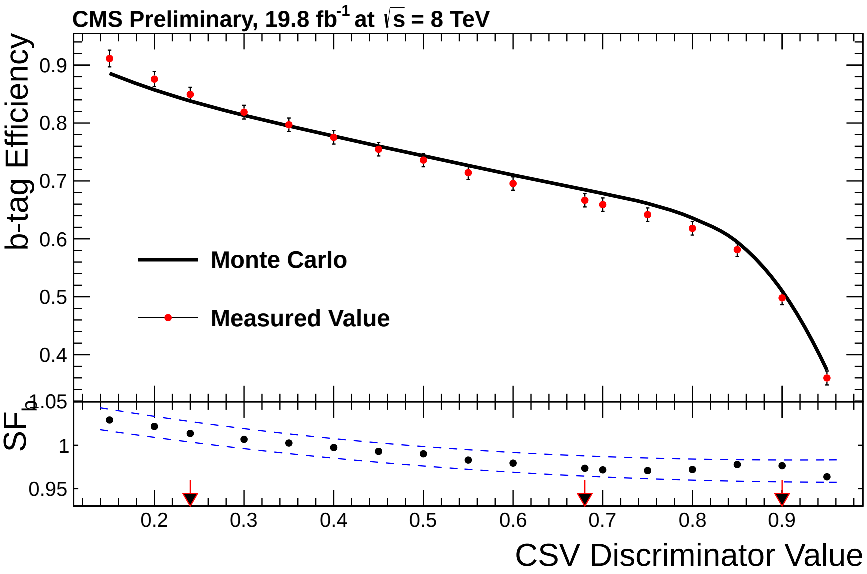

Figure 3.14: Tagging efficiency for b jets of the CSV algorithm measured in data and simulation. The bottom plot show the ratio of data and simulation efficencies, corresponding to the scale factors.

3.2.9 MUON SYSTEM

A good muon detection, precise momentum measurement and efficient triggering are of primary importance as muons represent a clear signature for a large number of physics processes. The CMS muon system [87] fulfills these purposes. It is hosted in the return yoke outside the solenoid, that shields the muon detectors from electromagnetic and hadronic particles that are not contained inside the calorimeters. Furthermore, the yoke provides a magnetic field between consecutive muon stations, allowing a momentum measurement independent from the inner tracker.

Three different types of subdetectors, sketched in Fig. 3.15, compose the muon system:

Drift Tubes (DT): four layers of Drift Tube chambers are placed in the barrel region ($|\eta| < 1.2$), where the particle rate and the residual magnetic field are low. The chamber segmentation follows the iron yoke structure, consisting of 5 wheels along the z-axis, each one divided into 12 azimuthal sectors. Each DT chamber, on average 2 m × 2.5 m in size, consists of 12 aluminium layers, arranged in three groups of four, each with up to 60 tubes: the middle group measures the coordinate along the direction parallel to the beam and the two outside groups measure the perpendicular coordinate. The muon position in each DT is reconstructed by measuring the drift time of the ionization electrons, and converting it into a distance from the wire. Each one of the 250 DT chambers has a resolution of $\sim 100\,\mu$m in $r\varphi$ and 1 mrad in φ.

Cathode Strip Chambers (CSC): in the two endcaps ($0.8 < |\eta| < 2.4$), where the flux of hadron punch-through and radiation is much higher and the magnetic field is strongly varying, 540 Cathode Strip Chambers are used. In each of the endcaps the chambers are arranged in 4 disks perpendicular to the beam, and in concentric rings (3 rings in the innermost station, 2 in the others). Each chamber has a spatial resolution of about 200 μm in r, and $75 - 150\,\mu$m in the $r\varphi$ coordinate.

Resistive Plate Chambers (RPC): In both the barrel and the endcaps, a system of 912 Resistive Plate Chambers is installed, ensuring redundancy to the measurement. RPCs provide a rougher spatial resolution than DTs and CSCs, but the fast response with a good time resolution (1 ns) is used for triggering purposes.

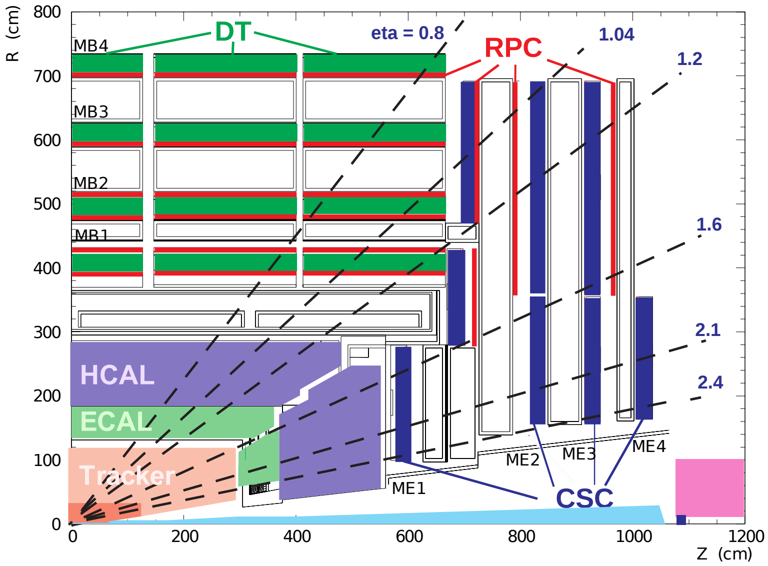

Figure 3.15: Section view of the CMS muon chambers and their coverage as a function of detector pseudorapidity [88].

3.2.10 MUON RECONSTRUCTION

In the standard CMS reconstruction for pp collisions [88], muon tracks are first reconstructed independently in the inner tracker (tracker track) and in the muon system (Standalone Muon track). Standalone Muons are reconstructed using segments and hits from DT, CSC, and RPC, using a Kalman filter technique, which provides direction, p_{T}, and charge of the muon candidate.

Based on those objects, two reconstruction approaches are used [88]: *Global Muon* (outside-in) and *Tracker Muon* (inside-out).

Global Muon reconstruction (outside-in): for each Standalone Muon track, a matching tracker track is found by comparing parameters of the two tracks propagated onto a common surface, and a Global Muon track is fitted combining hits from the tracker track and Standalone Muon track, using the Kalman-filter technique [73]. At large transverse momenta, $p_{\mathrm{T}} > 200\,\mathrm{GeV}$, the Global Muon fit can improve the momentum resolution compared to the tracker-only fit.

Tracker Muon reconstruction (inside-out): in this approach, all tracker tracks with $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}$ and the total momentum $p > 2.5\,\mathrm{GeV}$ are considered as possible muon candidates and are extrapolated to the muon system taking into account the magnetic field, the average expected energy losses, and multiple scattering in the detector material. If at least one muon segment (i.e., a short track stub made of DT or CSC hits) matches the extrapolated track, the corresponding tracker track qualifies as a Tracker Muon.

Tracker Muon reconstruction is more efficient than the Global Muon reconstruction at low momenta, $p_{\mathrm{T}} \lesssim 5\,\mathrm{GeV}$, because it requires only a single muon segment in the muon system, whereas Global Muon reconstruction is designed to have high efficiency for muons penetrating through more than one muon station and typically requires segments in at least two muon stations. Thanks to the high tracker-track efficiency and a very high efficiency of reconstructing segments in the muon system, about 99% of muons produced in pp collisions and having sufficiently high momentum are reconstructed either as a Global Muon or a Tracker Muon, and very often as both. Muons reconstructed only as Standalone Muon tracks have worse momentum resolution and less favorable collision muon to cosmic-ray muon ratio than the Global and Tracker Muons and are usually not used in physics analyses.

Muons used in the present analysis are based on the *Particle Flow Muon* selection, considering Global Muon or a Tracker Muon candidates and by applying minimal requirements on the track components in the muon system and taking into account a matching with small energy deposits in the calorimeters. In order to further increase purity, additional cuts are applied, matching the *Tight* identification working point. The normalized χ^2 of the muon track is required to be smaller then 10, and a minimum number of hits are requested in the muon system (> 0), pixel (> 0) and strips (> 5), with at least one matched station. Cuts on the impact parameters select muons with a d_{xy} and d_z smaller than 0.2 cm and 0.5 cm respectively.

The precise measurement of lepton momentum is one of the crucial points for physics analyses which aim at reconstructing narrow resonances with the best possible accuracy. For this reason, the reconstructed momenta of the muons are corrected both in data and in simulation in order to remove remaining local biases in the measured p_{T}, mainly due to the non-perfect knowledge of the detector alignment and magnetic field and the presence of material in front of the sensors. Corrections were derived with the *MuScleFit* [88, 89] algorithm on samples of real and simulated $\mathrm{Z} \to \mu^+\mu^-$ events. MuScleFit is based on an unbinned maximum likelihood fit which assigns a probability based on the dimuon invariant mass and the individual muon kinematics, taking into account the natural width of the resonance, the presence of possible biases in reconstructing the muon p_{T}, η, φ and charge, and the finite resolution of the detector. The last two terms enter in the likelihood though a parametrization and it is precisely the goal of the maximization to determine the best values of the parameters.

The corrections mentioned above calibrate the overall momentum scale and remove the dependency of the scale on the p_{T}, η, φ and charge of the muon, thereby improving also the momentum resolution. In the central part of the detector and for the lepton momentum range $p_{\mathrm{T}} < 100\,\mathrm{GeV}$ relevant for this analysis the typical muon momentum resolution after the above momentum correction and smearing is $\delta p_{\mathrm{T}}/p_{\mathrm{T}} \approx 1-3\%$. After the calibration, the relative momentum scale between data and simulation is consistent within 0.1% (Fig. 3.16).

3.2.11 LEPTON ISOLATION

Lepton isolation is defined as the ratio of the p_{T} sum of all charged and neutral particles in the event within a $\Delta R = 0.4$ cone around the lepton, divided by the lepton p_{T} itself:

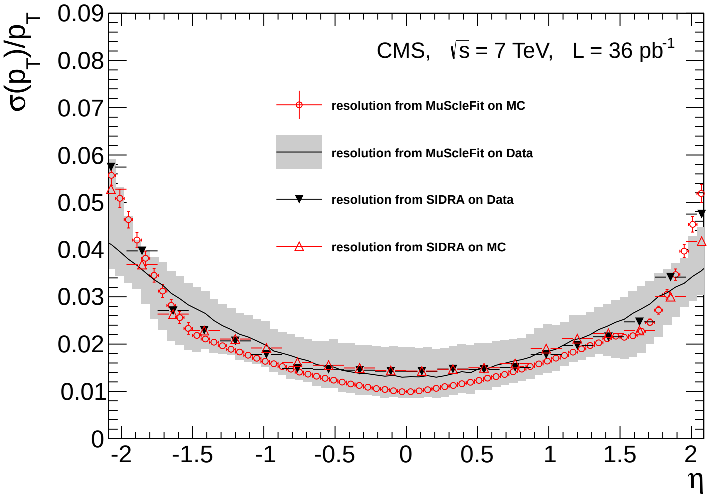

Figure 3.16: Relative transverse momentum resolution $\sigma(p_{\mathrm{T}})/p_{\mathrm{T}}$ in data and simulation measured by applying two different calibration algorithms, MuScleFit and SIDRA, to muons produced in the decays of Z bosons and passing a Tight selection. The thin line shows the result of MuScleFit on data, with the gray band representing the overall (statistical and systematic) 1σ uncertainty of the measurement. The circles are the result of MuScleFit on simulation [88].

$$I_{rel} = \left[\sum p_{\mathrm{T}}^{\mathrm{ch\ had}} + \max(\sum p_{\mathrm{T}}^{\mathrm{neu\ had}} + \sum p_{\mathrm{T}}^{\gamma} - 0.5 \cdot \sum p_{\mathrm{T}}^{\mathrm{pu\ ch\ had}},\ 0)\right] / p_{\mathrm{T}}^{\ell}$$

Besides charged particles and photons, an additional term is introduced to subtract the pileup contribution. Since charged hadrons coming from other vertices can be easily subtracted from the event and thus do not enter in the first term, the same procedure cannot be done for the neutral particles. The contribution of pileup neutral deposits is, with a good approximation, about half of the charged ones, and this term is subtracted from the neutral component. Isolated leptons are selected by applying a 0.15 cut for electrons, and 0.12 for muons.

3.2.12 MISSING ENERGY

The $E_{\mathrm{T}}^{\mathrm{miss}}$ is the imbalance in the transverse energy of all visible particles, and it is reconstructed with the particle flow algorithm [69, 90]. The *raw* $E_{\mathrm{T}}^{\mathrm{miss}}$ is defined as the inverse vectorial sum of the transverse momentum of all the reconstructed charged and neutral particle flow candidates: $\vec{E}_{\mathrm{T}}^{\mathrm{miss}} = -\sum_{i=0}^{all} \vec{p}_{\mathrm{T}i}$. The raw $E_{\mathrm{T}}^{\mathrm{miss}}$ is systematically different from true $E_{\mathrm{T}}^{\mathrm{miss}}$, for many reasons including the non-compensating nature of the calorimeters and detector misalignment. To better estimate the true $E_{\mathrm{T}}^{\mathrm{miss}}$, corrections can be applied:

Type-0: a mitigation for the degradation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction due to the pileup interactions, by applying the CHS algorithm. However, the $E_{\mathrm{T}}^{\mathrm{miss}}$ contribution from pileup neutral particles cannot be easily subtracted; the assumption is that the $E_{\mathrm{T}}^{\mathrm{miss}}$ contribution term of charged and neutral pileup particles are the same, and cancellation at the true level is exact: $\sum_{neuPU} \vec{p}_{\mathrm{T}_i}^{\,true} + \sum_{chPU} \vec{p}_{\mathrm{T}_i}^{\,true} = 0$. An additional $E_{\mathrm{T}}^{\mathrm{miss}}$ term is then added to the raw $E_{\mathrm{T}}^{\mathrm{miss}}$ to take into account the neutral PU contribution, which is equal to the charged one with a multiplicative scale factor taking into account calorimeter mismeasurements of low-p_{T} energy deposits.

Type-1: propagation of the jet energy corrections (JEC) to MET. The Type-I correction replaces the vector sum of transverse momenta of particles which can be clustered as jets with the vector sum of the transverse momenta of the jets to which JEC is applied.

Particle flow $E_{\mathrm{T}}^{\mathrm{miss}}$ with type-0 and type-1 corrections applied is currently the default one used by CMS physics analyses; Figure 3.17 show the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for data and Monte Carlo after the corrections. Another important variable, beside $E_{\mathrm{T}}^{\mathrm{miss}}$ itself, is the missing transverse energy significance ($\mathcal{S}$). While a first version of the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance algorithm is described in Ref. [90], the current definition is:

$$\mathcal{S} = \frac{\mathcal{L}(\vec{\varepsilon} = \sum \vec{\varepsilon}_i)}{\mathcal{L}(\vec{\varepsilon} = 0)}$$

The numerator expresses the likelihood of the hypothesis under test that the true value ($\vec{\varepsilon}$) of the missing transverse energy is equal to the observed value ($\sum \vec{\varepsilon}_i$). The denominator instead expresses the likelihood of the null hypothesis, that the true missing transverse energy is actually zero. Under the null hypothesis, observation of any non-zero missing transverse energy is attributed to resolution smearing. The discriminating power of the significance with respect to the $E_{\mathrm{T}}^{\mathrm{miss}}$ is shown in Fig. 3.17.

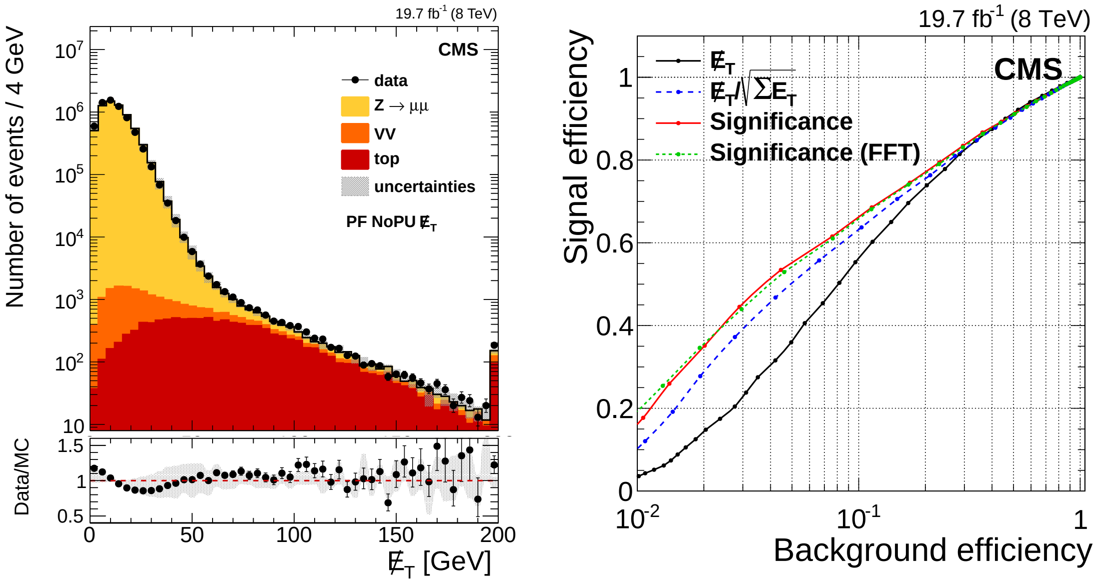

Figure 3.17: Right: No-PU $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in $Z \to \mu^+\mu^-$ events; the points in the lower panel of each plot show the data/MC ratio, including the statistical uncertainties of both data and simulation. Left: signal versus background efficiencies for $W \to e\nu$ for various $E_{\mathrm{T}}^{\mathrm{miss}}$ based discriminating variables [91].

Chapter 4

Search for the Higgs to $b\bar{b}$ produced in association with b quarks

At the end of 2011, with about 5 fb^{-1} of data delivered by LHC, there was no conclusive evidence of the SM Higgs boson, although some mass region were excluded, and small excesses over background expectations were observed by the ATLAS and CMS experiments with a combined statistical significance of about 3σ. Although these results were fully compatible with a light SM Higgs boson hypothesis due to the rather large uncertainty, some channels showed deviations from the expected values, posing the question if the hints of the new particle was due to the Higgs boson predicted by the SM or not, as suggested by some BSM interpretations. Furthermore, the Tevatron experiments had published few years earlier the results for the search for BSM Higgs bosons in the multi-b channels, showing an excess larger than 2σ in the low mass region. In this perspective, the search for a Higgs boson in channels suppressed in the SM was motivated by the theoretical and experimental scenario at the time. Even if the best sensitivity for high-$\tan\beta$ *scenarios belongs to the* $\Phi \to \tau\tau$ *channel, the Higgs boson decay into a pair of b quarks in the b quark associated production was a final state not yet studied at LHC. This process, characterized by an overwhelming irreducible QCD background, had a cross section too small to be viable in the SM, but a much larger rate was foreseen in BSM theories such as the MSSM for large* $\tan\beta$ *values. This chapter describes this search at CMS with the full 2011 dataset at* $\sqrt{s} = 7$ TeV *[92], with a particular emphasis on the* semi-leptonic *channel [93], which exploits the presence of a muon from heavy flavour decays.*

4.1 SIGNAL CHARACTERIZATION AND ANALYSIS STRATEGY

In the Minimal Supersymmetric Standard Model, the production of the Higgs boson in association with b quarks (Fig. 4.1) can be highly enhanced with respect to the standard model case, due to the enhanced Yukawa couplings to the down-type quarks. This in particular occurs for moderately high values of the $\tan\beta$ parameter of the theory. Moreover, the Higgs boson decays with large branching ratio into a pair of b quarks, thus resulting in a characteristic signature of four b quarks in the event final state.

The expected cross section times branching ratio, within the MSSM in the m_h^{max} benchmark scenario (see Sec. 2.3), calculated at next-to-next-to-leading order (NNLO) with the BBH@NNLO [94] and the FEYNHIGGS [95–98] software respectively, as a function of m_A and for different values

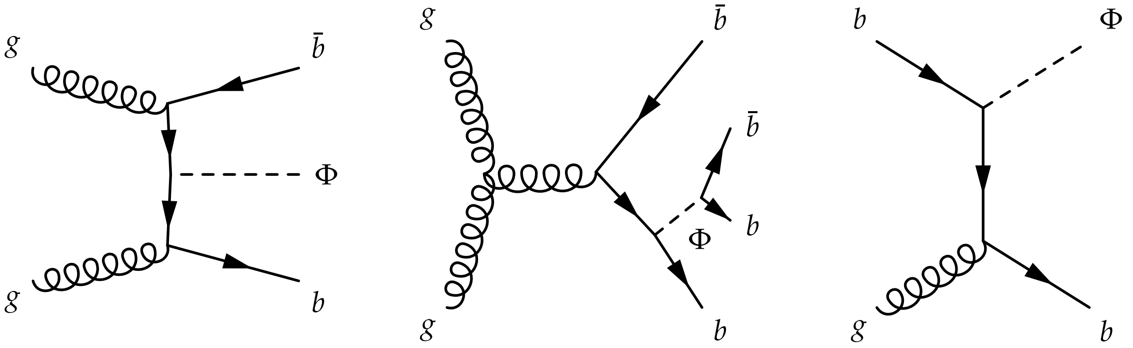

Figure 4.1: Leading-order diagrams for the production process of a Higgs boson in association with b quarks.

of tan β, are shown in Fig. 4.2. At low m_A and for low values of tan β, the $\sigma \times \mathcal{B}r$ is of the order of few tens of pb, but it increases to $\mathcal{O}(100 \text{ pb})$ for tan $\beta \gtrsim 30$. For large m_A the $\sigma \times \mathcal{B}r$ is significantly lower, due to the accessibility of other decay modes. The branching ratios of a Higgs boson into a $b\bar{b}$ are shown separately in Fig. 4.3; this decay mode is the dominant channel for a large portion of the m_Φ-tan β parameter space.

In this MSSM scenario, two of the three neutral Higgs bosons h, H and A (collectively denoted as Φ) are almost degenerate in mass. Given the expected resolution in the invariant mass $\sim 17 - 11\%$ GeV at $m_H = 120 - 250$ GeV, respectively, the reconstructed peaks of the two bosons are not distinguishable, but seen as coming from a single particle with a production cross section which is the sum of the two. This mass degeneration is taken into account throughout the analysis.

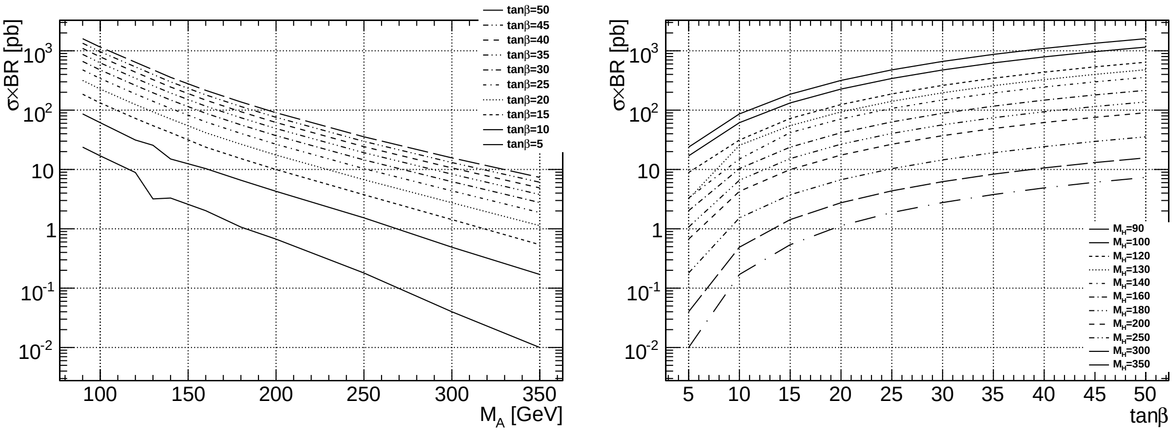

Figure 4.2: MSSM neutral Higgs production cross section times branching ratio for $b\bar{b}\Phi$ associated production at $\sqrt{s} = 7$ TeV and $\Phi \to b\bar{b}$ decay as a function of m_A (left) and tan β (right)

The p_T and η spectra of the four b quarks in simulated signal events are reported in Fig. 4.4 for a sample mass value $m_H = 120$ GeV. The dynamics of the production are such that the fourth b-quark jet often escapes detection in the forward region or at least falls outside the tracker fiducial region ($|\eta| < 2.5$) where the b quark identification algorithms are effective. As seen in Fig. 4.4, spectator b quarks have a considerably softer p_T spectrum, and consequently a lower reconstruction efficiency, while the b quarks produced in the Higgs boson decay generally have an harder p_T, and are emitted in the central region of the detector. These peculiarities allow to solve the possible combinations of jets used to reconstruct the Higgs boson candidate. The choice of selecting as Φ decay products the two highest-p_T b jets in the event is correct in

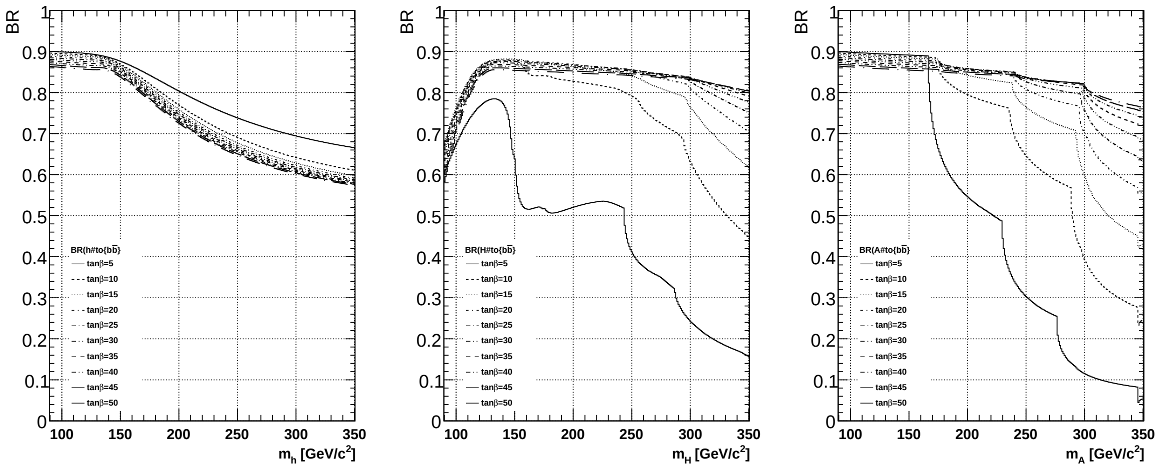

Figure 4.3: Branching ratio of neutral MSSM Higgs into $\mathrm{b\bar{b}}$, as a function of the Higgs boson mass and for various values of $\tan\beta$ for h (left), H (center) and A (right).

about 80% of the simulated events for a $m_\Phi = 120\,\mathrm{GeV}$, and increases with the boson mass. Furthermore, in about 67% of those events both of the jets are reconstructed with a p_T larger than 30 GeV.

Since the jet p_T spectra are rapidly falling above the 30 GeV threshold, a critical issue of the analysis is to maintain an acceptable trigger efficiency with a reasonably low p_T selections on jets. Another peculiarity of the signal is that the presence of four b quarks in the event implies a relatively high probability that the event contains one or more muons from the semileptonic b quark decays. The strategy of the semi-leptonic analysis is to exploit the presence of at least one muon in the event, requiring it at trigger level to reduce the p_T thresholds on jets maintaining an acceptable event rate. As a consequence of the different kinematic regimes between jets produced from massive objects and those from associated partons, the p_T spectrum of the muons emitted by b quarks from $\Phi \to \mathrm{b\bar{b}}$ is harder with respect to the others (Fig. 4.5). For this reason, the muon is required to be reconstructed inside one of the two jets selected to reconstruct the Higgs boson. For signal events with $m_\Phi = 120$ GeV, at least 20% of them have a reconstructed muon with $p_\mathrm{T} \geq 5\,\mathrm{GeV}$. This fraction is lower increasing the transverse momentum threshold, being about 7.0%, 4.5% and 3.5% for $p_\mathrm{T} \geq 12, 15$ and 17 GeV, respectively.

A total of 11 samples are produced with the PYTHIA generator, one for each Higgs mass hypothesis ($m_\Phi = 90, 100, 120, 130, 140, 160, 180, 200, 250, 300, 350\,\mathrm{GeV}$). The simulated production processes are the $gg \to \mathrm{b\bar{b}} + \Phi$, with two gluons in the initial state and $q\bar{q} \to \mathrm{b\bar{b}} + \Phi$ with two quark in the initial state. In both cases, the Higgs boson is forced to decay into b quarks.

The main background of the $\mathrm{b\bar{b}\Phi b\bar{b}}$ analysis are multijet (QCD) processes, that have a cross section several orders of magnitude larger than signal. Background events has to have at least three, high-p_T hadronic jets, originating from real or mistagged b quarks. Such events represent an irreducible background, which differs from the signal only for the dijet invariant mass (resonating only in signal events) and different distributions in some kinematic variables, such as the p_T spectra of the jets. An accurate prediction of the QCD yield and shape is required to preserve the sensitivity of the analysis, but simulation is not reliable for a quantitative background determination because of the uncertainties and imperfect modeling of multiple b jets production in gluon splitting and flavour excitation processes. For these reasons, two background prediction techniques using data (called *data-driven*) have been developed, while the simulated multijet sample is used only to validate the techniques in closure test.

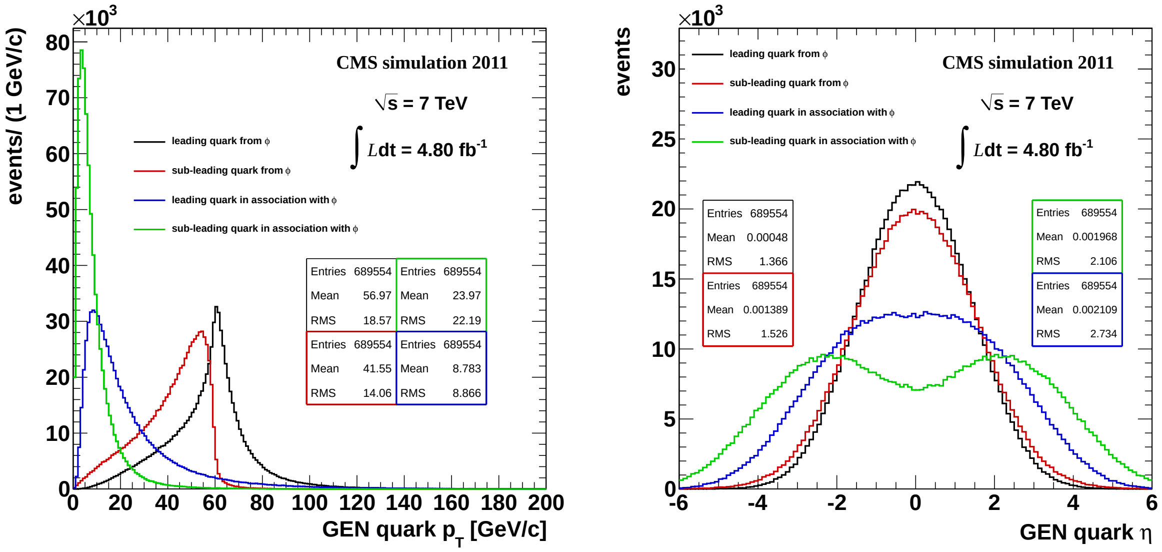

Figure 4.4: p_T (left) and η (right) generator level distributions of b-quarks in simulated b$\bar{\mathrm{b}}\Phi$b$\bar{\mathrm{b}}$ signal events. Black and red histograms are for the leading and sub-leading p_T b quarks from the Higgs boson decay, green and blue for the associated ones.

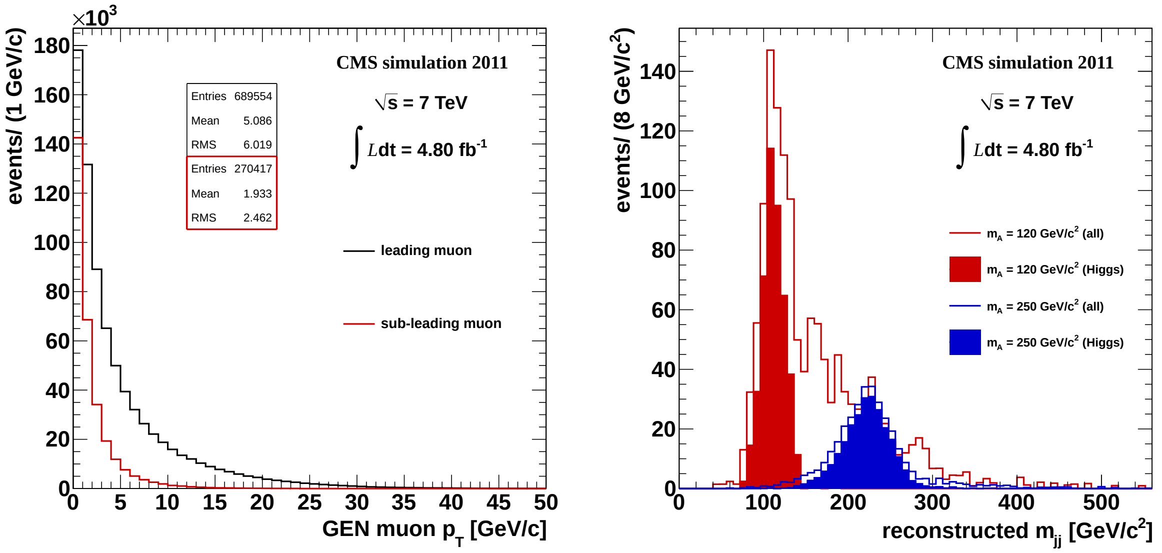

Figure 4.5: Left: transverse momentum spectra for the leading and sub-leading muons in the event. Right: reconstructed dijet invariant mass for Higgs bosons with $m_\Phi = 120$ (red) and 200 GeV/c^2 (blue). The filled histograms represent the invariant mass when the selected jets are originated from the b quarks from the Higgs boson decay. The number of events are normalized to 4.8 fb^{-1}, and calculating h or A/H cross section for $\tan\beta = 30$.

Other electroweak background processes that can mimic the signal signature are top quark pair production, which has a rather large cross section and two real, high-p_T b jets in the final state, and vector boson production with decays in at least two b quarks and additional jets, such as the $Z \to b\bar{b}$ or $VZ \to b\bar{b} + X$. However, the cross sections of electroweak backgrounds are generally small compared to the multijet production, which represents by far the most important background.

Signal and background processes are simulated with different event Monte Carlo (MC) generators, PYTHIA for signal, diboson and QCD samples and MADGRAPH for the $t\bar{t}$ and single boson with jets events. For the multijet sample, a filter is applied, and only those events containing at least one generated muon with $p_T \geq 15\,\text{GeV}$ are retained. Hadronization and fragmentation are simulated with PYTHIA, and resulting particles undergo a full simulation of the interactions with the detector. The pileup distribution of the simulated events is found to be different than the one measured in data during the 2011 data taking, and simulated events are reweighted to reproduce the analyzed data. Physics objects like jets and muons are reconstructed with the algorithms described in Chap. 3.

4.2 DATA SAMPLES AND TRIGGER

The analysis is performed on a data sample collected during the 2011 data taking period. Events are selected by triggers requiring at least two high-p_T jets, measured only with calorimetric information, in the central region of the detector. In order to maintain a trigger rate of about $\sim 5 \div 10$ Hz, jet b-tagging algorithms and the presence of a non-isolated muon with moderate p_T are required, allowing to maintain acceptable jet p_T thresholds.

The trigger selections start at the L1 stage by requiring the presence of a muon candidate, which, in the second part of the data-taking, is requested to be in coincidence with jet candidates. Although the muon reduces the overall signal efficiency of the selection due to the intrinsic semileptonic branching ratio (10%), the trigger efficiency is still sizable. The selections at HLT tighten the thresholds and add new requirements to cope with the instantaneous luminosity, which increased by more than one order of magnitude (from $2 \cdot 10^{32} cm^{-2} s^{-1}$ to $6 \cdot 10^{33} cm^{-2} s^{-1}$) during the 2011 run. The number of events and the corresponding integrated luminosity collected by each trigger path is reported in Tab. 4.1.

Path	HLT (*L1*) path name	Events	L [pb^{-1}]	ε
1	HLT_Mu12_CentralJet30_BtagIP *L1_SingleMu7*	3 027 717	180.9	3.64 ± 0.19
2	HLT_Mu12_DiCentralJet30_BtagIP3D *L1_SingleMu10*	4 532 555	537.1	2.28 ± 0.11
3	HLT_Mu12_DiCentralJet20_DiBtagIP3D1stTrack *L1_SingleMu10*	2 244 550	1108.6	1.66 ± 0.07
4	HLT_Mu12_eta2p1_DiCentralJet20_DiBtagIP3D1stTrack *L1_Mu10_Eta2p1_DoubleJet_16_8*	1 237 147	652.2	1.65 ± 0.06
	HLT_Mu12_eta2p1_DiCentralJet20_DiBtagIP3D1stTrack† *L1_Mu10_Eta2p1_DoubleJet_16_8*†	5 690 304	2326.8	
All		16 732 273	4805.7	

Table 4.1: HLT paths, corresponding data sample collected and estimated signal efficiency for a Higgs boson with mass $m_\Phi = 120\,\text{GeV}$. The last HLT path, marked with †, used the updated jet energy scale corrections at L1. The last column reports the estimated signal efficiency at trigger level for a single mass point with $m_\Phi = 120\,\text{GeV}$.

All the trigger paths require a reconstructed global muon with $p_T > 12\,\text{GeV}$, and one jet (*Path 1*) or two jets (*Path 2* onwards) with $p_T > 30\,\text{GeV}$ and $|\eta| < 2.6$. Jet b tagging, based on the TCHE algorithm described in Sec. 3.2.8 using 2D (*Path 1*) or 3D (*Path 2* onwards) impact parameter information, is applied at HLT. One b-tagged jet is required in *Path 1* and *Path 2*, and two in *Path 3* and *Path 4*.

The trigger efficiency for the different paths, with respect to the final event selection, is determined on data using different triggers. The total efficiency $\varepsilon_{\mu+j}$ of the combined muon+jets paths can be parametrized as $\varepsilon_{\mu+j} = \epsilon_{\mu} \cdot \epsilon_{j}$, where ϵ_{μ} is the HLT single-muon efficiency, and ϵ_{j} is efficiency for events satisfying the hadronic requirements to pass the final selection criteria, accounting for the reconstruction and jet b tagging at HLT. The factorization in two distinct components is possible due to the higher threshold on the muon p_{T} requested in the offline selection, where the turn-on efficiency curve has already reached the plateau level. Both ϵ_{μ} and ϵ_{j} are estimated with respect to an additional auxiliary prescaled trigger, *HLT_Mu12*. The ϵ_{μ} is evaluated by using prescaled single-muon triggers with p_{T} thresholds lower that 12 GeV, while ϵ_{j} is defined as the ratio of the number of events passing the various HLT paths of the analysis (with muon, jets and eventually b tagging) and the *HLT_Mu12* requirements, over the total number of events triggered by the *HLT_Mu12* path.

4.3 EVENT SELECTION

In order to reject the largest part of the backgrounds, only events satisfying the following requirements are retained:

- at least one good primary vertex
- a reconstructed muon with $p_{\mathrm{T}} \geq 15\,\mathrm{GeV}$, passing the *tight* quality cuts, with no explicit isolation requirements
- at least three PF jets, reconstructed with the anti-k_T algorithm within $|\eta| < 2.6$ and $E_{\mathrm{T}} > 30\,\mathrm{GeV}$ (for the two leading jets) and $E_{\mathrm{T}} > 20\,\mathrm{GeV}$ (for the others) after the jet energy corrections
- the angular separation between any pair of jets i, j must be $\Delta R_{ij} > 1$ to prevent ambiguities in b-tagging algorithms and to remove contributions from gluon splitting in background processes
- the two leading jets must have a CSV discriminator > 0.8
- the muon has to be used to reconstruct one of the two leading jets
- the third jet, ordered by p_{T}, has to have a CSV discriminator > 0.7.

These requirements select a sample of events containing three b-tagged jets (bbb), which is an optimal choice for the signal search, and called *signal region*. However, events samples with only one (bjj) or two b tags (bbj) have a sufficiently low signal purity to be considered safe from signal contamination and are used as the basis for the background modeling studies.

The data reduction flow after each selection is reported in Tab. 4.3 for data and for the different signal and backgrounds Monte Carlo simulated processes. Each contribution is normalized to the integrated luminosity, taking into account the estimated trigger efficiency. The third b-tagging requirement suppresses all the electroweak contributions to a negligible size, except those involving true b quarks in the final state such as $\mathrm{t\bar{t}}$ and $\mathrm{Z} \to \mathrm{b\bar{b}}$. In the latter, generated with PYTHIA, the jet multiplicity is rescaled to the one predicted by the MADGRAPH $\mathrm{Z} \to \mu^{+}\mu^{-}$ sample, which takes into account the contribution from higher-order diagrams involving more partons in the initial state. The difference between the two samples is almost a factor of two, with a large systematic uncertainty. The expected contribution from electroweak processes is negligible or relatively small with respect to the multijet background, and for this reason they do not receive any particular treatment.

Since data is collected using several trigger paths with different online requirements, the event kinematic properties, such as jet p_{T} spectra, are biased by the trigger selection. Due to the poor reliability of the trigger response in QCD Monte Carlo events, this effect could require the trigger paths to be analyzed independently. However, even if this is true after HLT, as a consequence of the offline selections the kinematic distributions of all the four trigger paths are comparable and do not show any particular discrepancy. As a conclusion, the data can be analyzed together, regardless of the trigger path.

4.3.1 CONTROL REGION DEFINITION

The problem represented by the simulated multijet background is that neither normalization nor kinematical distributions are able to reproduce accurately the data. Different *data-driven* techniques to estimate this background have to be finalized to predict the shape and yield. These procedures require a data sample with a negligible signal contamination, to be used as a starting base for signal region predictions. Since there is no single variable able to identify a signal-free region, multivariate analysis techniques allow to use several weakly-discriminant variables to obtain a reasonable separation between signal and background. This discriminator is not used to test the signal presence, but to select a signal-depleted region with one, two or three b-tagged jets where background studies can be performed. A likelihood discriminator $\mathcal{D}$, defined as:

$$\mathcal{D} = \frac{\prod_i S_i(x_i)}{\prod_i S_i(x_i) + \prod_i B_i(x_i)}$$

is built starting from the simulated signal and background normalized distributions S_i and B_i for the kinematic variables x_i listed in Tab. 4.2 and shown in Fig. 4.6. The choice of the best discrimination variables, and the variables themselves, depend on the Higgs boson mass. The considered Higgs boson mass points are divided into two regions, the *low mass* for $m_\Phi < 200\,\mathrm{GeV}$ and the *high mass* $m_\Phi \geq 200\,\mathrm{GeV}$, based on the $m_\Phi = 120,\ 130,\ 140\,\mathrm{GeV}$ and $m_\Phi = 250,\ 300,\ 350\,\mathrm{GeV}$. The likelihood discriminator distributions for signal and background, normalized to unit area, are shown in Fig. 4.7. The control region is defined as the sample of the events having a $\mathcal{D} < 0.4$ value.

Variable	Description
p_{Ti}	jet transverse momenta ($i = 1,2,3$)
$\Delta\varphi_{ij}$	azimuthal angle separation between jet pairs ($i,j = 1,2,3$)
$\Delta\eta_{ij}^{max,min}$	minimum and maximum η difference between jet pairs
ΔR_{ij}^{min}	minimum distance in the (η, φ) plane between jet pairs
$\Delta\eta_{12}$	pseudo-rapidity difference between the 1^{st} and 2^{nd} b jets
ΔR_{12}	distance in the (η, φ) plane between the 1^{st} and 2^{nd} b jets
$\Delta\varphi_{\Phi 3}$	azimuthal angle between the Φ candidate and the 3^{rd} jet
$\Delta\eta_{\Phi 3}$	η separation between the Φ candidate and the 3^{rd} jet
$\alpha_{\Phi 3}$	angle between the Φ candidate and the 3^{rd} jet
$\cos\theta^*$	cosine of the angle between Φ and one of the jets from the Higgs decay, in the Φ rest frame
N_{jets}	number of jets in the event with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$

Table 4.2: List of variables used to compute the discriminator; indices run on the number of jets $i, j = 1,2,3$, when jet 1 and 2 are used to reconstruct the Φ candidate.

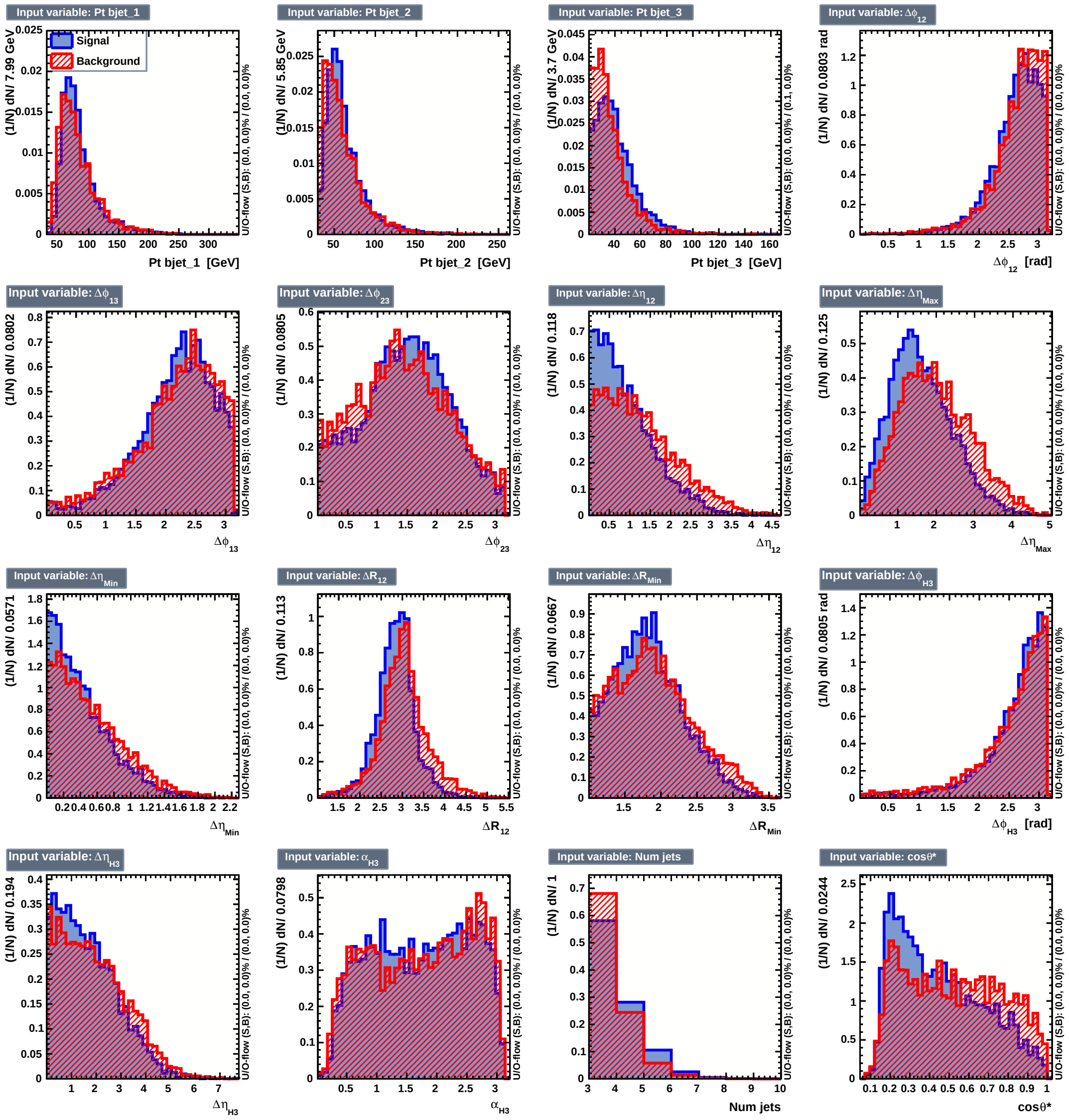

Figure 4.6: Normalized distributions for the kinematical variables used in the discriminator definition for $m_\Phi \leq 200\,\mathrm{GeV}$, for multijet simulation (red) and $m_\Phi = 120, 130, 140\,\mathrm{GeV}$ signal samples (blue).

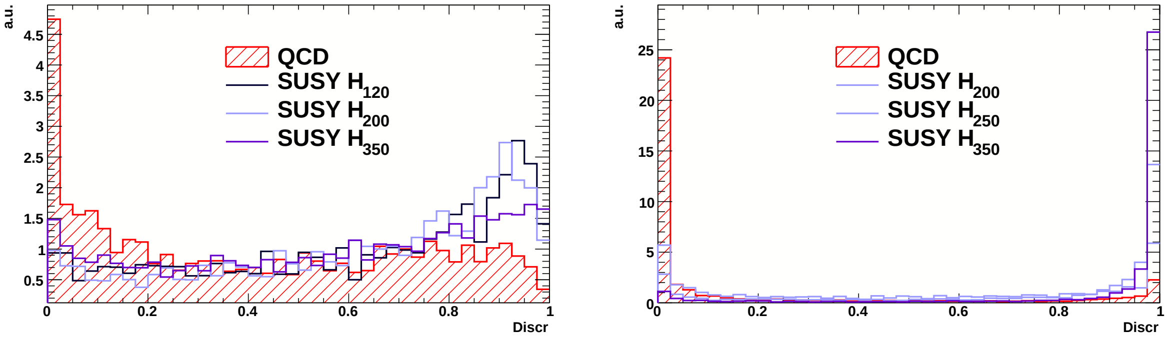

Figure 4.7: Likelihood output for the low mass region $m_\Phi \leq 200\,\mathrm{GeV}$ (left) and high mass region $m_\Phi \geq 200\,\mathrm{GeV}$ (right).

	Data						Background simulation					
Cut	*Path 1*	*Path 2*	*Path 3*	*Path 4*	*Path 4*†	All paths	QCD	$\mathrm{t\bar{t}}$	$\mathrm{Z \to b\bar{b}}$	ZZ	WZ	WW
σ (pb)	-						$2.97 \cdot 10^8$	152	3270	4.287	10.47	27.83
ϵ_{filter}	-						$2.86 \cdot 10^{-4}$	1.	0.695	1.	1.	1.
Trigger	3027717	4532555	2244550	1237147	5690304	16732273	406884036	730360	10920083	20599	50308	133723
$p_\mathrm{T}^\mu > 15\,\mathrm{GeV}$	1757902	2678935	1337394	742231	3222677	9739139	362703147	193972	212716	1848	6448	23448
$N_{jets} \geq 3$	665962	1245655	639616	404082	1556012	4511327	32221594	104007	24643	251	655	1855
$\Delta R_{ij} \geq 1$	513981	957884	498996	315284	1219439	3505584	26817799	77939	20890	191	495	1371
$CSV_1 > 0.8$	242982	492734	297838	184075	714506	1932135	7318818	29146	7872	35	54	54
$CSV_2 > 0.8$	52345	112428	162029	99175	387708	813685	1189598	8859	2238	8	8	3
μ in jet 1 or 2	50708	108551	156147	95760	374774	785940	1165417	3878	2114	6	4	1
$CSV_3 > 0.7$	3245	7323	12796	7623	29208	60195	72043	303	284	1.84	0.44	0.03

	Signal simulation										
m_Φ (GeV)	90	100	120	130	140	160	180	200	250	300	350
Trigger	3181487	2282039	1225179	942164	664435	425194	271531	178458	69624	29935	13310
$p_\mathrm{T}^\mu > 15\,\mathrm{GeV}$	79166	71272	53760	47703	37522	29513	21972	16306	8040	4014	1999
$N_{jets} \geq 3$	13416	13012	11321	10649	9028	8009	6546	5290	3012	1686	901
$\Delta R_{ij} \geq 1$	11709	11423	9796	9161	7689	6729	5386	4267	2381	1303	685
$CSV_1 > 0.8$	7452	7339	6416	5970	5133	4440	3545	2834	1584	850	441
$CSV_2 > 0.8$	4202	4078	3752	3405	3003	2662	2172	1747	986	531	271
μ in jet 1 or 2	3617	3508	3222	2937	2611	2321	1873	1509	860	465	239
$CSV_3 > 0.7$	1338	1324	1305	1160	1079	953	744	591	340	181	92
ε (%)	0.042	0.058	0.107	0.123	0.162	0.224	0.274	0.331	0.488	0.605	0.691

Table 4.3: Top: event number after each selection cut, for the various trigger paths in data, and the expected yield from simulation. For simulated backgrounds, the calculated cross section σ and the generation filter efficiency ϵ_{filter} are reported as well. The $\mathrm{Z \to b\bar{b}}$ sample needs to be rescaled by a factor ≈ 1.9 due to the correct jet multiplicity. Bottom: expected signal events after each selection cut and final signal efficiency, normalized to the resulting cross section for $\tan\beta = 30$. For all simulated samples, the trigger efficiency is included only starting from the cut $N_{jets} \geq 3$.

4.4 BACKGROUND DETERMINATION WITH B-TAGGING MATRICES METHOD

The idea behind the b-tagging matrices method is that every kinematic distributions of the multijet background in the three b-tag sample bbb can be predicted from data by using a larger and less pure two b-tag sample bbj, defined in Sec. 4.3. For any observable x, the prediction for its distribution $\mathcal{F}(x|bbb)$ in the three b jets sample can be estimated as:

$$\mathcal{F}(x|bbb) = \mathcal{F}(x|bbj) \times \mathcal{P}^{j3}_{b-tag}(j)$$

where $\mathcal{P}^{j3}_{b-tag}(j)$ is the probability of the third jet to be b tagged, for a certain b-tagging algorithm and working point, and $\mathcal{F}(x|bbj)$ is the observed distribution for x in the bbj sample.

The probability can be expanded, separating the efficiencies of the b-, c-, and light-jets, attributed to u, d, s partons or gluons can collectively denoted as l:

$$\mathcal{P}^{j3}_{b-tag}(j) = \varepsilon_b \cdot f_b + \varepsilon_c \cdot f_c + \varepsilon_l \cdot f_l$$

The f_b , f_c and f_l are the fractions of b, c and light quarks that originate the third jet in the bbj sample, respectively.

The b-tagging efficiency for jets originating form b or c, light quarks is extracted from simulated events. The efficiency is defined as the ratio between the number of jets coming from a given quark flavour, passing preselection requirements and the b-tag threshold of CSV$>$ 0.7, and the number of all jets of that specific flavour. The efficiency is parametrized with respect to the jet E_{T}, $|\eta|$ and separately for jets with an high and low number of charged tracks (≤ 10 and ≥ 10). The difference between b-tagging efficiency in data and Monte Carlo is considered by multiplying the efficiency in simulation by a scale factor, obtained from data measurements in independent samples [99]. For the working point CSVM (CSV$>$ 0.679), close to the CSV$>$ 0.7 cutoff adopted by the analysis, the scale factor is 0.95 ± 0.03.

The efficiency and flavour fractions depend on the third jet and event kinematics. The best solution, determined with a MC study, is to use a multiple parametrization for the heavy-flavour fraction:

$$f^{j3}_{b,c,l} = f\left(E^{j3}_{\mathrm{T}}, |\eta^{j3}|\right) \times f\left(\Delta R_{12}, \Delta R_{H3}\right)$$

where the first factor uses third jet's kinematic variables (transverse energy and pseudorapidity). The second factor, based on topological variables (the angular separation of the two leading jets ΔR_{12} and between the Higgs boson candidate and the third jet ΔR_{H3}), is used only for the shape of the distribution and it is normalized to unity.

The third jet flavour fractions are determined directly from data, using bbj events, kinematically as close as possible to the final selection. In order to avoid biases due to the presence of signal, the background prediction is performed without looking at data in the signal region (*blind analysis*). As a consequence, the flavour fractions are extracted from the control region, defined in Sec. 4.3.1.

4.4.1 MEASUREMENT OF THE HEAVY FLAVOUR FRACTION OF THE THIRD JET

The flavour content $f_{b,c,l}$ of the third most energetic jet in the control region bbj events is estimated by looking at two jet-related variables. The first is the *Tag Mass*, which is the invariant mass of all the tracks associated to the secondary vertex of the jet, if one secondary vertex is reconstructed; the second is the output of the *JetBProbability* b-tagging algorithm, described in Sec. 3.2.8. Being a track-based discriminator, it can be used even if no secondary vertex is present.

The *bbj* sample is divided in different intervals of the four variables used for the parametrization of the third jet flavour content (Sec. 4.4). The intervals are chosen to have a good estimation of the flavour fractions in the whole phase space, but retaining enough data statistics to perform the fit. For each interval, a likelihood fit of the two variables is performed on data to obtain the fractions, using templates for the Tag Mass and JetBProbability distributions of b, c and light-partons obtained from Monte Carlo simulated events. In order to improve the available statistics in simulation, the b-tag requirement on the first two leading jets is relaxed, having care to check that no bias is introduced.

The templates are used to fit data using a two-dimensional likelihood in the *Tag Mass* and *JetBProbability* variables if the first is defined, and one-dimensional in the *JetBProbability* otherwise. The fit is constrained by the $f_b + f_c + f_l = 1$ condition. In order to achieve the correct treatment of parameter errors near the boundaries, the fractions f_b, f_c, f_l are not directly fitted, but they are replaced with the following parameter set:

$$\begin{aligned} \xi_1 &= \frac{1}{2}(1 + \sin f_b) \\ \xi_2 &= \frac{1}{2}(1 - \xi_1)(1 + \sin f_c) \\ \xi_3 &= 1 - \xi_1 - \xi_2 \end{aligned}$$

The likelihood function to be maximized is:

$$\mathcal{L}^{(2D)} = \prod_i \mathcal{P}_{TagMass}(n_i, \mu_i) \cdot \mathcal{P}_{JetBP}(n_i, \mu_i)$$

$$\mathcal{P}(n_i, \mu_i) = \frac{e^{\mu_i}(\mu_i)^{n_i}}{n_i!}$$

where n_i and μ_i are respectively the number of events and the expected events in each bin; μ_i depends on the number n_i^b, n_i^c, n_i^l of events in the i-th bin of the three templates:

$$\mu_i = \left[\xi_1 n_i^b + \xi_2 n_i^c + \xi_3 n_i^l\right] \cdot \sum_i^{bins} n_i$$

Once the likelihood is maximized, the result values of ξ_1, ξ_2, ξ_3 and their errors are returned, from which the final f_b, f_c, f_l values can be easily extracted. The relative fractions are summed weighted with the number of events:

$$f_{b,c,l} = \frac{f_{b,c,l}^{2D} \cdot N^{2D} + f_{b,c,l}^{1D} \cdot N^{1D}}{N^{2D} + N^{1D}}$$

where the 2D and 1D superscripts refer to the events fitted with the two and one dimensional likelihood. As an example, the template fits to the data for a given considered interval of the third jet variables are shown in Fig. 4.8. The *Tag Mass* and *JetBProbability* variables are correlated, but this does not represent an issue since the fit errors are not used in the analysis. As a closure test of the procedure, the results for the fitted flavour fractions in the simulated data in the control region are compared in Fig. 4.9 with the true flavour fractions in simulation. The comparison with the Monte Carlo true values shows a small bias of the order of 6% averaged on the whole parameter space. As it will be shown in the next section, this has a negligible effect on the final prediction of the background shape and normalization. The results of the extracted fractions in data are shown in Fig. 4.10, divided between the samples collected with single and double online b tag trigger paths; as expected, f_b is higher in the latter.

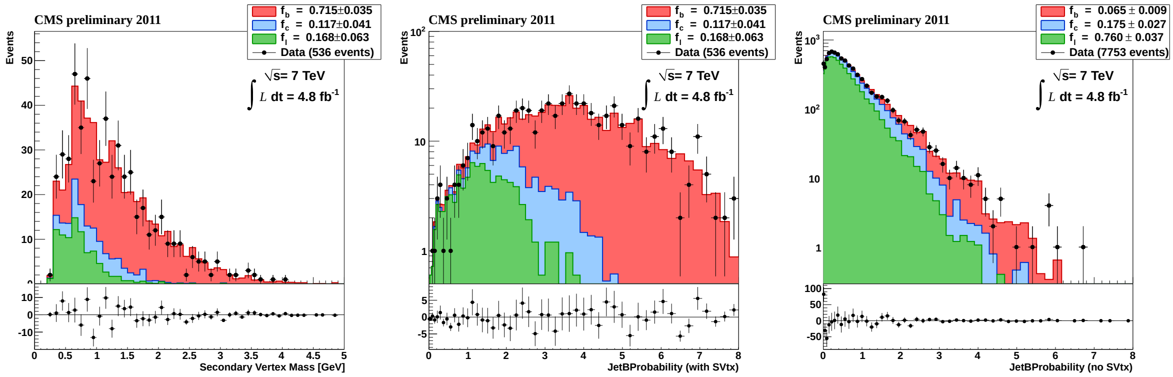

Figure 4.8: Example of a fit for one bin of the E_{T}, $|\eta|$ parametrization. The red, green and blue histograms are respectively the templates for b, c and light quarks. Black dots represents data. Left and center: two-dimensional fit to *Tag Mass* and *JetBProbability*. Right: fit to *JetBProbability* only when no secondary vertex is reconstructed.

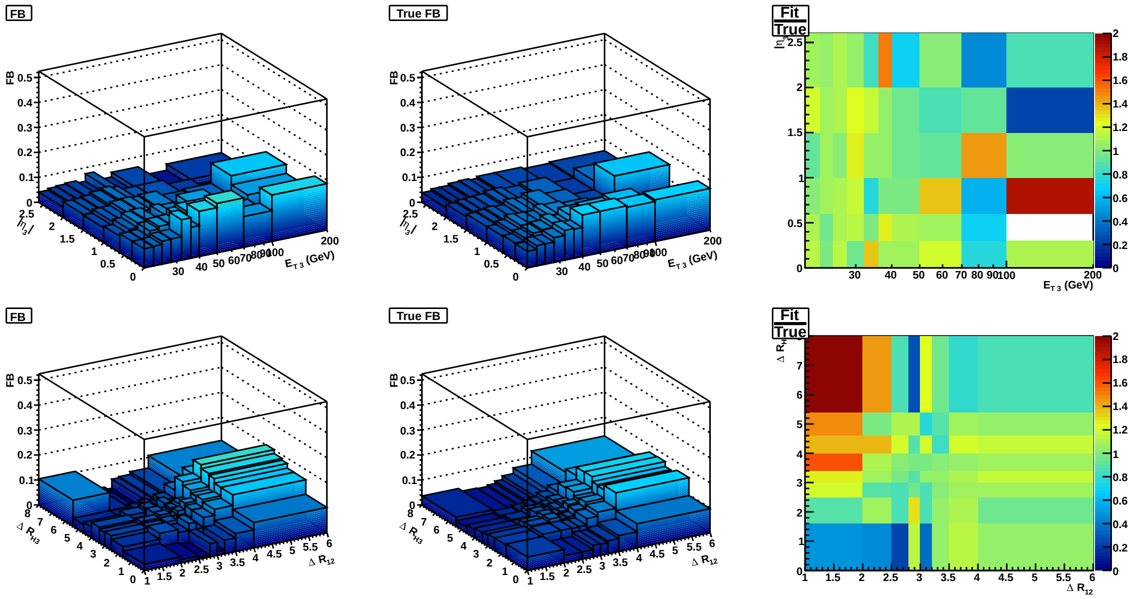

Figure 4.9: Comparison between fit results and QCD Monte Carlo truth. Top is for the E_{T}, η parametrization, bottom the ΔR_{12}, ΔR_{H3}) parametrization. Left: f_b as resulting from the likelihood fit, center: f_b from MC truth, right: ratio between fitted and truth f_b.

4.4.2 PREDICTION FROM TWO TO THREE B-TAGGED JETS

By using the b-tagging efficiencies and flavour fractions, relevant kinematic variables can be predicted in the bbb sample starting from the two-tagged bbj sample, both in the control region

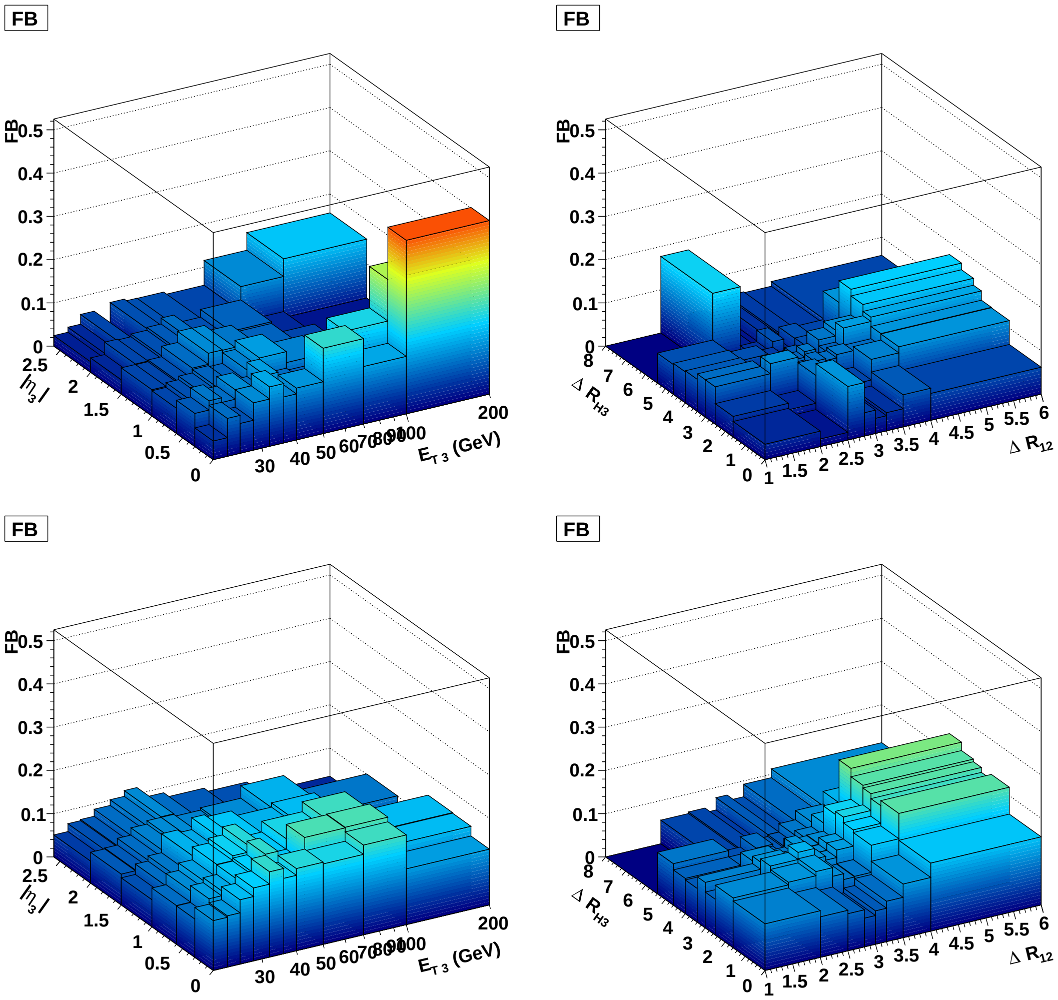

Figure 4.10: Fitted b-flavour fraction f_b on data, shown separately for trigger paths requiring at least one (top) or two (bottom) b-tagged jets in the online selections, and for the p_T, η parametrization (left) and ΔR_{12}, ΔR_{H3} parametrization (right).

and in the signal region. The predicted and observed distribution of some key variables, such as the invariant mass of the two leading jets, are shown in Fig. 4.11. As a reference, the Higgs signal template for $m_\Phi = 120\,\text{GeV}$ or $m_\Phi = 250\,\text{GeV}$, are superimposed and normalized to the cross section obtained with $\tan\beta = 30$. In the Monte Carlo control and signal region the shape and normalization are in good agreement, showing a full closure of the technique. Also the normalization shows a nice agreement within the statistical uncertainties of the simulated sample available, which however has almost a factor ten less statistics than the real data. It is worth stressing that the signal region is a completely separate sample with respect to the one where the b-tagging efficiencies and flavour fractions are obtained. For what concerns the data, the shape agreement is also good in the control region. However, the prediction of the overall normalization factor shows an overestimation of about 10%; this disagreement, and the corresponding uncertainty, is used as a correction factor for the signal region in data, as described in Sec. 4.6.

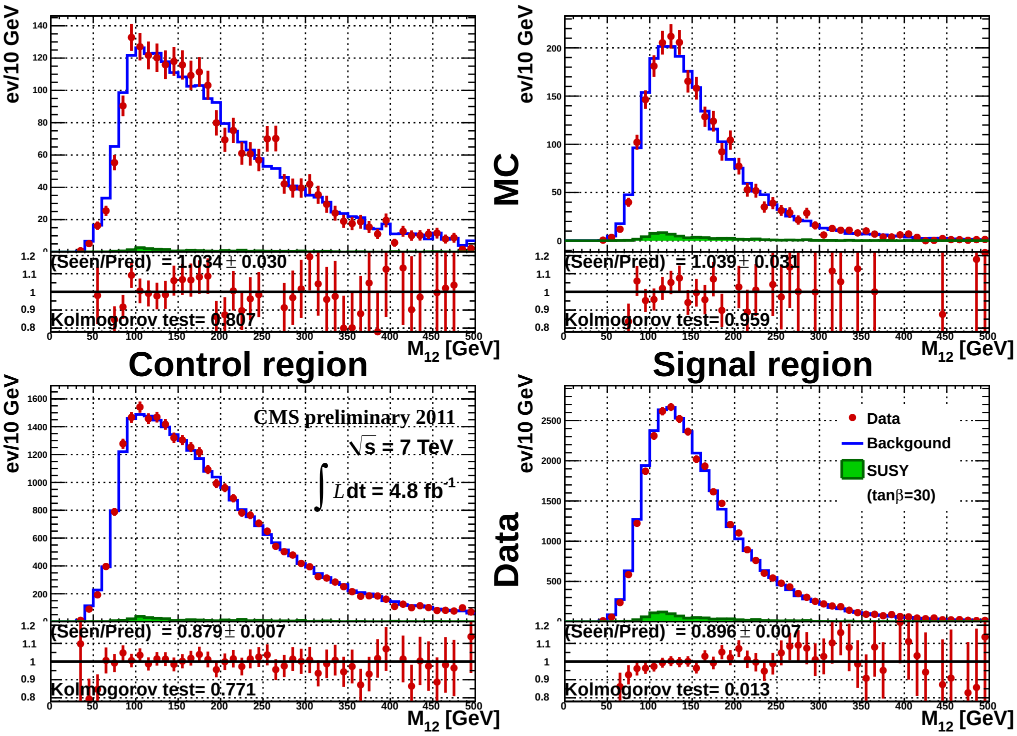

Figure 4.11: Invariant mass of the first two jets (M_{12}), for simulation (top) and data (bottom) and for control (left) and signal regions (right), showing the background prediction (blue) and measured data or simulation (red) for low mass range ($m_{\Phi} < 200\,\text{GeV}$). The signal corresponds $m_{\Phi} = 120\,\text{GeV}$ and $\tan\beta = 30$.

4.5 THE BACKGROUND SHAPE PREDICTION WITH THE HYPERBALL ALGORITHM

An estimation of shape and normalization of the background is possible with an alternative and independent method to the one presented in Sec. 4.4. This method, known as the *Hyperball* method, is also data-driven. The algorithm assigns a weight to each events, corresponding to the probability that the three jets in the event are b tagged. The weight is calculated as the number of *bbb* events divided by the number of *bjj* events contained in the volume (an *hyper-ellipsoid*) of the multi-dimensional space of the kinematic variables centered on the coordinates of the event itself.

The prediction of the relevant kinematic variables that are used for the signal search is performed starting from a sample of events selected using the criteria listed in Sec. 4.3, except for b tagging that here is applied only on the leading jet. In order to keep the Hyperball algorithm results completely orthogonal from those of the b-tagging matrix method described in the previous section, events where the two leading jets are both b tagged are excluded. The criterion used to identify the sample of similar events to those in the signal region from the control region sample is based on the multi-dimensional distance defined by the following expression:

$$D^2 = \sum_{i=1}^{n_V} w_i^2 \left(x_i - y_{\alpha i}\right)^2$$

where the x_i are the variables that characterize the signal region events, the $y_{\alpha i}$ are the corresponding variables for the α event in the control region. The range of each variable x_i is divided into ten intervals, with different sizes chosen to have all intervals equally populated; for each interval the fraction of three b-tagged events, f_i, is computed. For any value of the variable x_i, the weight w_i is calculated by choosing the two intervals with $\bar{x}_i$ averages immediately smaller ($\bar{x}_{iL}$ and bigger $\bar{x}_{iR}$ than $\bar{x}_i$, and computing the ratio of the differences between the *bbb* events fraction and the x_i average distance:

$$w_i = \frac{f_{i_R} - f_{i_L}}{\bar{x}_{i_R} - \bar{x}_{i_L}}$$

To reduce the bias introduced by assuming that the average *bbb* events fraction does correspond to the fraction at the average $\bar{x}_i$, a least-square fit of the fraction itself to a quadratic function of x_i is performed, and its value at $\bar{x}_i$ is taken. The above procedure estimates the gradient of the b-tagging probability in each of the directions of space as if it is independent on the particular coordinate.

The estimation of a b-tag probability from the ratio of the number of *bbb* and *bjj* events in the hyper-ellipsoid could result in a bias when the tested event has some variable near to its boundary value. This bias is cured by identifying which variables are near to the corresponding thresholds and performing a linear interpolation. When all the energies and pseudorapidities in the tested event are not close to the above defined thresholds, the 3 b-tags probability for the test event is computed by a weighted ratio:

$$P = \frac{\Sigma_\beta 1/D_\beta^2}{\Sigma_\alpha 1/D_\alpha^2}$$

where in the numerator and denominator the sum run over *bbb* and *bjj* events. In this way the most distant events have a smaller weight.

Applying the Hyperball method to real data, a specific procedure tunes the most important parameters of the algorithm, such as the number of training events, the number of events in the hyper-ellipsoid, number of variables, and number of intervals of the invariant mass distributions between the two leading jets. As any other multivariate approach, the Hyperball method is sensitive to the size of the training sample, in particular from the training sample, that should generally be chosen as large as possible. For this reason the *bjj* sample is chosen as a starting point. The variables used are the E_{T} and η of the three jets, the number of charged tracks of the second and third jet, the azimuthal difference between the first two jets and the angular separation between the second and third jet, the invariant mass of the two leading jets, and the p_{T} of the system of jet 1 and 2, jet 2 and 3, and the three jets combined. It is verified that by increasing the number of variables the prediction power rapidly worsens, especially the normalization of the three b-tagged sample. The comparison of the predicted and observed distribution, in control region only, are shown in Fig. 4.12, for both the low and high mass regions.

4.6 SYSTEMATIC UNCERTAINTIES

The background prediction is affected by uncertainties that limit the precision with which the signal and background are known. Two main categories of systematic uncertainties are identified, affecting the signal and background normalization or shape distributions.

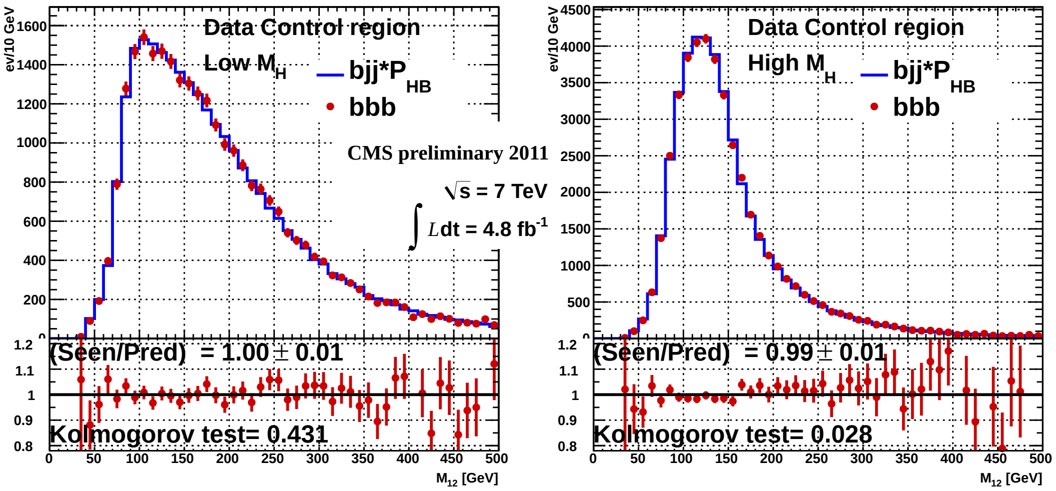

Figure 4.12: Invariant mass of the two leading jets as predicted by the Hyperball method (blue line) compared to the observed data distribution (red dots) for events in the control region. The results are shown for both low (left) and high mass range (right). The predicted and observed distributions are normalized to the same area, and the ratio between the true events is shown on the bottom of the distributions together with the Kolmogorov test result.

4.6.1 COMBINATION OF THE TWO BACKGROUND PREDICTION METHODS AND SHAPE UNCERTAINTIES

The methods described in Sec. 4.4 and 4.5 provide two independent background templates of the Higgs boson invariant mass, B_{bm} and B_{hb}. Because of their orthogonality, the two methods can be directly compared; their agreement confirms their trustworthiness. The total normalization of the predicted background is chosen to be given by the b-tagging matrix method, and therefore the Hyperball background template is scaled accordingly. Shapes are combined by performing a weighted bin-by-bin average:

$$x_w = \frac{x_{bm}\sigma_{bm}^2 + x_{hb}\sigma_{hb}^2}{\sigma_{bm}^2 + \sigma_{hb}^2}$$

where x and σ are the number of events and their statistical error in a given bin for the two methods. The weighted average statistical error σ_w is rescaled by the χ^2, as described in Ref. [24]:

$$\sigma_w^2 = \chi^2 \left(\frac{1}{\sigma_{bm}} + \frac{1}{\sigma_{hb}} \right)^{-1} \quad , \quad \chi^2 = \frac{1}{N-1} \left(\left[\frac{x_{bm} - x_w}{\sigma x_{bm}} \right]^2 + \left[\frac{x_{hb} - x_w}{\sigma_{x_{hb}}} \right]^2 \right)$$

The difference between the unscaled and scaled statistical error is interpreted as a systematic uncertainty on the shape of the background prediction. In Fig. 4.13 are shown the predicted background distributions obtained with the two methods compared to the weighted average one, for the two leading jet invariant mass.

4.6.2 NORMALIZATION UNCERTAINTIES

Normalization uncertainties affecting only the signal samples or the background estimation are separated. Uncertainties in physics object reconstruction and energy determination, limited knowledge of the parton distribution functions or collected luminosity, affect only the

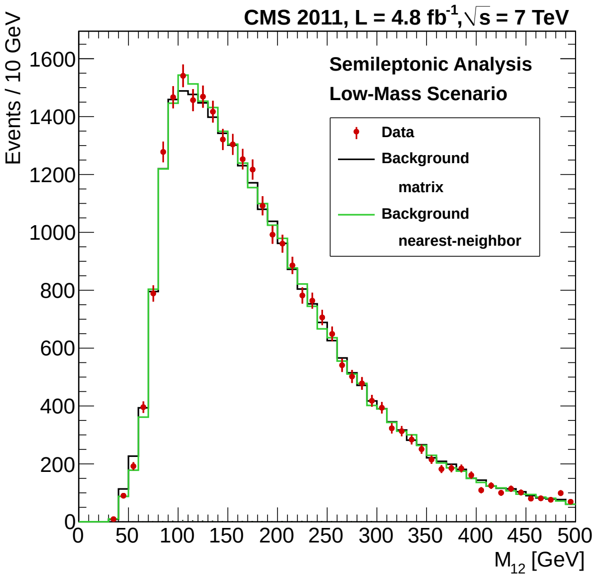

Figure 4.13: Background predictions from the b-tagging matrix and the Hyperball method compared with their weighted average for low (left) and high mass range (right). Statistical uncertainties of the weighted average in each bin are rescaled according to the method discussed in the text.

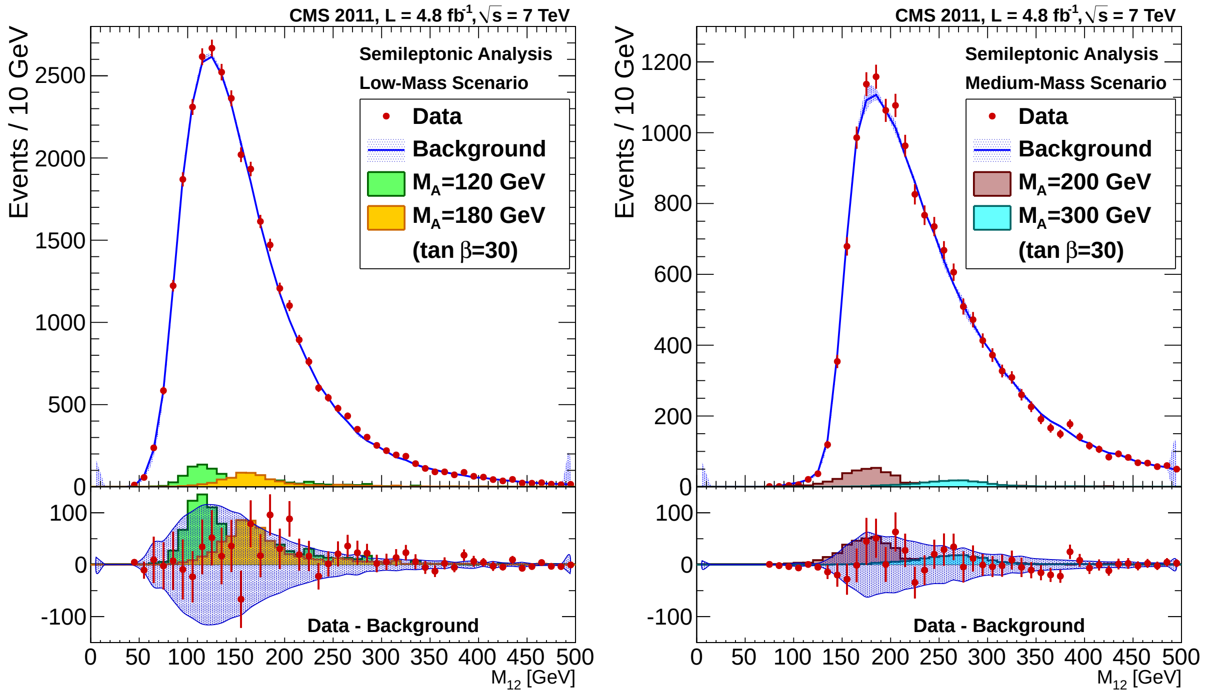

Figure 4.14: The distributions show the two leading jet invariant mass for the predicted backgrounds in the signal regions (blue line) compared to the observed data (red dots), for low (left) and high (right) Higgs mass regions. The expected signal contribution for different Higgs boson mass hypothesis is also shown, with a cross section obtained for $\tan\beta = 30$. Below the distributions, the difference between the observed data and predicted background is shown, together with the systematic uncertainties on the background.

signal, since background normalization is determined with other methods based on data. The uncertainty sources for signal are the following:

Trigger: the trigger turn-on curves are derived from data, as described in Sec. 4.2, along with their uncertainties ($\approx 5 \sim 3\%$ for $m_\Phi = 120\,\mathrm{GeV}$ and $m_\Phi = 250\,\mathrm{GeV}$) coming from the limited statistics of the prescaled triggers used as reference.

b-tagging efficiency: the uncertainty in the data over MC scale factor is extracted from a dedicated study [99], using a b-enriched sample from top quark decay, and accounts for $\approx 4\%$ per jet, totaling $\approx 12\%$ for the three jets.

Jet energy scale: the uncertainty in the jet energy corrections (see Sec. 3.2.7) is estimated by the standard procedure of scaling up and down the jet E_T in the event by the measured uncertainty [100]. The relative change in the amount of events passing the final selection is $^{+2.5}_{-3.1}\%$.

Jet energy resolution: the contribution due to the resolution of the jet energy measurement is obtained by changing randomly the momenta of each jet, according to the corresponding uncertainty for the p_T and η [100]. The procedure is repeated several times, obtaining different event yields, and the corresponding uncertainty is the observed standard deviation of the expected signal yield, which results to be $\pm 1.9\%$.

Muon momentum scale and resolution: the effects are small given the good performance of the muon measurement, and are estimated as $\approx 0.2\%$ and $\approx 0.6\%$, respectively [88].

Luminosity: the uncertainty on the measured integrated luminosity in 2011 is 2.2% [101].

The uncertainty sources that affect only the background are:

Comparison of observed and predicted background: the ratio of the predicted and observed distributions in the data control region is fitted with the lowest order polynomial compatible with the data, as shown in Fig. 4.15. Given the good agreement between the *bbb* data and the prediction from *bbj*, a simple constant provide a good fit, as tested with a Fisher test with increasing order polynomials. The fitted values are 0.877 ± 0.007 and 0.885 ± 0.006 in the low and high Higgs mass regions, respectively. These values are used as normalization correction factors in the signal regions and as systematic uncertainties of the background normalization prediction, respectively.

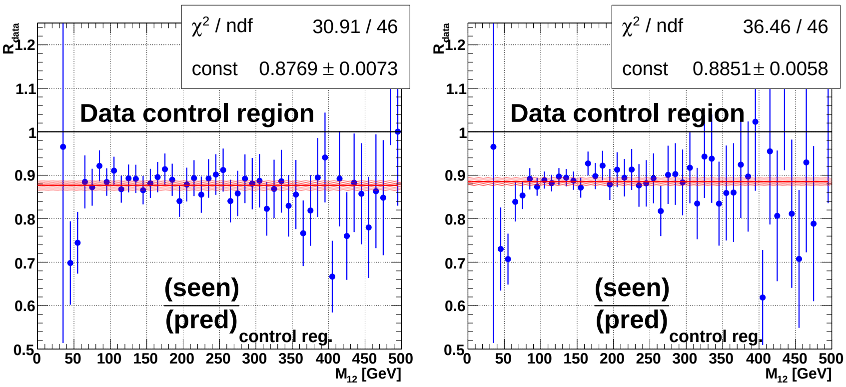

Figure 4.15: Ratio of distribution of *bbb* events (seen) and prediction from *bbj* (pred) as a function of M_{12} in the control region for data, fitted with a constant function for the low (left) and high mass (right).

Extrapolation from control to signal region: the b-tagging matrix method obtains an estimate of the signal region normalization by looking at the control region only. The correction factor is simply the ratio of the data/prediction ratios in simulation:

$$R = \frac{\left(\frac{N_{observed}}{N_{predicted}}\right)_{SR}}{\left(\frac{N_{observed}}{N_{predicted}}\right)_{CR}}$$

This correction is 1.01 ± 0.042 for low mass and 1.02 ± 0.05 for the high mass region. These results are taken as a bias for the extrapolation from control and signal region, and their errors as a systematic uncertainty in the normalization of the background prediction.

4.6.3 MODEL-DEPENDENT UNCERTAINTIES

For the computation of the limits in the m_{A}, $\tan\beta$ parameter space of the MSSM, additional model-dependent systematic uncertainties have to be taken into account:

Factorization and renormalization of QCD scale: singularities that appear in perturbative QCD calculation are removed by factorization and renormalization techniques [102]. The theoretical systematic uncertainty on these calculations for the $pp \rightarrow b + \Phi$ process is in the $6 - 28\%$ range, depending on m_{Φ}.

Underlying event and parton showering: the modeling of the underlying event and of the parton showering process are affected by a systematic uncertainty estimated to be 4%.

PDF : the effect is estimated reweighting on a per-event basis the PDF values obtained by varying up and down their parameters by the corresponding uncertainty, as suggested by the PDF4LHC working group [103]. The uncertainty is then represented by the weighted number of events passing the offline selection, resulting $^{+2.5\%}_{-2.7\%}$ for $m_{\Phi} = 120\,\mathrm{GeV}$, and $^{+4.7\%}_{-4.4\%}$ for $m_{\Phi} = 250\,\mathrm{GeV}$.

4.7 COMBINATION WITH THE ALL-HADRONIC ANALYSIS

An alternative approach in the MSSM Higgs boson search is pursued by a different analysis, targeting the same final state but with different trigger, selections and strategy, without relying on the presence of the muon in the event (henceforth called *all-hadronic*) [104]. Analogously with the *semi-leptonic* analysis, a peak in the invariant mass of the two leading jets is sought looking at the distribution of a three b-tagged jet sample.

The trigger relies on the online identification of at least two b jets, as in the semi-leptonic analysis, but dropping the muon requirement implies an higher rate and the consequent need to increase the online thresholds on the jets to keep the trigger rates manageable. Without the muon, the signal efficiency increases, but the background yield is larger as well. The efficiency of the trigger for signal events passing the offline selection is $47 - 67\%$, for a Higgs boson mass in the range of $90 - 350\,\mathrm{GeV}$. Three different categories are defined according to the trigger selections: in the first and second ones the leading jet p_{T} threshold is set to $p_{\mathrm{T}} > 46$ and $60\,\mathrm{GeV}$, and $p_{\mathrm{T}} > 38$ and $53\,\mathrm{GeV}$ for the next-to-leading jet. The third category is similar to the first but requires a third jet with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$. The triggers with lower thresholds allow for a better exploration of the low mass region (first and third categories, and $m_{\phi} < 180$ GeV), albeit with smaller integrated luminosity ($2.7\,\mathrm{fb}^{-1}$). The higher-threshold triggers (first and second categories) allow to collect an higher integrated luminosity ($4.0\,\mathrm{fb}^{-1}$), but they are sensitive only to high masses ($180 \leq m_{\phi} \leq 350$ GeV).

The offline selections require at least three reconstructed jets with $|\eta| < 2.2$, passing the p_T cuts of 46, 38 and 20 GeV in the low, and 60, 53 and 20 GeV in the medium mass scenario. Analogously with the semi-leptonic analysis, a signal-enriched triple-b jet and a double-b jet control region samples are defined after tight CSV b tagging application. A compact variable sensitive to the three-jet flavour composition is constructed on a per-vent basis depending on the value of the secondary vertex mass of the three jets. For each jet, an index B_j is set equal to zero for jets with no reconstructed secondary vertex or where the secondary vertex mass is below 1 GeV, 1 if the vertex mass between $1 - 2$ GeV, 2 if larger than 2 GeV. The three indices B_1, B_2, and B_3 are combined in an event b-tag variable X_{123}, defined as $X_{123} = X_{12} + X_3$, where $X_{12} = 0, 1$ or 2 depending on whether $B_1 + B_2 < 2, 2 \leq B_1 + B_2 < 3$ or $B_1 + B_2 \geq 3$, respectively, and $X_3 = 0$ if $B_3 < 2$, and $X_3 = 3$ otherwise. By construction, the event b-tag variable X_{123} can have six possible values ranging from 0 to 5.

Due to the unreliability of the simulated multijet background, a method based on control data samples, similar to the one used in Ref. [59] is applied. The background model is constructed from templates that are derived from the double-b tag sample. Events are divided into the following categories depending on the number and the p_T-ranking of the b-tagged jets in the event. From these categories, nine background templates are constructed by weighting each untagged jet with the b-tagging probability assuming that its true flavour corresponds to either a light parton, a c quark or a b quark. Similar templates are combined together, reducing their number to five. Each background template is a distribution in the two-dimensional space spanned by M_{12}, the dijet mass of the two leading jets, and the event b-tag variable X_{123}.

A signal template is obtained for each considered value of the Higgs boson mass by performing the full selection on the events of the corresponding simulated signal sample. The signal is extracted by simultaneously fitting the M_{12} and X_{123} variables a linear combination of signal and background templates to data in the triple-b tag sample. The normalization of background templates is a free parameters, but constrained to be positive. The M_{12} and X_{123} projections of the background templates after the fit are shown in Fig. 4.16.

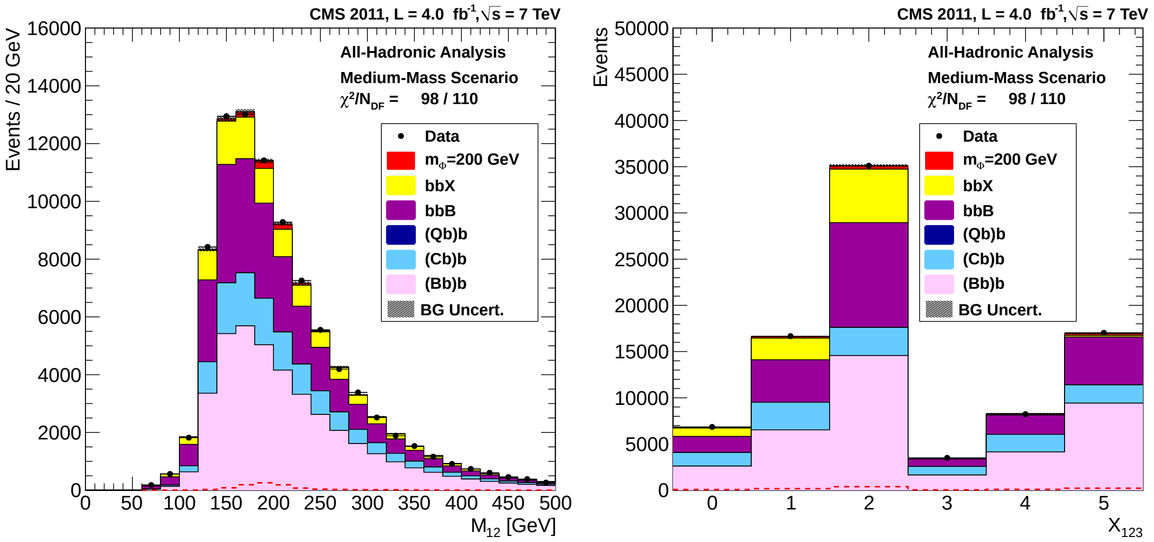

Figure 4.16: Dijet mass distribution, M_{12} (left), and the distribution of the event b-tag variable X_{123} (right) in the medium mass scenario, with the background-only fit, and including an additional signal template for a MSSM Higgs boson with a mass of 200 GeV.

4.8 SIGNAL EXTRACTION AND UPPER LIMITS

In the semi-leptonic analysis, the presence of a MSSM Higgs boson in the triple b-tagged data sample is tested with a binned likelihood fit to the data using the background and signal templates. The background shape and normalization is obtained from the data-driven methods described in Sec. 4.4 and 4.5, and the signal templates are obtained from the Monte Carlo simulation for different Higgs mass hypotheses. The most discriminating variable, used for the fit, is the invariant mass of the two leading jets M_{12}, shown in Fig. 4.14.

No significant deviation from the background expectations is observed in either analysis, and the $\mathrm{CL_s}$ [105, 106] criterion is used to combine both results and determine the 95% confidence level (CL) limit on the signal contribution in the data, using the ROOSTATS [107] package. To avoid correlations, in the *all-hadronic* analysis the events common to the *semi-leptonic* case are removed from the triple-b-tag samples. The fractions of events removed in the *all-hadronic* data samples are 2.3% and 2.7% for the low and medium mass scenarios, respectively. The requirement of a muon in the semi-leptonic analysis and the harder kinematic selections of the all-hadronic analysis are responsible for such small overlap. Common events in the simulated signal samples are also removed, although they are found to have negligible effect on the shape of the signal templates. The *all-hadronic* signature has a generally larger signal efficiency, but requires higher thresholds for jet energies, while the presence of a muon in the *semi-leptonic* signature allows for lower thresholds at the cost of lower signal efficiency. As a result, both signatures are comparable in sensitivity.

The results are shown in Fig. 4.17 in terms of the product of the total cross section and the branching fraction. There is generally good agreement between the observed and expected upper limits within statistical errors, and no indication of a signal is seen. The observed upper limits range from about 312 pb at $m_\Phi = 90\,\mathrm{GeV}$ to about 4 pb at $m_\Phi = 350\,\mathrm{GeV}$. The individual observed limits of the *all-hadronic* and *semi-leptonic* analyses are also displayed separately.

These limits can be interpreted as exclusion limits in the m_A, $\tan\beta$ parameter space of the MSSM model. The combined values of cross section times branching ratio are converted in the corresponding $\tan\beta$ value, for the m_h^{max} benchmark scenario as described in Sec. 4.1, and are shown in Fig. 4.17. The upper limits for the semi-leptonic and all-hadronic individual analyses are shown as well; the sensitivity of the two is comparable, but the semi-leptonic is slightly more stringent.

Figure 4.18 shows the results in the scenario with $\mu = -200\,\mathrm{GeV}$, together with previous limits set by the Tevatron experiments [59] in the multi-b jet final state, and by LEP [56], as described in Sec. 2.6 and 2.5. In particular, no excess over the expected SM background is found for high values of $\tan\beta$ and for a resonance in the $100 - 150\,\mathrm{GeV}$ mass range, as previously reported by CDF and D0. The result of this work, using only the data collected in 2011 at the LHC with a center-of-mass of $\sqrt{s} = 7\,\mathrm{TeV}$, extends the sensitivity for MSSM searches in the $\Phi \to \mathrm{b\bar{b}}$ decay mode to much lower values of $\tan\beta$, excluding the region where the excess was reported, and provides the most stringent limits on neutral Higgs boson decay in the $\mathrm{b\bar{b}}$ channel, produced in association with b quarks.

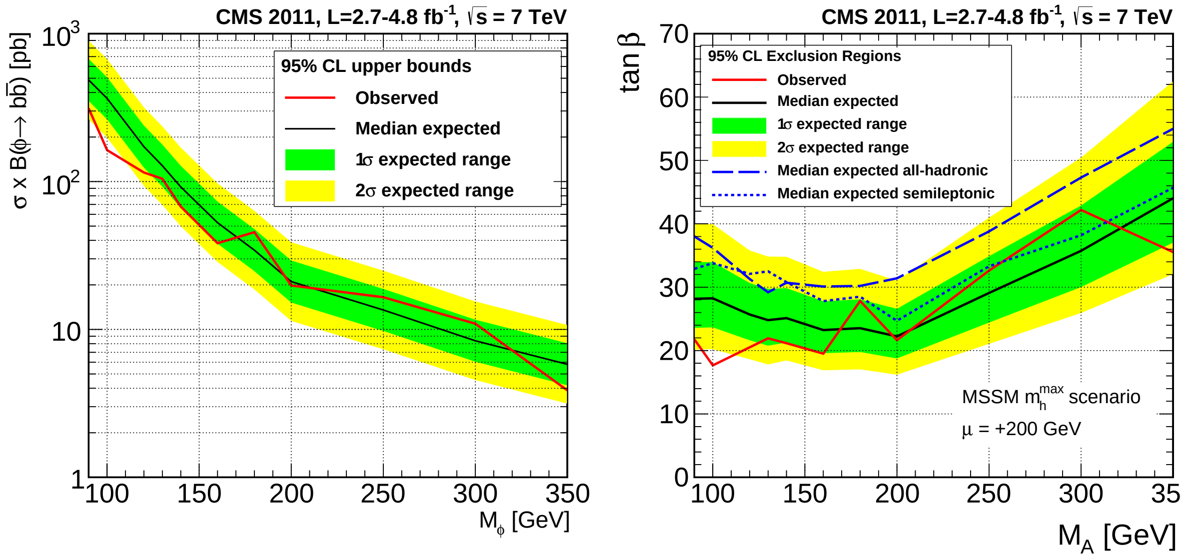

Figure 4.17: Left: observed and expected upper limits for the cross section times branching fraction at 95% CL, including statistical and systematic uncertainties for the combined all-hadronic and semi-leptonic results, in the $pp \rightarrow b\Phi$, $\Phi \rightarrow b\bar{b}$ channel, with $\Phi = h, H, A$. 68% and 95% CL ranges for the expected upper limit are also shown. Right: observed upper limits at 95% CL on $\tan\beta$ as a function of m_A, in the m_h^{max} benchmark scenario with $\mu = +200$ GeV for the individual *all-hadronic* and *semi-leptonic* results and for the combination of the two.

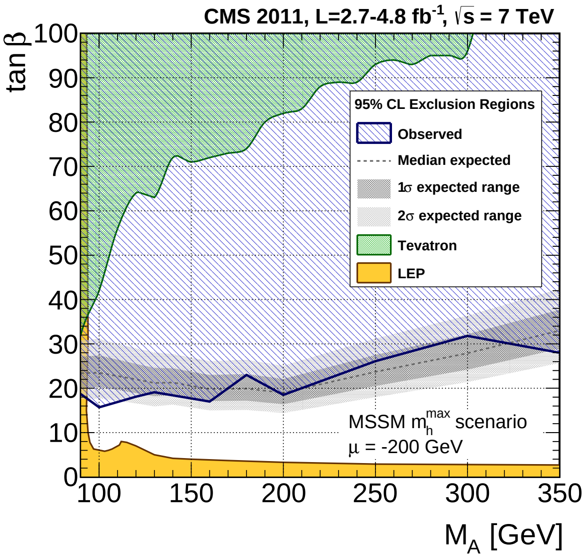

Figure 4.18: Exclusion in the m_A-$\tan\beta$ plane in the m_h^{max} benchmark scenario with $\mu = -200$ GeV for the combined all-hadronic and semi-leptonic results. Previous exclusion regions from LEP [56] and Tevatron in the multi-b jet channel [59] are overlaid.

Chapter 5

Search for the $A \to Zh \to \ell\ell b\bar{b}$ process

After the discovery of the Higgs boson in 2012, the perspectives for studies of a BSM Higgs sector changed considerably. Searches for this extended sector can be performed either by measuring deviations in the expected values of the couplings of the discovered h *boson to other SM particles, or via direct searches in final states disfavored by the SM. An interesting way to probe this kind of new physics is to search for bosons that decay into final states that contain the SM-like boson found in 2012, whose mass has already been measured. This chapter describes a search for a heavy pseudoscalar Higgs* A*, decaying into a* Z *and a* h *boson, both on-shell, with the* Z *decaying into a pair of $\ell^+\ell^-$ leptons (electrons or muons, i.e. ℓ being e or μ), and the* h *into a pair of* $\mathrm{b\bar{b}}$ *quarks [108]. The presence of three reconstructible resonances provides the opportunity to make use of several analysis techniques to increase the signal sensitivity. This channel is particularly constraining for the 2HDM and MSSM Higgs sectors in the low* $\tan\beta$ *regime, allowing to perform a search complementary to the search in the* $\mathrm{b\bar{b}\Phi b\bar{b}}$ *channel, which is instead sensitive to the high* $\tan\beta$ *region.*

5.1 INTRODUCTION AND ANALYSIS STRATEGY

This search for an heavy pseudoscalar boson A, decaying into a Z boson and a light h boson in the $\ell\ell\mathrm{b\bar{b}}$ final state, is performed on almost $20\,\mathrm{fb}^{-1}$ of data collected by CMS during 2012. The $\mathrm{A} \to \mathrm{Zh} \to \ell\ell\mathrm{b\bar{b}}$ analysis strategy is to reconstruct the Z, h, and A boson candidates from the visible decay products in the event. The final state is relatively easy to select using the online trigger since there are two high-p_T leptons, and easy to reconstruct because of the presence of well-measured decay products. The signal is expected to manifest itself as a peak in the four-body invariant mass ($m_{\ell\ell bb}$) spectrum over the SM expected continuum. The main standard model backgrounds are due to Z boson production with at least two accompanying b jets (genuine or mistagged), and $\mathrm{t\bar{t}} \to \mathrm{WW} \to \ell\nu\ell\nu\mathrm{b\bar{b}}$ decays. These backgrounds are evaluated and normalized directly using appropriate control regions in data. The standard model h boson produced in association with a Z boson also yields a potential background, but it differs from the signal because the $m_{\ell\ell bb}$ mass is not resonant. Signal sensitivity is improved by pursuing two different strategies:

1. exploit the signal topology and features to achieve optimal signal (S) efficiency and background (B) rejection using a multivariate analysis technique

2. improve the resolution of the reconstructed $m_{\ell\ell bb}$ invariant mass, exploiting the known value of the h boson mass to perform a kinematic fit on the jets.

In most 2HDM formulations, the A boson is produced predominantly through gluon-gluon fusion and decays to on-shell Z and h bosons provided $m_{\mathrm{A}} \gtrsim m_{\mathrm{h}} + m_{\mathrm{Z}} \approx 220\,\mathrm{GeV}$. This channel can be fruitfully studied for $m_{\mathrm{A}} \lesssim 2m_{top}$ and in the low-tan β region of the parameter space, where the decay $\mathrm{A} \to \mathrm{Zh}$ is generally dominant, but, depending on the model, it can still be sizable for higher values of m_{A}. The $\mathrm{h} \to \mathrm{b\bar{b}}$ decay has the advantage to have large branching fraction for most of the parameter space in 2HDM and MSSM. The expected sensitivity of the analysis, obtained by extrapolating the results of the standard model $\mathrm{h} \to \mathrm{b\bar{b}}$ search, is shown in Fig. 5.1. The expected exclusion region in the MSSM parameter space is obtained within a customized benchmark scenario, obtained by trading the information on the light Higgs mass for the radiative contributions of SUSY sparticles [109]. The $\mathrm{A} \to \mathrm{Zh} \to \ell\ell\mathrm{b\bar{b}}$ channel seems to be promising, being more stringent than the $\mathrm{H} \to \mathrm{hh} \to \mathrm{b\bar{b}b\bar{b}}$ and the $\mathrm{H} \to \mathrm{VV}$ searches in the low-tan β region and below the top quark kinematic threshold, in a portion of the parameter space where the $\Phi \to \mathrm{b\bar{b}}$ and $\Phi \to \tau\tau$ searches are not sensitive.

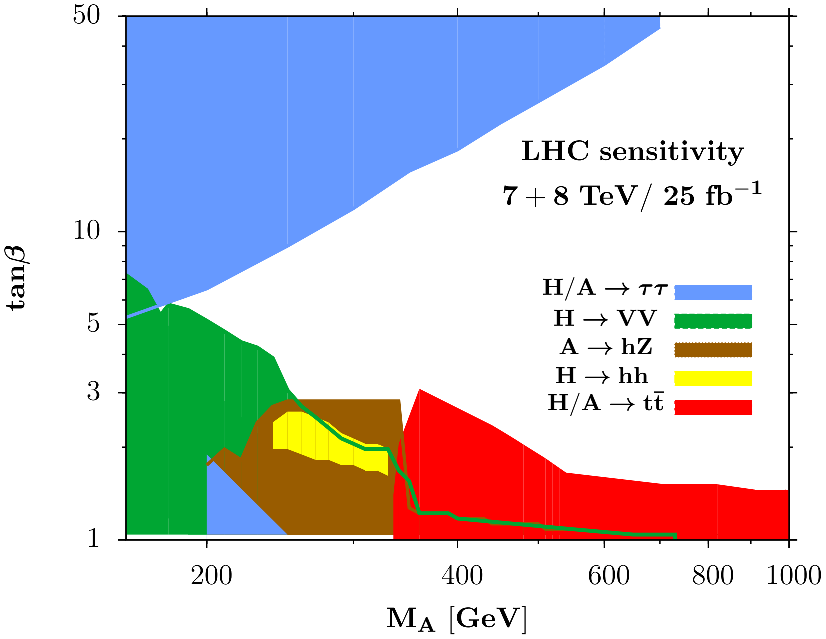

Figure 5.1: Expected sensitivity for different channels in the MSSM low-tan β region. The benchmark scenario is custom-defined, fixing the radiative corrections to m_{h}. The $\mathrm{A} \to \mathrm{Zh} \to \ell\ell\mathrm{b\bar{b}}$ analysis reach is obtained by projecting the results of the standard model search in the same final state [109].

Preliminary studies for this channel at CMS can be found in Ref. [110].

5.2 SIGNAL, BACKGROUND AND DATA SAMPLES

5.2.1 SIGNAL

The pseudoscalar boson A is assumed to have a narrow width, and to be produced via the gluon-gluon fusion process. Since no full next-to-leading-order (NLO) generator capable of handling the virtual loop diagrams as the ones shown in Fig. 5.2 is available for the considered process, an approximated gluon-gluon-(pseudo)scalar vertex has to be introduced within the

Higgs Effective Field Theory (HEFT) in the 2HDM. The A boson is forced to decay into on-shell Z and h bosons, and the Z boson is forced to acquire a transverse polarization with respect to its flight direction in order to preserve to total angular momentum. The Z is forced to decay only into electrons or muons, and the h into a pair of b quarks. Other decay modes are negligible after the applied analysis selections. The contribution from $\mathrm{Z} \to \tau\tau \to \ell\ell + \mathrm{X}$ is estimated to be 5 orders of magnitude smaller than the prompt $\mathrm{Z} \to \ell\ell$ decays. Because of the steeply falling sensitivity to the MSSM as $\tan\beta$ increases, the focus of the present analysis is on the gluon-gluon process alone, which is largely dominant for $\tan\beta \lesssim 3$ over other production processes; associated production with b quarks plays a role only for $\tan\beta \gtrsim 4$. In Fig. 5.3 production cross sections via the gluon-gluon fusion and associated production are compared for different $\tan\beta$ values.

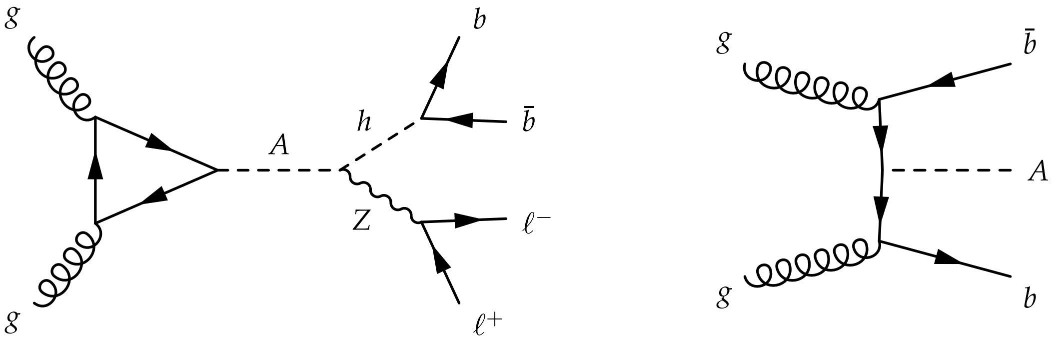

Figure 5.2: Feynman diagrams of the two main pseudoscalar Higgs boson processes. On the left is shown the gluon-gluon fusion, and the A decay into a Z and h bosons, in the $\ell\ell\mathrm{b\bar{b}}$ final state. On the right is the associated production with b quarks, dominant for high $\tan\beta$.

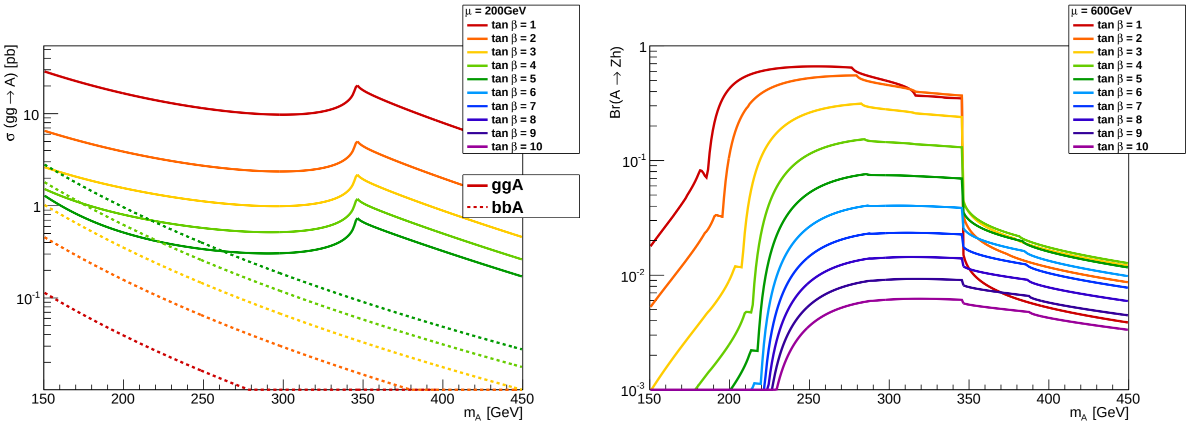

Figure 5.3: Left: production cross-section for the pseudoscalar boson A in the gluon-gluon fusion (solid lines) and associated production (dotted lines). Right: branching ratio $\mathcal{B}r(\mathrm{A} \to \mathrm{Zh})$ calculated by FEYNHIGGS in the m_h^{max} MSSM scenario with $\mu = 600\,\mathrm{GeV}$.

The newly-discovered SM-like Higgs boson is assumed to be the lightest scalar through the analysis, introducing the second resonance with a known mass in the decay and adding an additional constraint to the event kinematics. In the generated signal samples the light Higgs boson mass is fixed at $m_\mathrm{h} = 125\,\mathrm{GeV}$ and its width to few MeV, as predicted by the SM.

The $\mathrm{A} \to \mathrm{Zh} \to \ell\ell\mathrm{b\bar{b}}$ channel has its maximum sensitivity in a restricted interval of m_A masses, ranging from $m_\mathrm{A} \approx m_\mathrm{Z} + m_\mathrm{h}$ where the Z boson becomes on-shell to about $m_\mathrm{A} \approx 2 \cdot m_{top}$ where the branching ratio $\mathcal{B}r(\mathrm{A} \to \mathrm{Zh})$ sharply decreases because of the opening of the $\mathrm{A} \to \mathrm{t\bar{t}}$

channel. In this mass interval, a total of 6 signal samples were produced, at mass intervals separated by the typical experimental resolution of the 4-body invariant mass, which is of about 25 GeV. Additionally, since in some scenarios the A boson couplings with up-like quarks are suppressed, three more mass points are simulated to test the sensitivity to higher masses. In summary, nine mass points with 500 000 events each are simulated: 225, 250, 275, 300, 325, 350, 400, 500 and 600 GeV.

Other than the mass, another important property of the pseudoscalar boson is its width. Generally speaking, it fully depends on the model considered and its parameters. However, since A does not decay into vector bosons, the width of the pseudoscalar is generally smaller than the one of the scalar Higgs(es), even for higher masses, at least until the $\mathrm{t\bar{t}}$ decay threshold; in this case, a strong dependence on the model parameters is present (see Sec. 5.2.1.2). For $m_{\mathrm{A}} \lesssim 2m_{top}$, the typical A width ranges from some MeV to a few GeV. For this reason, the A natural width is assumed to be smaller than the experimental resolution (*narrow width approximation*), and it is kept fixed to an average value $\Gamma_{\mathrm{A}} = 0.1\,\mathrm{GeV}$.

Signal samples are generated with MADGRAPH [111], and the parton showering and hadronization are performed with PYTHIA using the $Z_2^\star$ tune [112]. The CTEQ6L [113] parton distribution functions are used in all simulated samples. The generated Monte Carlo events, including the pileup interactions, are processed through a full detector simulation based on GEANT4 [114] and reconstructed with the same algorithms used for real data. The PYTHIA generator has been excluded, since it does not take into proper account the polarization of the Z boson.

5.2.1.1 ASSOCIATED PRODUCTION

Even if the present analysis is designed around the gluon-gluon fusion production process, for certain regions of the 2HDM/MSSM parameter space the A production is association with b quarks can be sizable. The presence of an additional b quark in the final state can further enhance the signal sensitivity, and reduce the background yield. However, this analysis aims primarily to detect the $\mathrm{A} \to \mathrm{Zh} \to \ell\ell\mathrm{b\bar{b}}$ decay chain, trying to be as model-independent as possible on the A production mode. The associated production is studied by producing a specific sample with the MADGRAPH generator, with only 10 000 events for the $m_{\mathrm{A}} = 300\,\mathrm{GeV}$ mass point. The p_{T} and η distributions of the four b quarks at generation level are reported in Fig. 5.4. The plots show that the associated b quarks are generally produced with a softer p_{T} spectrum and at higher pseudorapidity than the b quarks from the Higgs decay, and only in a small fraction of events the leading spectator b quark (and even less for the sub-leading) is above the kinematic and geometrical thresholds, implying that in the majority of the events the additional b jets can not even be reconstructed or identified.

5.2.1.2 WIDTH

If the A boson natural width is smaller than the experimental resolution the exclusion upper limit does not directly depend on the width of the resonance. In order to check the validity of this assumption, a scan of the width Γ_{A} is performed within Type-I and Type-II 2HDM and in the MSSM parameter space. The result is that for $m_{\mathrm{A}} < 2m_{top}$ the width is smaller than the experimental resolution (≈ few GeV) for almost any parameter of the theory. For $m_{\mathrm{A}} > 2m_{top}$, with the accessibility of the $\mathrm{A} \to \mathrm{t\bar{t}}$ channel, Γ_{A} shows sharp variations near the kinematical threshold. Figure 5.5 and 5.6 report the A width above $2m_{top}$, which becomes strongly parameter-dependent and can easily be about $10 \div 30$ GeV or even larger in some particular regions, e.g. very low $\tan\beta \ll 1$, or $\cos(\beta - \alpha) \approx \pm 1$.

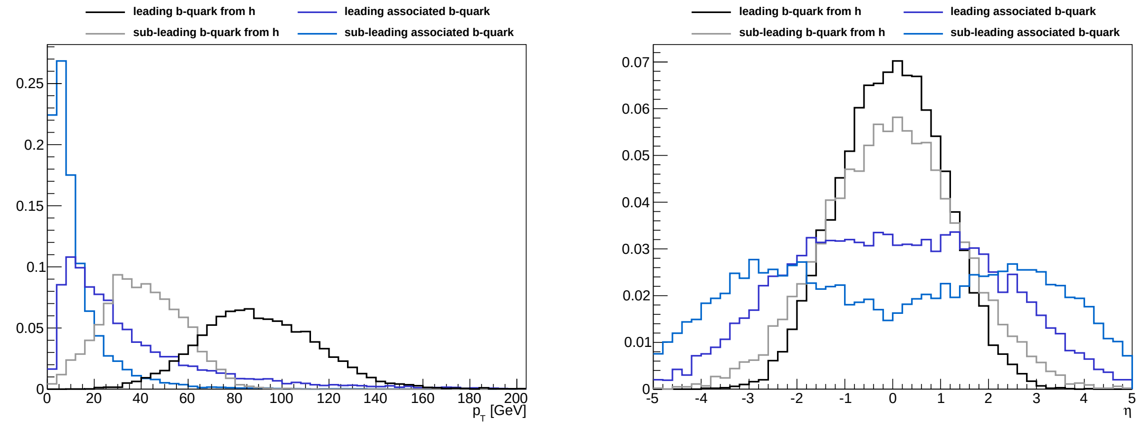

Figure 5.4: p_{T} (left) and η (right) distributions at generation level of the b quarks from h decay (black and gray lines) and from associated production (blue and azure lines).

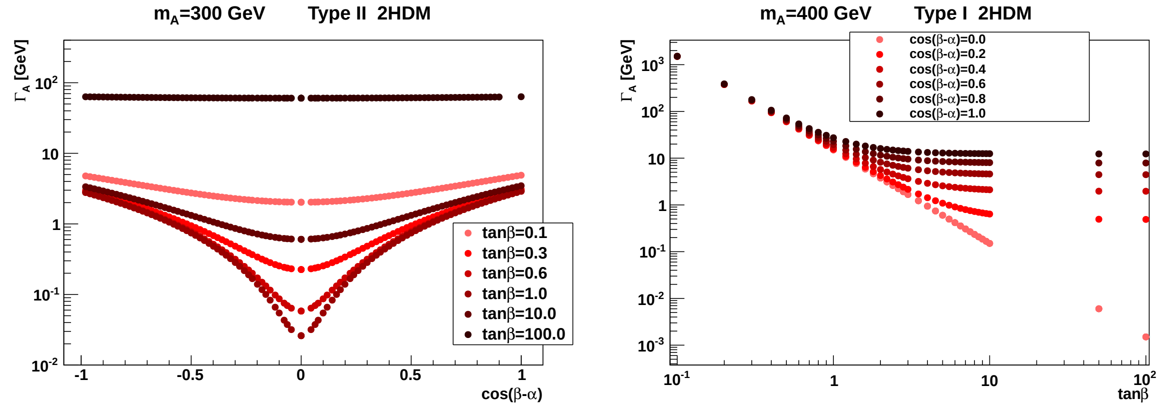

Figure 5.5: Γ_{A} as a function of $\cos(\beta - \alpha)$ in Type-II 2HDM for $m_{\mathrm{A}} = 300\,\mathrm{GeV}$ (left) and as a function of $\tan\beta$ in Type-II 2HDM for $m_{\mathrm{A}} = 400\,\mathrm{GeV}$ (right).

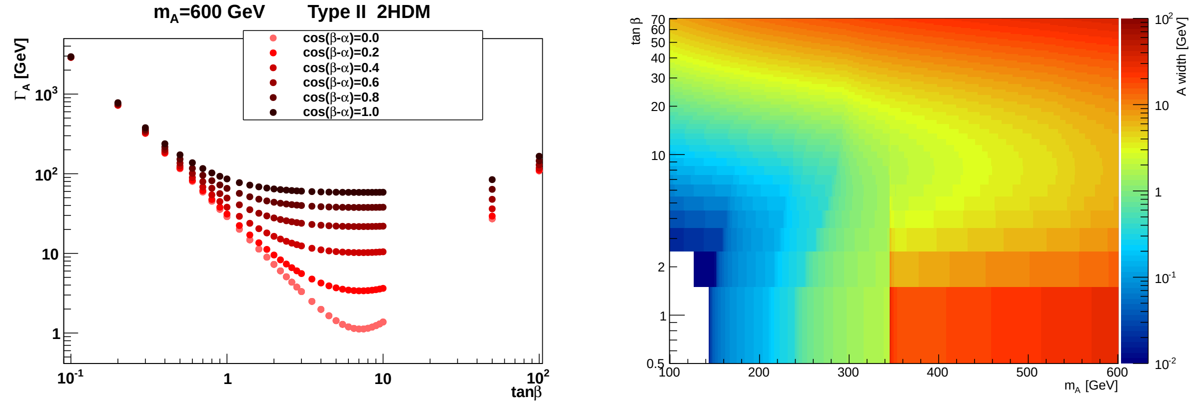

Figure 5.6: Γ_{A} as a function of $\tan\beta$ and for different values of $\cos(\beta - \alpha)$ (right) for $m_{\mathrm{A}} = 600\,\mathrm{GeV}$ in Type-II 2HDM (left) and in the $m_{\mathrm{A}} - \tan\beta$ plane in the MSSM m_h^{max} benchmark scenario (right).

5.2.2 BACKGROUNDS

All physics processes yielding two opposite-charge, same-flavour isolated leptons accompanied by at least two jets have to be considered as possible sources of background. The most relevant ones are described in the following and summarized in Tab. 5.1, and the cross section used to normalize SM backgrounds are measured by CMS or calculated at (N)NLO by the Standard Model Cross Section Working Group [12].

Drell-Yan: The production of single Z/γ^* bosons in association with one or more partons or gluons in the final state, resulting in two or more reconstructed jets, is topologically similar to the searched signal, but its final state objects feature a generally softer p_T spectrum, a non-resonant and rapidly falling dijet mass distribution, and other less distinctive characteristics (effective spin and colour radiation) that should theoretically distinguish it from the signal. Before b tagging, the contribution from light jets (initiated by *udsc* quarks or gluons) dominates, while after the application of b tagging the primary contribution in the signal region is from $Z + b\bar{b}$. Simulated Drell-Yan events are rescaled to the measured cross section [115]. The MADGRAPH $Z + b\bar{b}$ cross section is rescaled to the 76.75 pb value calculated with the MCFM software [116] treating b quarks as massive particles [12] by multiplying $Z + b$ and $Z + b\bar{b}$ normalization by a factor ≈ 1.3. Since the Monte Carlo is produced in the 5-flavour scheme ($m_b = 0$), $Z + b$ events are $Z + b\bar{b}$ events where one b-quark has $p_T < 10\,\mathrm{GeV}$ and does not pass generation cuts, and are treated as the same physics process. In order to increase the Monte Carlo statistics in certain regions of the phase-space, specific exclusive Drell-Yan samples are generated with different intervals in the number of partons involved in the hard interaction (from 1 to 4), p_T of the Z ($50 - 70$, $70 - 100$, > 100, > 180 GeV) and sum of the p_T of the involved partons ($200 - 400$, $400 - 600$, > 600 GeV). These datasets are merged with the inclusive sample, taking into account the event population in each interval, in order to minimize the statistical uncertainty in the final sample.

Top: Production of $t\bar{t}$ pairs represents a particularly challenging background at the LHC, given its large production cross section. These events always contain two energetic b jets and two W bosons that may decay to high p_T, isolated leptons. The primary handles to reduce the $t\bar{t}$ background are topological, exploiting the larger jet multiplicity and the presence of missing energy. Both semi-leptonic and fully-leptonic $t\bar{t}$ decays have been considered, the latter being much more important than the former for the final state considered. Other top quark production processes, not dominant for the purposes of this analysis, are the following:

Single Top : Single top events can be more difficult to reject with respect to $t\bar{t}$ events, but fortunately their cross section is low enough that they are no cause of concerns. Several production processes are present, but only those involving two real W bosons in the final state produce a sizable yield of dilepton plus dijet events.

t-channel : top is produced after a quark-gluon interaction with the exchange of a virtual W. For t-channel single top, the total cross section has been measured [117], while the t over $\bar{t}$ ratio is taken from MC.

s-channel : top is produced in association with an anti-bottom, after the annihilation of a pair of quarks in a weak vertex.

tW-channel : top is produced in association with a charged vector boson in a weak process, from a gluon-bottom pair in the initial state. This channel has the largest background contribution among single-top channels due to the presence of two real W in the final state.

Top+V : the associated production of top pairs and a vector boson (V=W, Z) may lead to a final state including many leptons and jets, but the cross section of such processes is very small.

Dibosons: The production of two vector bosons in the SM is a rare process, with a similar kinematics to that of the signal; the boost of the bosons is generally high. Three processes are considered [118–121]:

WW : very few events survive the data preselection, since both W should decay leptonically and their invariant mass fall in the Z mass window; the request of two jets results in a very low efficiency.

WZ : similar to the previous case, there is a low probability that such events can contaminate the signal region, because the hadronically-decaying W jets have to be both mistagged.

ZZ : even if its cross section is small, ZZ production is an important background because one Z can decay leptonically, while the other into b jets, leading to a final state very similar to the signal. The only difference is the value of the dijet invariant mass, that peaks at the Z mass instead of the h mass.

Other backgrounds are :

W *+jets*: consisting in the production of single W vector boson [115] in association with quarks or gluons, which includes an additional real or fake lepton that fires the trigger and mimics the final state signature. Despite the large cross section, very few events pass analysis cuts.

QCD: despite its enormous cross section at LHC, the probability to produce two real or spurious high-p_T isolated lepton candidates is very low. The QCD dataset is enriched in its muon content via a filter at generation level and a cut on the p_T of the hard-interaction $\hat{p}_T > 20\,\mathrm{GeV}$. Few events survive lepton selection, none of them make it to the signal region. Even if specific lepton-enriched samples are used, the knowledge of this background is limited by the finite MC statistics.

SM Higgs: the recently discovered SM-like scalar boson has to be included as background too, since its production in association with a Z boson is topologically very similar to the searched signal, as it features an on-shell Z and a couple of resonating b jets at $m_h = 125\,\mathrm{GeV}$. The Zh cross section is estimated to be $\sigma = 408 \pm 4\,\mathrm{pb}$, while its branching ratio to $h \to b\bar{b}$ is $\mathcal{B}r = 0.566 \pm 0.003$ for a 125 GeV Higgs [12].

All simulated events are produced with the MADGRAPH [111] matrix element generator, with the exclusion of the multijet QCD sample and the diboson samples, simulated with PYTHIA [124]. For SM Higgs production associated with a Z boson and single top processes, the POWHEG [125] generator is used. As the signal samples, all datasets are processed using GEANT4 [114] with a detailed CMS detector simulation, including the pileup events. The simulated distribution of the expected number of interactions in the same bunch crossing is reweighted to the one observed in 2012 data.

Process	Measured (pb)	Calculated (pb)	Source
W+jets	**35643** ± 560	36703 ± 424	[115]
Z+jets ($m_{\ell\ell} > 50$)	**3503** ± 171	3531.9 ± 39.4	[115]
Z+jets ($m_{\ell\ell} < 50$)	**13123** ± 640	11050 ± *n.d.*	[115]
t$\bar{\text{t}}$ (semi-leptonic)	**105.2** ± 5.6	93.64 ± 2.18	[122]
t$\bar{\text{t}}$ (full-leptonic)	**25.32** ± 1.34	23.64 ± 0.52	[122]
t (t-channel)	**49.9** ± 9.1	56.4 ± 2.4	[117]
$\bar{t}$ (t-channel)	**28.8** ± 5.5	30.7 ± 1.2	[117]
t (s-channel)	-	**3.79** ± 0.15	[12]
$\bar{t}$ (s-channel)	-	**1.76** ± 0.08	[12]
t (tW-channel)	**11.7** ± 2.7	11.1 ± 0.8	[123]
$\bar{t}$ (tW-channel)	**11.7** ± 2.7	11.1 ± 0.8	[123]
WW	**69.9** ± 7.0	55.48 ± *n.d.*	[121]
WZ	**24.6** ± 1.7	21.91 ± *n.d.*	[120]
ZZ	**7.7** ± 0.9	7.7 ± *n.d.*	[119]
t$\bar{\text{t}}$ + W	-	**0.232** ± 0.03	[12]
t$\bar{\text{t}}$ + Z	-	**0.2057** ± 0.02	[12]
ZH	-	$\mathbf{23.3 \cdot 10^{-3}}$	[12]

Table 5.1: The cross section for the considered backgrounds, multiplied by their branching ratios when meaningful. Each different sample notation includes all the decays taken into account in the calculation. Measured cross sections were used when available, (N)NLO calculations otherwise. The bold numbers are those effectively used in the analysis.

5.2.3 DATA

The data sample used in this analysis has been collected during 2012, at a center-of-mass energy of 8 TeV, and the total integrated luminosity amounts to $\int \mathcal{L}dt = 19.7 \pm 0.5\,\text{fb}^{-1}$ [126]. The analyzed events have been selected by two different trigger paths:

HLT_Mu17_Mu8: requires two muons with p_T larger than 17 and 8 GeV, respectively, within $|\eta| < 2.4$ and without any selection on muon isolation;

HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL: demands two isolated electrons with p_T larger than 17 and 8 GeV, respectively, satisfying online quality and isolation requirements.

The trigger paths are simulated in Monte Carlo events, and required to be fired as in the data sample. Trigger scale factors are applied to take into account residual efficiency differences with the trigger simulation, and to cope with the different efficiency of the trigger menus during the 2012 data taking period. The data samples used are listed in Tab. 5.2.

5.3 EVENT SELECTION AND BACKGROUND CONTROL REGIONS

5.3.1 PRESELECTION AND SIGNAL REGION

Data and Monte Carlo events have to meet certain requirements to be considered suitable for the analysis. A basic screening, called *preselection*, selects only events firing the double-muon or double-electron trigger. At least two opposite-sign, same-flavour leptons, have to be reconstructed within geometrical acceptance ($|\eta| < 2.5$ for electrons, $|\eta| < 2.4$ for muons); the p_T threshold is set to 20 GeV for the lepton with largest p_T, and to 10 GeV for the lepton with

	Double-muon		Double-electron	
Period	Events	L (fb^{-1})	Events	L (fb^{-1})
Run A	1 340 460	0.912	1 162 721	0.876
Run B	12 251 486	4.511	5 868 916	4.511
Run C	13 894 075	7.017	9 306 031	7.055
Run D	12 536 714	7.369	6 178 749	7.369
Total	39 961 200	19.67 ± 0.51	22 516 417	19.71 ± 0.51

Table 5.2: Summary of data samples used in this analysis. The triggers considered are *HLT_Mu17_Mu8* for the double-muon dataset, and *HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL* for the double-electron one.

next-highest p_{T}. The Z boson candidate is formed from the two highest-p_{T}, opposite-charge, same-flavour leptons, and must have an invariant mass larger than 50 GeV. In addition, at least two jets are required with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$, within $|\eta| < 2.4$, and with angular separation relative to each lepton of $\Delta R > 0.5$.

A signal-enriched region (henceforth called *signal region* or SR) is selected requiring two leptons and at least two b-tagged jets, identified with the CSV b-tagging algorithm, with the *tight* working point on the leading CSV jet (also denoted as CSVT, or CSV$>$ 0.920) and the *loose* (CSVL, or CSV$>$ 0.405) on the subleading one. The h candidate is built from the two jets that have the highest b tagging CSV value. With this choice, the jets originating from the Higgs decay are correctly selected in about 80% of the events after preselection cuts, and $\geq$ 97% in the signal region. The fraction is lower, $\approx$ 60% and $\approx$ 80%, for events produced in the b quark associated production due to the higher b-jet multiplicity in the event. Additional jets are ignored, and any event not fulfilling these conditions is discarded. Finally, the invariant masses of the lepton and jet pairs should be compatible with those of the Z and h boson ($80 < m_{\ell\ell} < 100$ and $90 < m_{\mathrm{bb}} < 140\,\mathrm{GeV}$ respectively), and a small missing energy ($E_{\mathrm{T}}^{\mathrm{miss}} < 60\,\mathrm{GeV}$) has to be observed in the detector.

The data reduction with the selections described above is summarized in Tab. 5.6 and Fig. 5.7. The multijet events are suppressed effectively by the requirement of two leptons, and reduced to a negligible level by the Z selection, like the W +jets. The requirement of two b-tagged jets reduces strongly the large Z +jets background. The most important background after the analysis cuts are $\mathrm{Z} + \mathrm{b\bar{b}}$, $\mathrm{Z} + b$, $\mathrm{t\bar{t}}$, Z+jets, while few single top and VV events, mostly Z Z, still survive.

5.3.2 BACKGROUND CONTROL REGIONS

Dedicated control regions (CR) are defined to check both the normalization and shape in crucial kinematical distributions of the most important backgrounds by inverting the selections that enhance signal. Drell-Yan backgrounds are considered separately as a function of the number of b jets, distinguishing Z +jets (no b jets), Z +b (1 b jet) and Z +$\mathrm{b\bar{b}}$ (2 b jets), and their corresponding control regions selected by requiring the proper number of b-tagged jets. The tight and loose working-point selections can be used to tag the leading and next-to-leading jets in the event. In the Z +b control region, a b tag veto is applied to jets other than the b-tagged jet. Additional kinematic selections include a cut on the dilepton invariant mass around m_{Z}, low $E_{\mathrm{T}}^{\mathrm{miss}}$ and a dijet invariant mass veto close to the Higgs boson mass. The $\mathrm{t\bar{t}}$ control region is

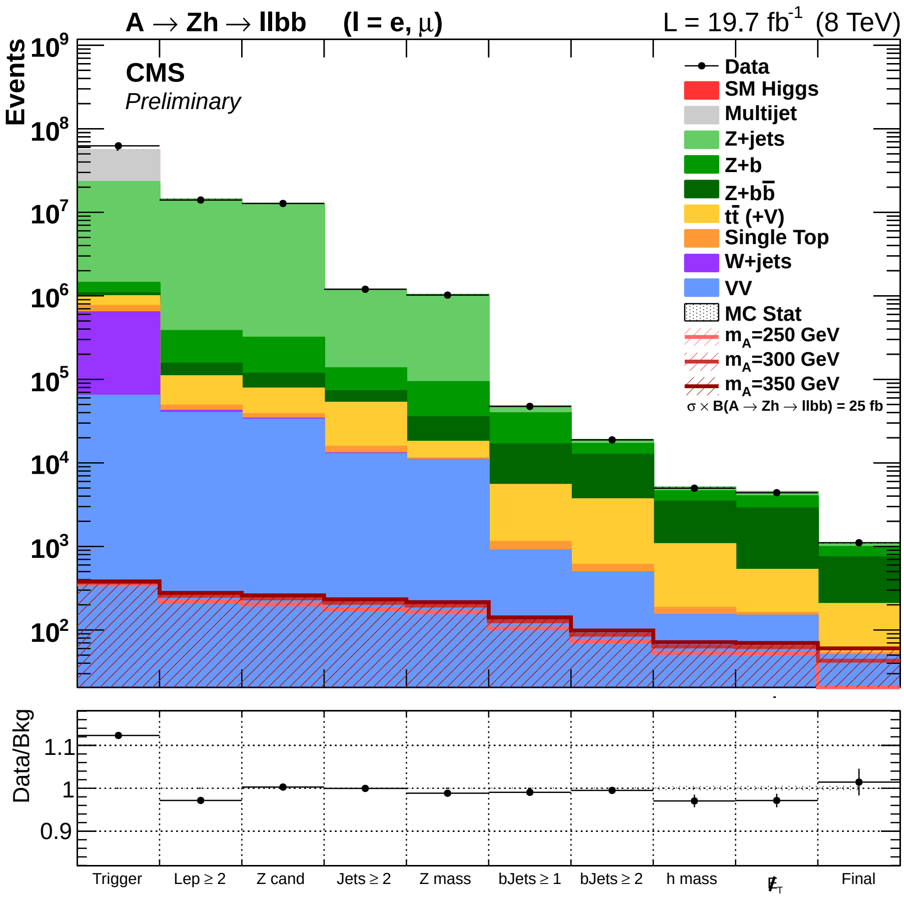

Figure 5.7: Events reduction for data and background after selections. Also three signals sample, with m_{A} =250, 300, and 350 GeV are shown: the signals are normalized to a cross section times branching ratio of 25 fb. The *Final* label identifies the event yield after the final selections for the intermediate mass cut-based analysis.

defined by inverting the $m_{\ell\ell}$ and $E_{\mathrm{T}}^{\mathrm{miss}}$ cutoffs, dropping the dijet veto, and requiring at least two b-tagged jets.

Control Region	Z mass (GeV)	h mass (GeV)	CSV$_1$	CSV$_2$	$E_{\mathrm{T}}^{\mathrm{miss}}$ (GeV)
Z +0 b-jets	$80 << 100$	$< 90, > 140$	-	-	-
Z +1 b-jets	$80 << 100$	$< 90, > 140$	Tight	not Loose	< 40
Z +2 b-jets	$80 << 100$	$< 90, > 140$	Tight	Loose	< 40
Top	$< 80, > 100$	-	Tight	Loose	> 40

Table 5.3: Control regions selections. Signal is avoided by vetoing $m_{\ell\ell}$ close to the Z mass, or m_{bb} within resolution with m_{h}.

Corrective scale factors for normalization, reported in Table 5.4, are obtained from a simultaneous likelihood fit to the difference between data and simulation in the four control regions, and are applied in the following steps of the analysis. Scale factors account for not only cross section discrepancies, but also for potential residual differences in physics objects selection efficiencies. The $\mathrm{t\bar{t}}$ and Z +jets scale factors are dominated by the corresponding control region, the former since it is the only control region where its presence is dominant, the latter given the larger statistics of the corresponding control region. The Z +b$\bar{b}$ and Z +b ones are instead determined both by Z +b$\bar{b}$ and Z +b regions.

The yield, and distributions of less significant backgrounds from diboson, single top quark, and Vh events, are taken from simulation, using measured cross sections.

	Z +jets	Z +b	Z +b$\bar{b}$	t$\bar{t}$
Scale Factors	1.069 ± 0.002	0.945 ± 0.012	1.008 ± 0.020	0.984 ± 0.010

Table 5.4: Rescaling factors for control regions obtained for the four main backgrounds. The uncertainties correspond to just statistical contributions.

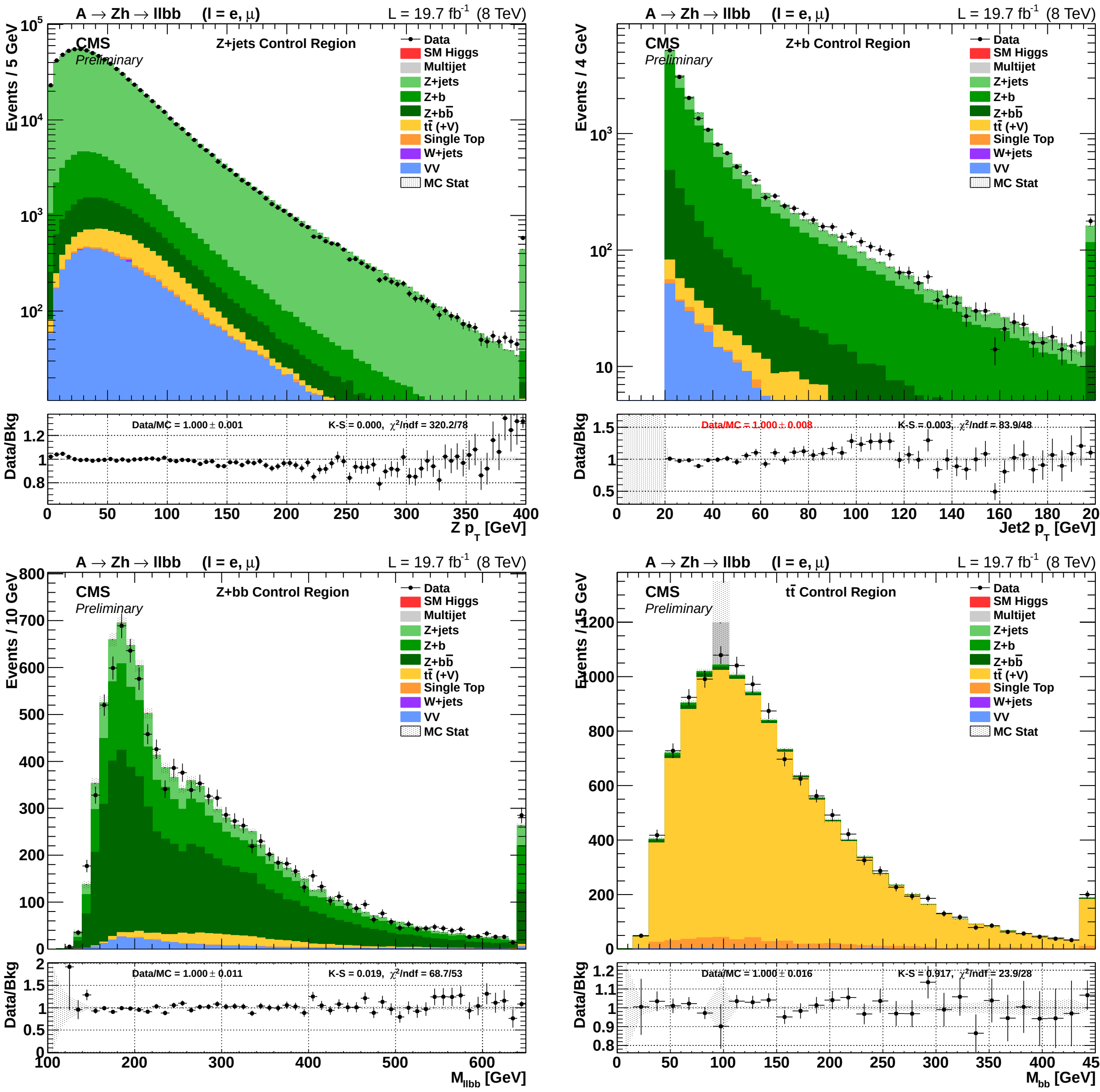

Table 5.5: Comparison between data and simulation for some of the most important kinematic variables in the four control region. Top left: Z candidate transverse momentum in the Z +jets control region. Top right: sub-leading jet p_T in the Z +b control region. Bottom left: four body invariant mass ($m_{\ell\ell bb}$) in the Z +b$\bar{b}$ control region. Bottom right: dijet invariant mass in the top control region.

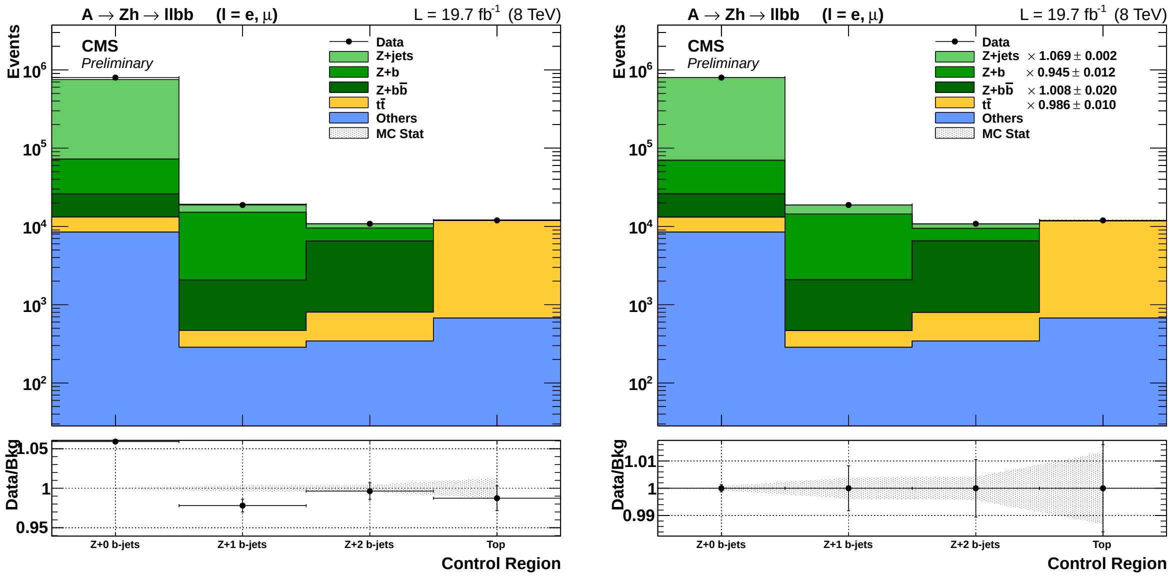

Figure 5.8: MC yield before (left) and after (right) the likelihood fit, with scale factors applied.

5.3.3 MULTIJET ESTIMATION

Multijet contamination in control and signal regions is evaluated with data, due to the very small selection efficiency for this kind of background. The yield is predicted with the *ABCD* method, inverting lepton isolation criteria in the plane of the two variables, sketched in Fig. 5.9, enhancing the QCD contribution with respect the other physics processes. The method is applied to either control regions (Z +b$\bar{b}$, t$\bar{t}$) or the signal region. The predicted yield in the region D, where the multijet contribution is estimated, is given by the ratio of the number of QCD events observed in the regions A, B, C by the simple relation $N_D = N_A / N_B \times N_C$.

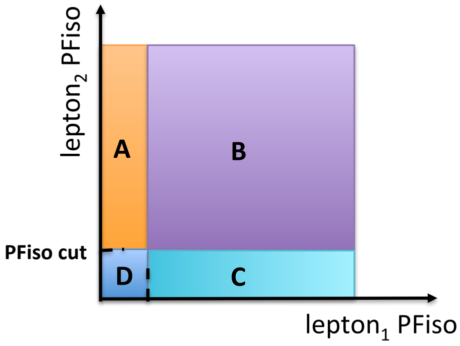

Figure 5.9: Graphical representation of the regions used in the ABCD method.

Due to the lack of simulated samples with an adequate statistics, the QCD contribution is considered as the difference between data and the electroweak backgrounds. Results using the QCD simulation are considered only as a cross check, and are fully compatible with the data-driven method, although with a larger statistical error. The multijet contamination is estimated to be 6.5 ± 2.3 in the Z +b$\bar{b}$ control region, and 3.0 ± 1.7 in the signal region, and the contribution of this background is considered negligible.

5.3.4 SIMULATION MISMODELINGS

The simulated Drell-Yan backgrounds show some disagreement in few variables when compared to data in the corresponding control regions. The first variable to be considered is the pseudorapidity of the two b-tagged jets, which show an nonphysical peak around 0, not visible in data. Investigations evidence that the peak is present only for b partons, treated as massless by MADGRAPH but massive in PYTHIA. Since no exact method for removing the peak has been found, the Monte Carlo is reweighted to match the data, and the difference is considered as a shape systematic uncertainty on the simulated background.

The second, and most important effect is the disagreement, observed in the Z +b$\bar{\text{b}}$ control region, of some distributions such as the $\cos\theta^*$ angle and centrality of the four main objects in their rest frame (defined in Sec. 5.4), which reflects also on other key variables, like the 4-body invariant mass. In particular, the centrality distribution exhibits a clear linear trend, visible in all Drell-Yan and t$\bar{\text{t}}$ control regions. A per-event weight is calculated as the ratio of the centrality in data over simulation in the Z +b$\bar{\text{b}}$ control region using a first degree polynomial, and it is employed to correct the Drell-Yan and t$\bar{\text{t}}$ backgrounds. The difference between the unweighted and weighted simulation distributions is not negligible, and is taken as a source of MC modeling systematic uncertainty. After the reweighting, an improvement in simulation agreement is observed in all distributions; the linear trend observed in the m_{A} distribution disappears as well, as visible in Fig. 5.10.

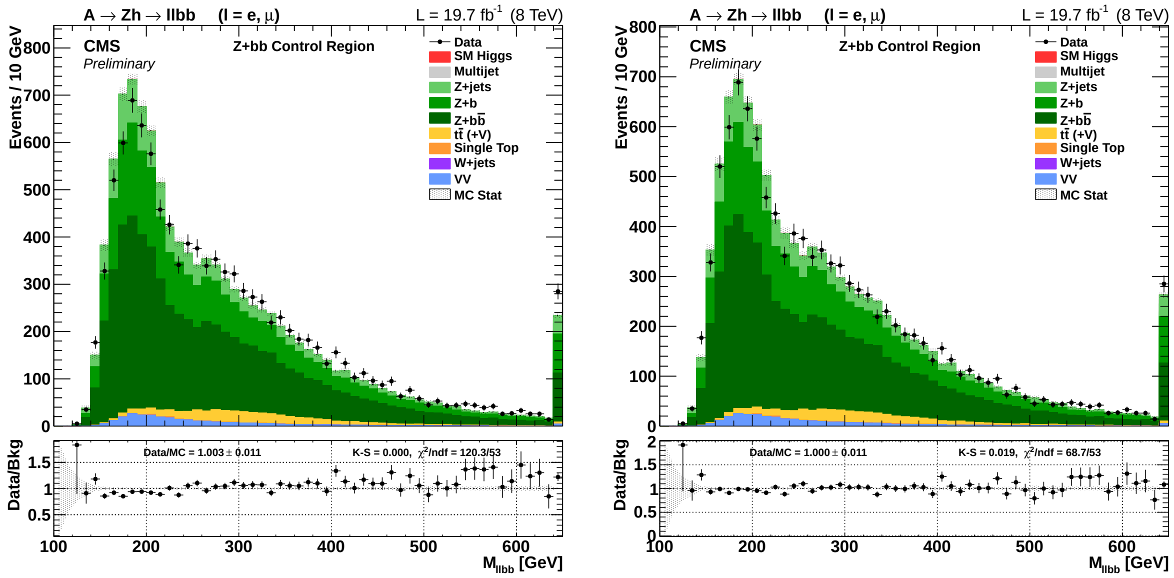

Figure 5.10: 4-body invariant mass in the Z+heavy flavour control region before (left) and after (right) the reweighting.

Dataset	Trigger	Lep ≥ 2	Z cand	Jets ≥ 2	Z mass	bJets ≥ 1	bJets ≥ 2	h mass	$E_{\mathrm{T}}^{\mathrm{miss}}$	Final
Data	62 486 296	14 024 408	12 750 833	1 198 888	1 021 320	47 491	18 863	4 979	4 400	1 108
VV	64 289	39 903	33 237	12 836	10 890	897	491	152	149	50
W+jets	574 850	2 285	941	232	41	–	–	–	–	–
Single Top	123 266	6 751	4 496	2 656	466	241	117	33	11	3
tt (+V)	238 632	60 837	39 548	37 271	6 627	4 347	3 087	889	368	152
Z+bb	76 722	45 332	38 739	19 250	17 482	11 233	8 824	2 368	2 332	537
Z+b	360 589	226 791	199 773	63 834	57 952	22 965	4 407	1 166	1 150	245
Z+jets	21 818 907	13 964 489	12 394 090	1 062 179	939 520	8 205	1 993	495	491	89
Multijet	32 377 032	87 741	3 526	911	127	–	–	–	–	–
SM Zh	190	125	115	98	91	56	38	28	27	14
All Bkg.	55 634 480	14 434 255	12 714 468	1 199 268	1 033 197	47 944	18 957	5 131	4 529	1 092
$m_A = 250\,\mathrm{GeV}$	351	217	202	176	163	104	72	53	52	11
$m_A = 300\,\mathrm{GeV}$	370	256	238	211	195	127	87	63	62	42
$m_A = 350\,\mathrm{GeV}$	381	278	259	232	215	141	98	71	69	60

Table 5.6: Event selection for data and background after preselections (top) and analysis cuts (bottom). Also three signals samples, with m_{A} =250, 300, and 350 GeV are shown: the signals are normalized to a cross section times branching ratio of 25 fb. The *Final* label identifies the event yield after the final selections for the intermediate mass cut-based analysis. The *Final* label identifies the event yield after the final selections for the intermediate mass cut-based analysis.

5.4 DISCRIMINATING VARIABLES

The $\mathrm{A} \to \mathrm{Zh} \to \ell\ell\mathrm{b}\bar{\mathrm{b}}$ decay yields a very characteristic signature with respect to the irreducible backgrounds. The possibility to reject the latter, besides the information on the mass of the resonances, relies in finding characteristic variables of the signal process. Two approaches can be attempted: the first and most immediate one is to consider variables which are natural from an experimental point of view, such as the final objects transverse momenta, invariant masses, pseudorapidity and azimuthal separation, and many more. This can be thought of as a type of bottom-up parametrization, and the full set of variables can be found in Sec. 5.4.1.

The alternative is a top-down parametrization, motivated by the physical process. A good starting point are the variables that characterize the pseudoscalar production, such as the polar angle, and follow the decay chain up to the final objects (leptons and quarks) in their rest frame. This procedure is described in Sec. 5.4.2. Additional useful information can be carried by event variables, built from all the objects reconstructed in the event (Sec. 5.4.3); in Sec. 5.4.4 also radiation-specific variables are investigated, exploiting the colour properties between the quarks originating the two jets.

5.4.1 KINEMATICS

The basic step is to use the final objects' four momenta in the lab frame and see if any visible difference between the backgrounds and the signal samples are present. Different four-vectors can be combined if the resulting variable does not loose too much information. A good starting point is the kinematic of the reconstructed A candidate, namely the object composed by the four observable objects (leptons and jets). The p_{T} tails are usually larger for the signal with respect to backgrounds, making it interesting for a possible search in the high-p_{T} regime. Unfortunately, the signal yield is so small that the advantage of such an approach is not clear. Some differences can be observed if looking at the pseudorapidity, the signal being produced in a more forward region, but this quantity heavily relies on details of the PDF in non-well-constrained regions of parton x and Q^2.

More interesting are the A decay products, the Z and h candidates. The p_{T} of these objects, especially the Z, is particularly effective at discriminating backgrounds, provided the A is massive enough to give the daughters the required momentum. Indeed, the Z p_{T} spectrum is steeply falling in Drell-Yan processes, initiated by a $q\bar{q}$ state resulting the vector boson to be produced at smaller transverse momentum. On the other hand, also the h boost is equally effective, except at low m_{A} where is overwhelmed by Z +jets events. Also the angular separation between the two intermediate candidates is investigated, but it is strongly correlated with A p_{T}.

While a real, on-shell Z boson is present almost always in both signal and background, there is a stronger motivation to look at the dijet system more carefully, since its origin is completely different in signal and background. In signal, the two b jets come from a single, massive object, but in Drell-Yan processes they come from additional gluons or quarks involved in the hard scattering and are largely uncorrelated. The simplest variables are:

- the p_{T} of the two jets, since the $Z + bb$ irreducible background has usually a softer p_{T} spectrum
- $\Delta\eta_{\mathrm{b\bar{b}}} = |\eta_{b1} - \eta_{b2}|$ because signal tends to emit b jets with a closer angle with respect to the DY+jets backgrounds
- $\Delta\varphi_{\mathrm{b\bar{b}}} = \pi - ||\varphi_{b1} - \varphi_{b2}| - \pi|$, for the same reason as above
- $\Delta R_{\mathrm{b\bar{b}}} = \sqrt{\Delta\eta^2_{\mathrm{b\bar{b}}} + \Delta\varphi^2_{\mathrm{b\bar{b}}}}$ gives the separation in the $\eta - \varphi$ plane, and it is quite corre-

lated to the h candidate boost

Taking $\Delta R_{\mathrm{b\bar{b}}}$ instead of $\Delta\eta_{\mathrm{b\bar{b}}}$ and $\Delta\varphi_{\mathrm{b\bar{b}}}$ is often advantageous for different reasons, but it entails a loss of useful information. To overcome this issue, another polar angle, called *twist*, can be introduced. Thinking of $\Delta\eta$, $\Delta\varphi$ as 2D Cartesian coordinates, the polar coordinate is the well-known ΔR, while the azimuthal one is the twist angle τ, defined as:

$$\tau \equiv \tan^{-1}\frac{\Delta\varphi}{\Delta\eta}$$

Twist is a longitudinal boost-invariant version of the rotation of the h b$\bar{\mathrm{b}}$ plane with respect to the beam-h plane, and it is zero when the particles are separated along the cylinder in η, and $\pi/2$ when separated around the cylinder in φ.

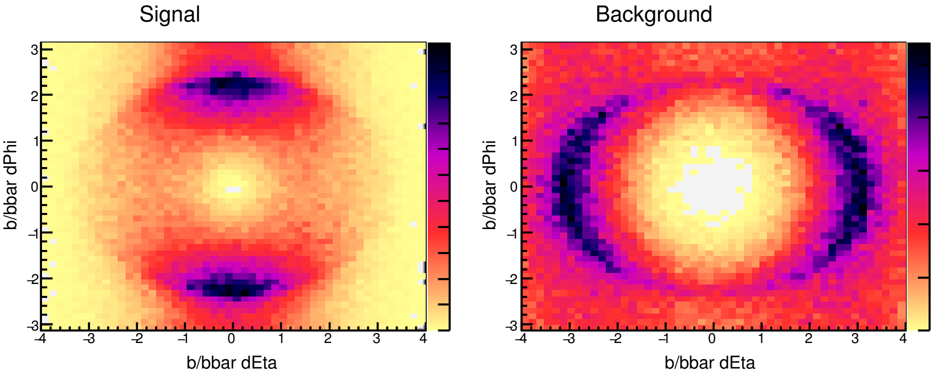

Figure 5.11: $\Delta\eta_{b\bar{b}}$ vs $\Delta\phi_{b\bar{b}}$ for the Higgs boson signal (left) and the gg initiated $Zb\bar{b}$ background (right). These distributions are obtained at the parton level, and for a moderate h boost ($p_{\mathrm{T}} > 50$ GeV) [127].

The twist angle has an additional advantage. While $\Delta R_{\mathrm{b\bar{b}}}$ is strongly correlated with the Higgs boson p_{T}, twist, being orthogonal to $\Delta R_{\mathrm{b\bar{b}}}$ ends up being very uncorrelated with p_{T}^{H}, becoming useful even at high Higgs boson p_{T}. Twist is also uncorrelated with m_{bb}, since m_{bb} is basically determined by $\Delta R_{\mathrm{b\bar{b}}}$ and p_{T}^{H}, so it makes a nearly ideal complimentary discriminant to $\Delta R_{\mathrm{b\bar{b}}}$.

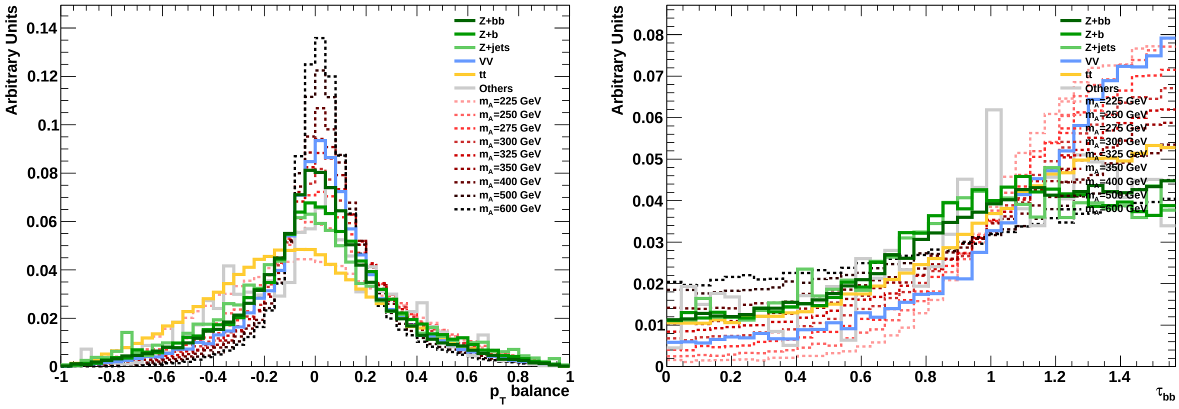

Figure 5.12: p_{T} balance (left) and twist angle between the two jets (right).

5.4.2 ANGULAR VARIABLES

There are several features in the signal A $\rightarrow$ Zh $\rightarrow$ $\ell\ell$b$\bar{\text{b}}$ decay kinematics which can help discriminate against the background. In fact, five helicity-dependent angular observables fully describe kinematics in a 2 $\rightarrow$ 1 $\rightarrow$ 2 $\rightarrow$ 4 decay [128]; they are independent on the three invariant masses of the particles involved (m_{A}, m_{Z}, m_{h}) and on the longitudinal and transverse momenta, but typically they have weaker discrimination power and rely on process production dynamics. These variables are largely uncorrelated and are more attractive to be used in a collective discriminator rather than independently.

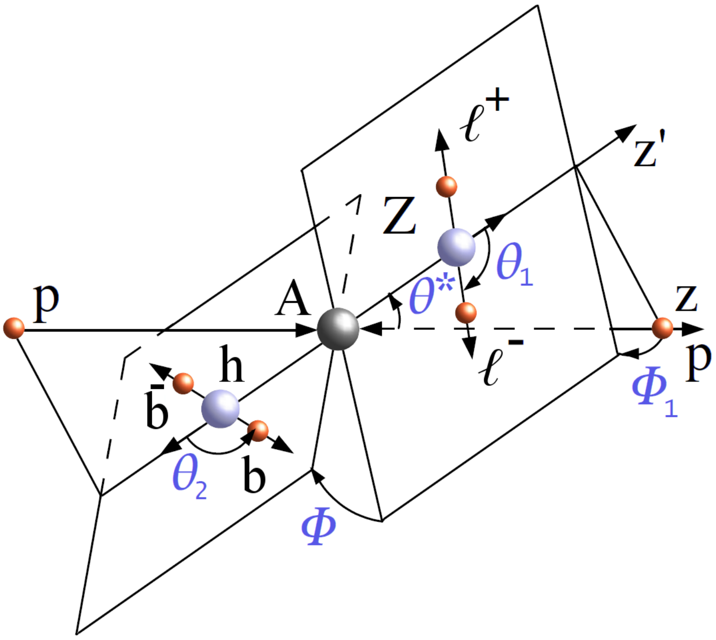

Figure 5.13: Diagram of the $A \rightarrow Zh \rightarrow \ell\ell bb$ decay and the definition of the five helicity angles.

The five angles are sketched in Fig. 5.13; two of them are production angles (θ^*, Φ_1), the remaining three decay angles (θ_1, θ_2, Φ). Their normalized distributions are shown in Fig. 5.14 and are defined as:

θ^*: angle between the Z flight direction and the beam in the A rest frame

θ_1: angle of the negatively charged lepton and the Z flight direction in the Z rest frame

θ_2: same as θ_1, but for h and jets. Since there is no unambiguous way to know the charge of the b parton, this variable is symmetrized, and the most energetic jet is chosen by convention

Φ: angle between the Z and h decay planes

Φ_1: angle between the Z decay plane and the plane of the Z flight direction and the beams.

A *likelihood discriminant* is constructed starting from the signal and background probability distribution functions s_i and b_i, parametrized by proper functions. A likelihood has been chosen for its simplicity and robustness with respect to other multivariate methods, and because variables have a low correlation. The likelihood ratio discriminator is defined as:

$$\mathcal{D} = \frac{\prod_i s_i(x_i)}{\prod_i s_i(x_i) + \prod_i b_i(x_i)}$$

The training of the likelihood has been performed with simulated events in the signal region, and a single signal sample, $m_A = 350\,\text{GeV}$, is selected; the other signal samples do not differ significantly. The five angular variables for background and signal samples, and the final discriminator output are shown in Fig. 5.14.

Due to the relatively small discrimination power, the Likelihood discriminant is used only in the cut-based analysis; being almost all the discrimination due to $\cos\theta^*$ and $\cos\theta_1$, these two angles are directly used as input of other multivariate methods in the final steps of the analysis.

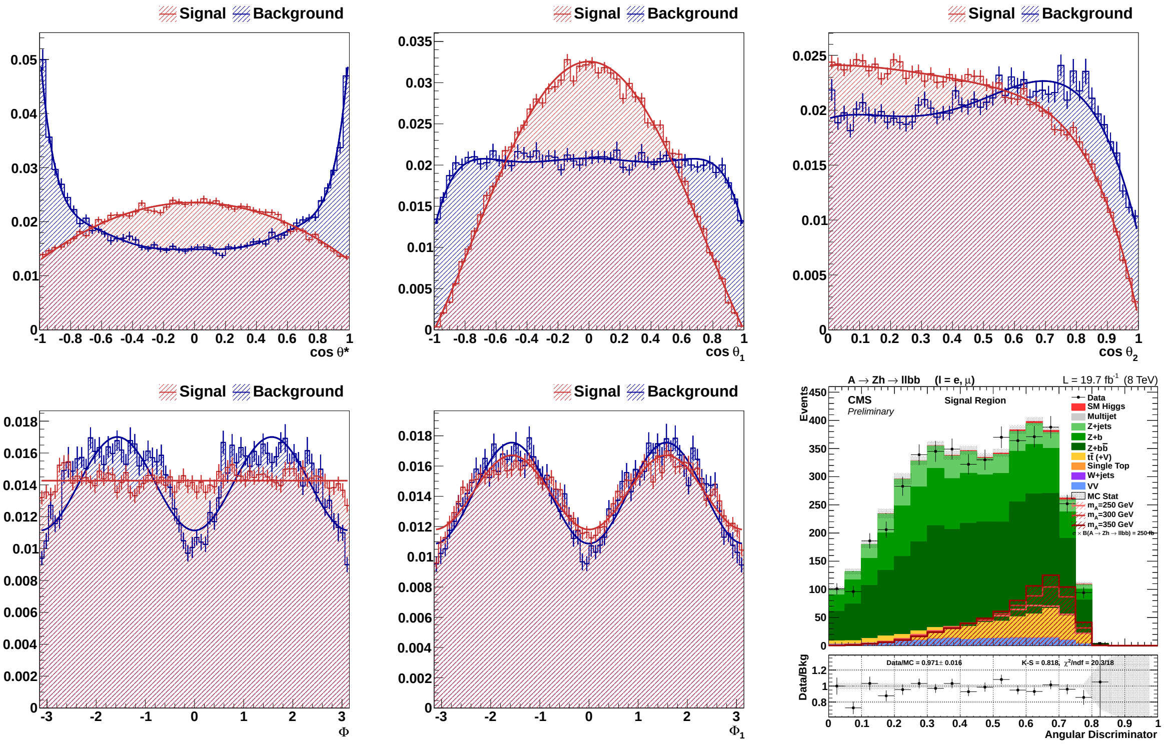

Figure 5.14: Signal (red) and background (blue) probability distribution functions for the five angles. Bottom left: angular discriminator output in the signal region.

5.4.3 EVENT SHAPE

Beside the 4-momenta of the four main objects, additional information could be carried by variables calculated with all objects in the event. The most natural ones are:

H_T: scalar sum of E_T of all hadronic objects in the event passing a p_T threshold $> 20\,\text{GeV}$

S_T: scalar sum of H_T, leptons p_T and E_T^{miss}

Centrality: defined as $\sum p_T / E_{vis}$ the sum of the p_T of all objects divided by the visible energy

In addition to these, other event shape variables can be calculated. These involve the eigenvalues of a tensor, the *Sphericity Tensor*, composed by objects' 3-momenta [127]:

$$\frac{1}{\Sigma_i |\vec{p}_i|^2} \sum_i \begin{pmatrix} p_x p_x & p_x p_y & p_x p_z \\ p_y p_x & p_y p_y & p_y p_z \\ p_z p_x & p_z p_y & p_z p_z \end{pmatrix}$$

The eigenvalues of this matrix are computed, then ordered and normalized $\lambda_1 \geq \lambda_2 \geq \lambda_3$ with $\lambda_1 + \lambda_2 + \lambda_3 = 1$. The event shapes are then defined as in Ref. [129].

Sphericity: $S = \frac{3}{2}(\lambda_2 + \lambda_3)$ where $0 \leq S \leq 1$. A 2-jet event has $S \approx 0$ while an isotropic one has $S \approx 1$.

Aplanarity: $A = \frac{3}{2}\lambda_3$ where $0 \leq A \leq 1/2$. A planar event has $A \approx 0$ while an isotropic one has $A \approx 1/2$.

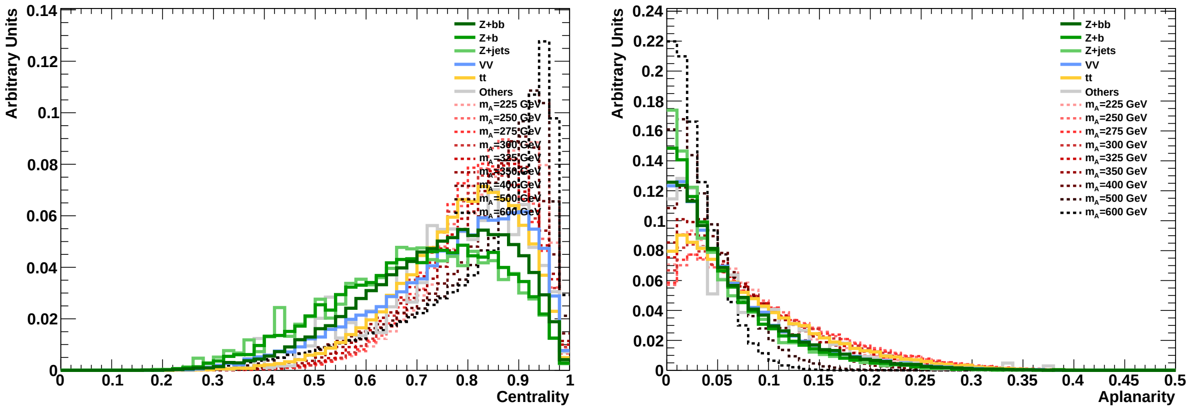

Figure 5.15: Centrality calculated in the A rest frame (left) and Aplanarity (right).

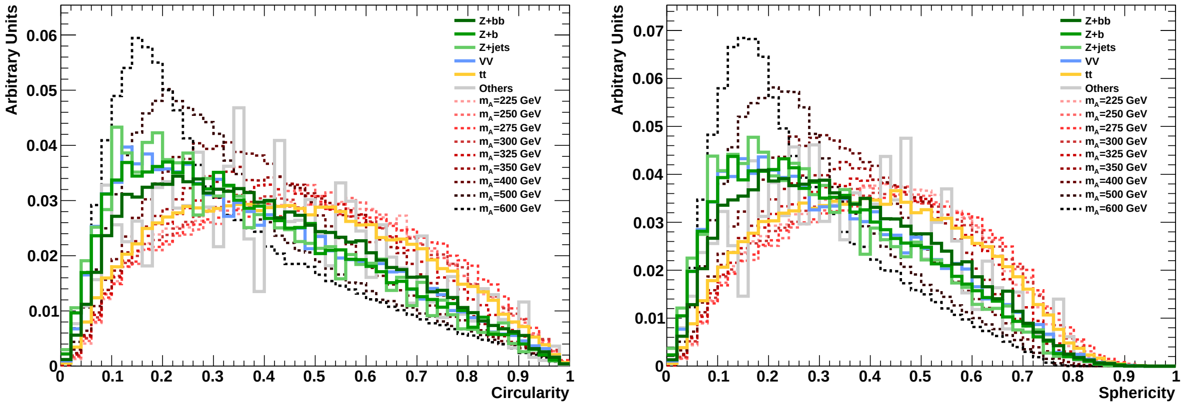

Figure 5.16: Circularity (left) and Sphericity (right).

5.4.4 RADIATION

One peculiar difference between signal and background lays in the fact that the two quarks produced by the Higgs decay are colour-connected to each other (the Higgs is a colour singlet), while in the background they are often colour-connected to the proton remnants that travel down the beam pipe. This is shown schematically in Fig. 5.17. This difference is independent on the event kinematics, and, if observable, it is complementary to kinematical variables and can add useful information in a multivariate search.

Since colour flow is physical, it can be possible to extract the colour connections of an event only if they have an observable effect on the distribution of the observable hadrons. The *pull vector* is an observable quantity at hadron colliders specifically designed to measure the colour

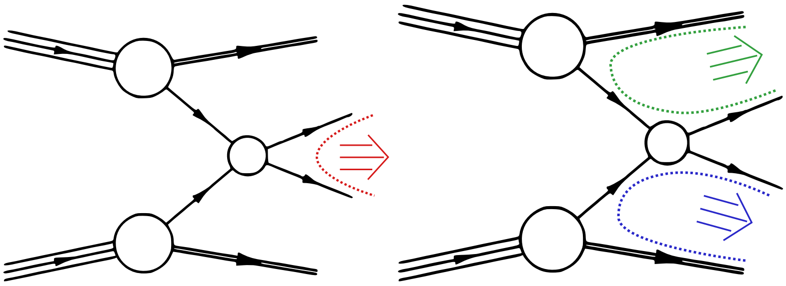

Figure 5.17: Colour connections for signals (left) and backgrounds (right).

flow [130]. It is a p_{T} weighted moment vector that tends to point toward the colour-connected partner of the jet's initiating quark, and it is defined as:

$$\vec{t} = \sum_{i \in \text{jet}} \frac{p_T^i \, |r_i|}{p_T^{\text{jet}}} \, \vec{r}_i$$

where $\vec{r}_i = (\Delta y_i, \Delta\varphi_i) = \vec{c}_i - \vec{J}$ with $\vec{J} = (y_J, \varphi_J)$ is the location of the jet and $\vec{c}_i$ is the position of a particle with transverse momentum p_T^i. For $\vec{r}_i$ and $\vec{t}$ calculation, only the charged particle flow objects used in the jet reconstruction are used.

The most effective way to use the pull vector is to calculate the *pull angle*, which is the angle between the pull vector and some other vector in the event, since its magnitude has no physical meaning. For signal, the pull vectors are expected to point towards each other, thus for each jet the pull angle is defined as the angular separation between the pull vectors and the connecting direction between the two jets in the y, φ 2D-plane. In background the pull vectors should point along the closest beam, so two additional angles can be defined. All four pull angles are sketched in Fig. 5.18.

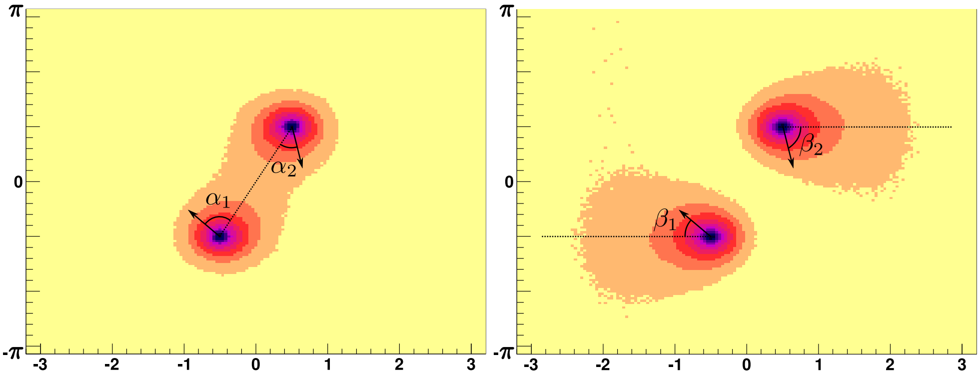

Figure 5.18: A single parton-level signal event showered millions of times (left) and a single parton-level background event with identical kinematics but different colour connections. The colour scale shows the average showered p_{T} density in the y, φ plane. The colour connections are shown with the thin lines, and examples of pull vectors (defined in a per-event basis) with arrows. The angles $\alpha_{1,2}$ and $\beta_{1,2}$ are illustrated as the angles of the pull vectors with respect to the signal and background-like connection lines.

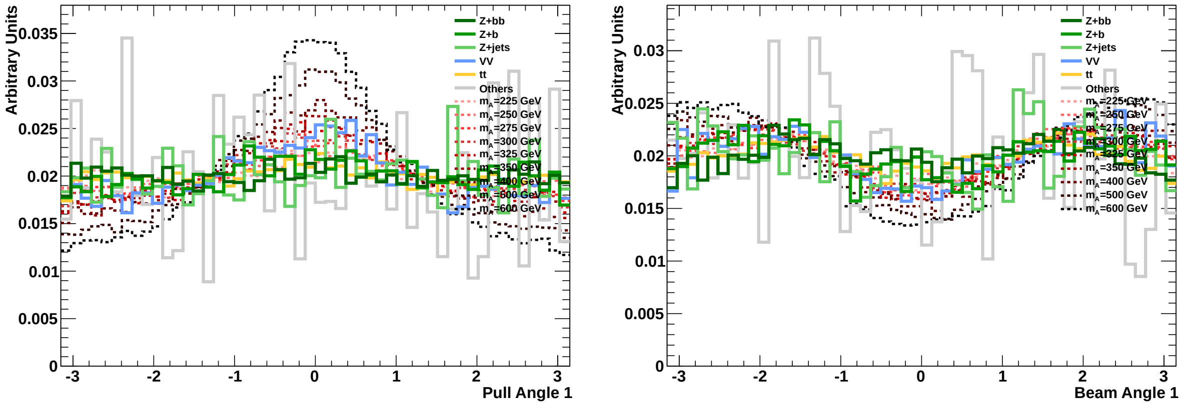

Figure 5.19: Colour angles for the leading jet: pull angle (left) and beam angle (right).

5.5 CUT-BASED ANALYSIS

Before starting a fully-optimized analysis with complex multivariate techniques, a simpler and more robust cut-based analysis has been developed. It is meant as a baseline benchmark useful to compare the gain obtained with every analysis improvements (kinematic fit, multivariate algorithms, two-dimensional fit), as well as performing an almost model-independent search for general $m_{\ell\ell bb}$ resonances, being less dependent on signal properties. Even if the results of the cut-based are not the best and definitive ones, they can give realistic results since the cut-based represents a self-consistent analysis, complete with the proper systematic uncertainties.

The cut-based analysis is developed on top of the selections described in Sec. 5.3, separately for three different signal mass ranges: low ($m_{\mathrm{A}} = 225, 250, 275\,\mathrm{GeV}$), intermediate ($m_{\mathrm{A}} = 300, 325, 350\,\mathrm{GeV}$), and high mass ($m_{\mathrm{A}} = 400, 500, 600\,\mathrm{GeV}$), due to the difference in the kinematic regimes. The cut-based analysis aims to optimize the significance of an hypothetical signal by applying cuts to the most discriminating variables, namely the boost of the boson candidates Z and h, the angular discriminator output $\mathcal{D}$, the ΔR between the two jet and lepton pairs and the S_{T}, all defined in Sec. 5.4. Starting from the signal region, several combinations of cuts are tested, counting the number of signal (S) and background (B) events within the $m_{\ell\ell bb}$ peak resolution after the selections. The figure of merit chosen is $Q = 2 \cdot (\sqrt{S+B} - \sqrt{B})$, which gives more appropriate significance values when S and B are small [131]. The resulting list of cuts is reported in Tab. 5.7; the distributions for $p_{\mathrm{T}}^{\mathrm{Z}}$, $\Delta R_{\mathrm{b\bar{b}}}$, and S_{T} in the signal region are shown in Fig. 5.20. The distributions of the invariant mass $m_{\ell\ell bb}$ for the three mass region for background and a signal with $\sigma \times \mathcal{B}r(\mathrm{A} \to \mathrm{Zh} \to \ell\ell\mathrm{b\bar{b}}) = 25\,\mathrm{fb}$ are shown in Fig. 5.21; Tab. 5.8 reports the expected number of signal and background events.

5.5.1 RESULTS OF THE CUT-BASED ANALYSIS

After the final selections, the distribution of the four body invariant mass $m_{\ell\ell bb}$ is exploited in the search for a possible signal. The mass templates for signals and backgrounds are fitted to the data, and a 95% CL upper limit is obtained using the $\mathrm{CL_s}$ method [105, 106]. The four main backgrounds, dibosons and other minor backgrounds are considered separately, and signal templates for mass values between the simulated mass points are obtained through the *histogram interpolation* technique, explained in App. A.2. All the normalization and shape systematic uncertainties, described in Sec. 5.8, are taken into account. As a cross-check, an exclusion limit with simple cut-and-count is extracted, giving comparable but $\sim 5 \div 30\%$ worse results. The 95% CL upper limits on the total $\sigma \times \mathcal{B}r$ (Fig. 5.22) are reported in Tab. 5.9.

	Low Mass	Intermediate Mass	High Mass
$m_{\ell\ell}$ (GeV)		$80 < m_{\ell\ell} < 100$	
CSV_1		Tight	
CSV_2		Loose	
m_{bb} (GeV)		$90 < m_{\mathrm{bb}} < 140$	
$E_{\mathrm{T}}^{\mathrm{miss}}$ (GeV)		< 60	
$\mathcal{D}$		> 0.30	
p_{T}^{Z}(GeV)	-	> 60	> 80
$\Delta R_{\mathrm{b\bar{b}}}$	< 3.25	< 3.25	< 2.50
S_{T}(GeV)	> 200	> 250	> 275

Table 5.7: Selections for the cut-based analysis, separately for each mass region.

	Low Mass Region		Intermediate Mass Region		High Mass Region	
Dataset	Events	Fraction	Events	Fraction	Events	Fraction
Data	2409 ± 49	-	1108 ± 33	-	565 ± 24	-
VV	98.0 ± 1.7	4%	50.3 ± 1.2	4.6%	33.4 ± 1.0	6%
W+jets	0 ± 0	0%	0 ± 0	0%	0 ± 0	0%
Single Top	7.3 ± 1.8	0.3%	3.4 ± 1.2	0.31%	1.8 ± 0.9	0.32%
tt (+V)	290 ± 3	12%	152.1 ± 2.5	14%	69.3 ± 1.7	12%
Z+bb	1260 ± 11	52%	537 ± 5	49%	281 ± 4	50%
Z+b	555 ± 7	23%	245 ± 3	22%	118.9 ± 2.2	21%
Z+jets	204 ± 5	8.4%	89.0 ± 1.9	8.2%	46.0 ± 1.2	8.2%
Multijet	0 ± 0	0%	0 ± 0	0%	0 ± 0	0%
SM Zh	22.38 ± 0.11	0.92%	14.73 ± 0.09	1.3%	10.72 ± 0.07	1.9%
All Bkg.	2437 ± 15	-	1092 ± 7	-	561 ± 5	-
$m_{\mathrm{A}} = 225\,\mathrm{GeV}$	29.06 ± 0.18	6.20%				
$m_{\mathrm{A}} = 250\,\mathrm{GeV}$	42.56 ± 0.21	9.05%				
$m_{\mathrm{A}} = 275\,\mathrm{GeV}$	51.10 ± 0.23	10.9%				
$m_{\mathrm{A}} = 300\,\mathrm{GeV}$			42.87 ± 0.21	9.15%		
$m_{\mathrm{A}} = 325\,\mathrm{GeV}$			53.58 ± 0.24	11.4%		
$m_{\mathrm{A}} = 350\,\mathrm{GeV}$			60.21 ± 0.25	12.8%		
$m_{\mathrm{A}} = 400\,\mathrm{GeV}$					63.04 ± 0.26	13.3%
$m_{\mathrm{A}} = 500\,\mathrm{GeV}$					77.19 ± 0.29	16.3%
$m_{\mathrm{A}} = 600\,\mathrm{GeV}$					82.21 ± 0.30	17.3%

Table 5.8: Expected and observed events for the Low, Intermediate and High Mass cut-based selections. Together with the number of expected events also the event fractions are reported for backgrounds, while for signal the selection efficiency. The signal yield is referred to a $\sigma \times \mathcal{B}r(\mathrm{A} \to \mathrm{Zh} \to \ell\ell\mathrm{b\bar{b}}) = 25\,\mathrm{fb}$.

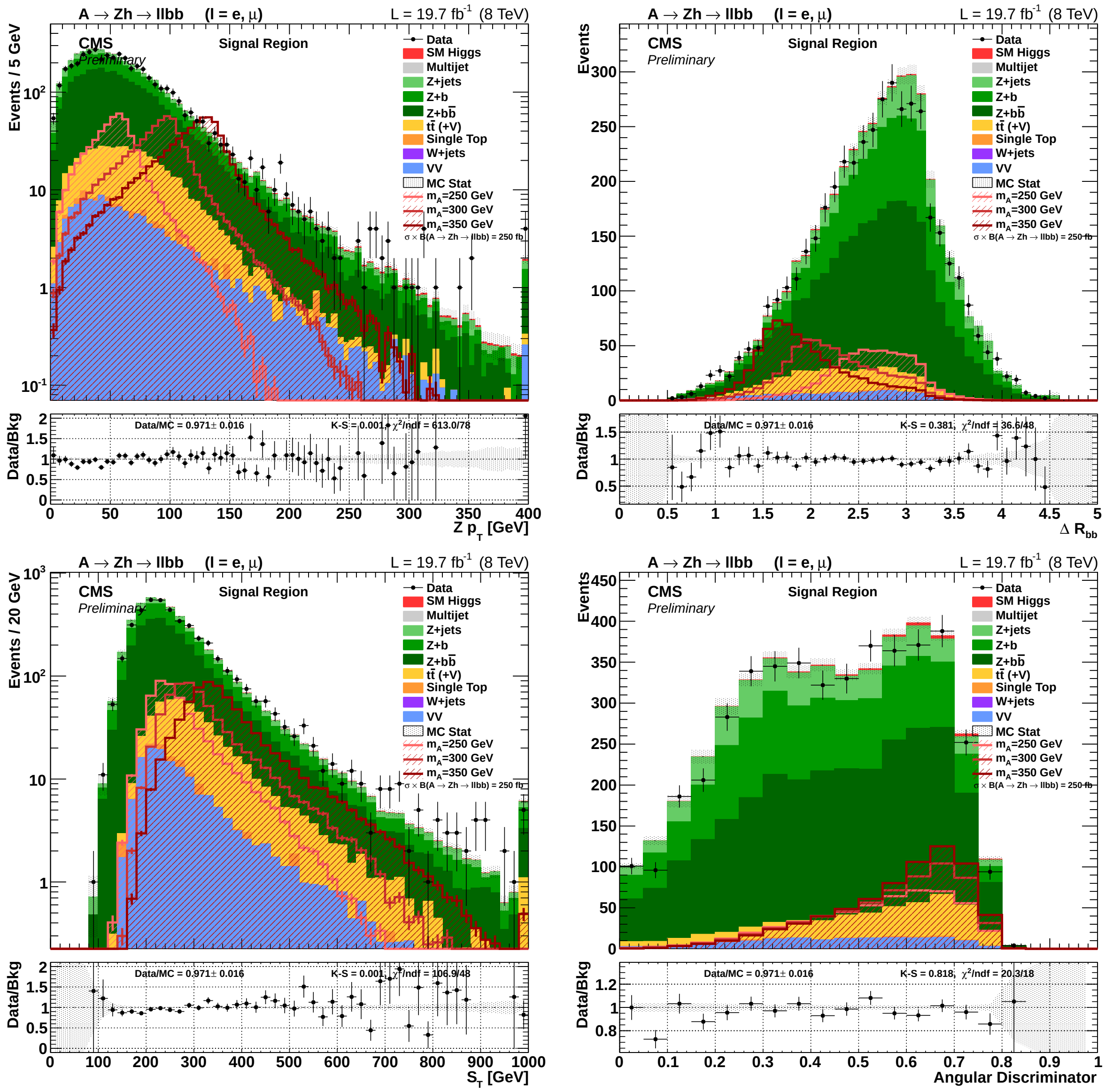

Figure 5.20: Distribution for p_T^Z (top left), $\Delta R_{b\bar{b}}$ (top right), S_T (bottom left), and the Angular Discriminator (bottom right) for signal and backgrounds in the signal region.

m_A (GeV)	225	250	275	300	325	350	400	500	600
Observed (fb)	24.0	39.4	36.3	41.3	33.7	17.8	5.06	7.38	3.55
Expected (fb)	47.4	45.4	38.1	29.4	21.7	16.5	10.0	5.23	3.27
Expected -1σ (fb)	33.8	32.5	27.3	21.0	15.5	11.7	7.09	3.65	2.23
Expected -2σ (fb)	25.3	24.4	20.5	15.7	11.6	8.73	5.27	2.69	1.62
Expected $+1\sigma$ (fb)	67.2	64.0	53.8	41.6	30.8	23.4	14.4	7.63	4.85
Expected $+2\sigma$ (fb)	91.5	86.8	73.0	56.7	41.9	32.1	20.0	10.7	7.02

Table 5.9: Expected upper limit for the cut-based analysis.

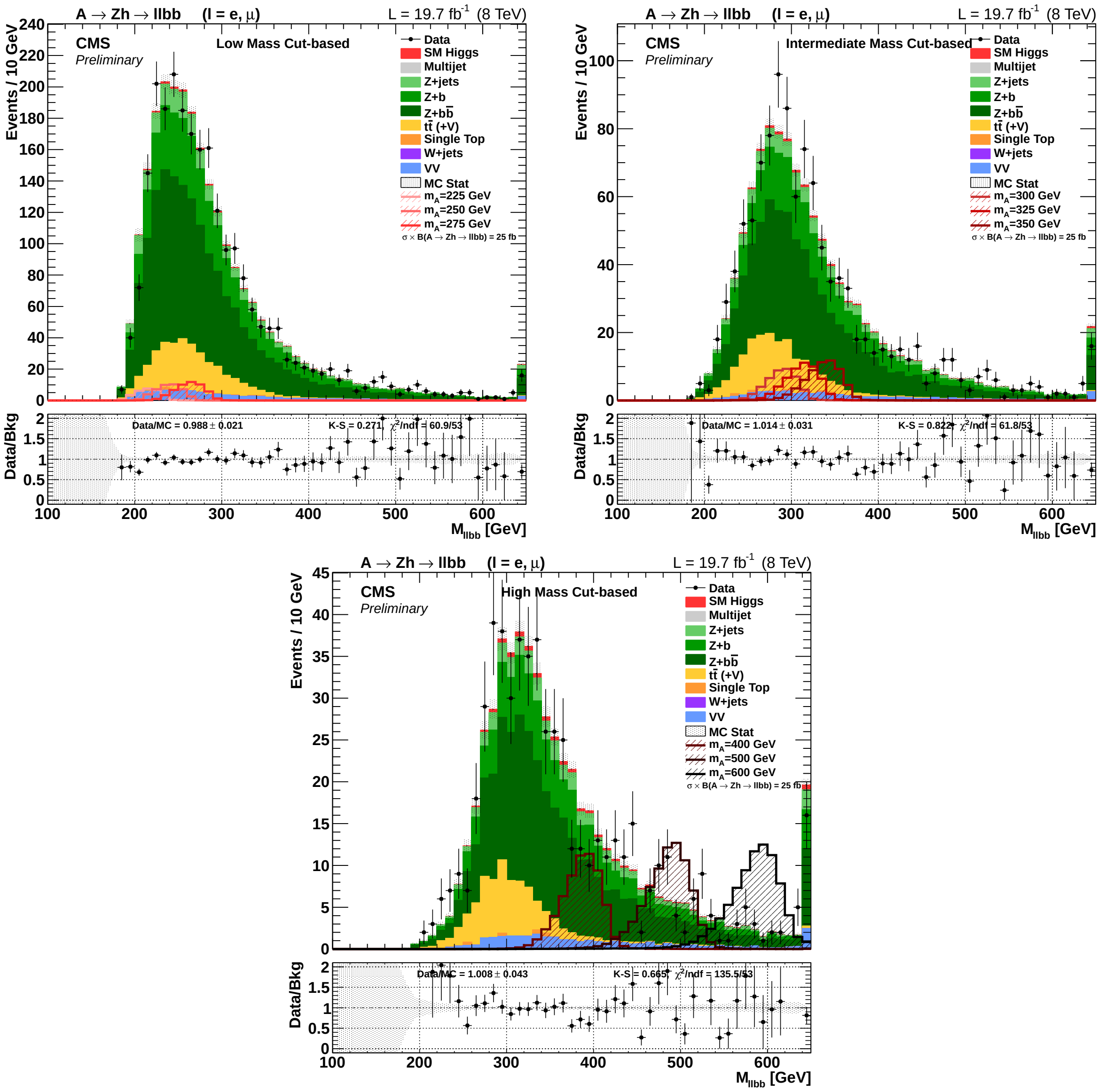

Figure 5.21: Four-body invariant mass $m_{\ell\ell bb}$ for the low (top left), intermediate (top right), and high mass region (bottom) after the final cut-based analysis selections.

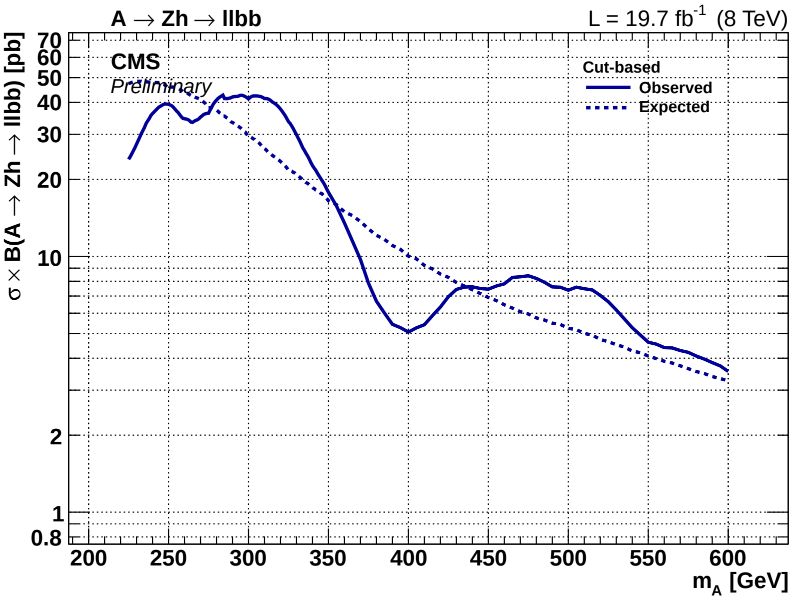

Figure 5.22: Expected and observed upper limit with the cut-based analysis.

5.6 KINEMATIC FIT

This search has a final state with two resonances of known mass, Z $\to \ell\ell$ and h $\to$ b$\bar{\text{b}}$. The $m_{\ell\ell bb}$ resolution is dominated by the dijet resolution, as leptons are better measured than jets. An important feature of the signal is that the two b jets originate from the decay of the h boson, whose mass is known with better precision than provided by the b$\bar{\text{b}}$ mass resolution [33]. A *kinematic fit*, performed with the TKINFITTER software [132] based on Lagrange multipliers, varies the jet p_{T}, η, and φ values within their resolution, to constrain the dijet invariant mass to m_{h} =125 GeV.

The p_{T} resolution of the two jets $\Delta p_{\text{T}}/p_{\text{T}}$, shown in Fig. 5.23, is studied in a MC sample and parametrized separately in the barrel and endcaps regions with a second degree polynomial. The procedure for jet η and φ is analogous. The kinematic fit leaves the jet invariant mass unchanged, and the jet energy is varied accordingly. The resolution is computed using the 4-momentum of the generated particles (with the exclusion of neutrinos, if present) clustered with same jet algorithms of reconstructed jets.

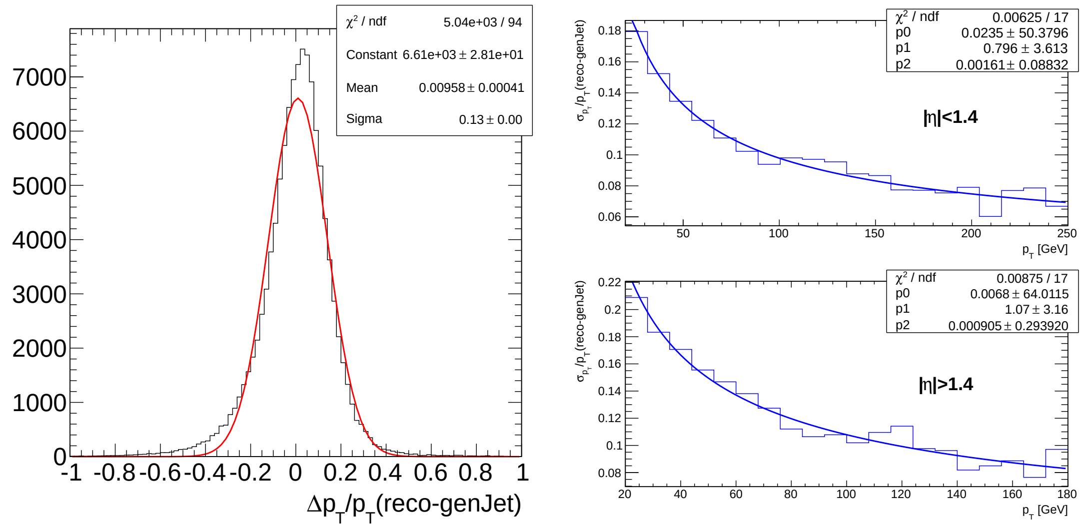

Figure 5.23: Relative resolution on p_{T} for b jets ($\Delta p_{\text{T}}/p_{\text{T}}$) (left) and its dependence on p_{T} for barrel (left-top) and endcap region (left-bottom) with functional fit

The goodness of the kinematic fit is tested by defining a χ^2 of the constrained vs unconstrained jets 4-momenta:

$$\chi^2 = \sum_{\substack{x=p_T,\eta,\phi \\ i=jet_1,jet_2}} \left(\frac{x_i^{\text{w/o KinFit}} - x_i^{\text{w/ KinFit}}}{\sigma_{x_i}} \right)^2 .$$

The χ^2 distribution is shown in Fig. 5.26, for data and background and signal MC samples, together with the unconstrained b$\bar{\text{b}}$ invariant mass as a function of the χ^2 of the kinematic fit. The χ^2 is larger for events with an unconstrained b$\bar{\text{b}}$ invariant mass far from the nominal m_{h}. Since the dijet invariant mass does not carry any further information after the kinematic fit, the χ^2 is used in subsequent steps of the analysis as a discriminant in place of m_{bb}.

The impact of the kinematic fit on the four-body invariant mass $m_{\ell\ell bb}$ before and after the kinematic fit are compared in Fig. 5.25 and 5.28. The results of a Gaussian fit on the bulk of the

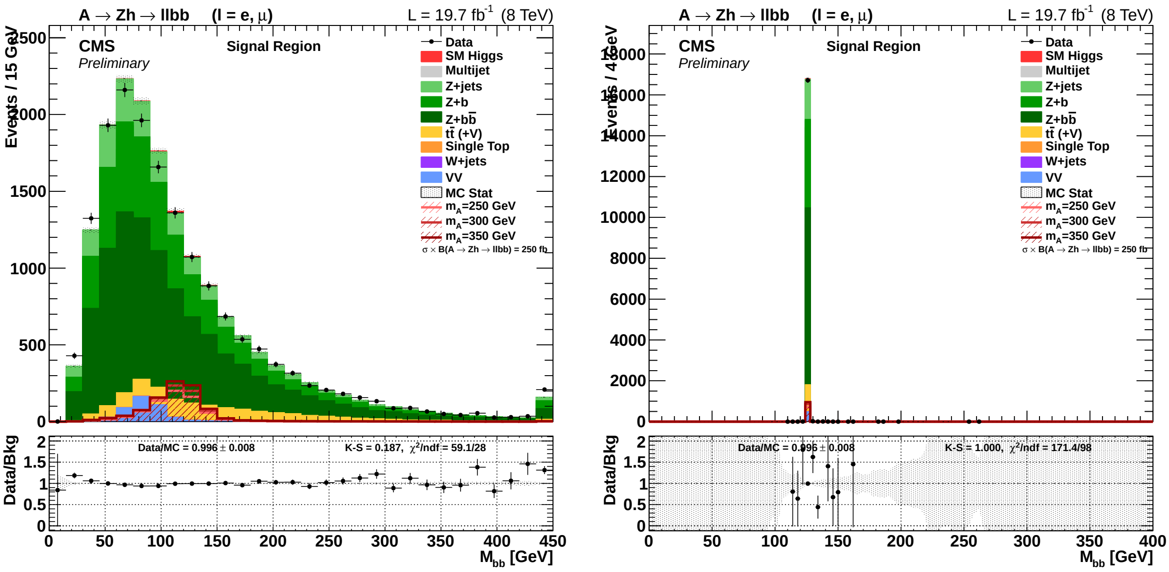

Figure 5.24: Invariant mass of the two b jets before (left) and after (right) the kinematic fit.

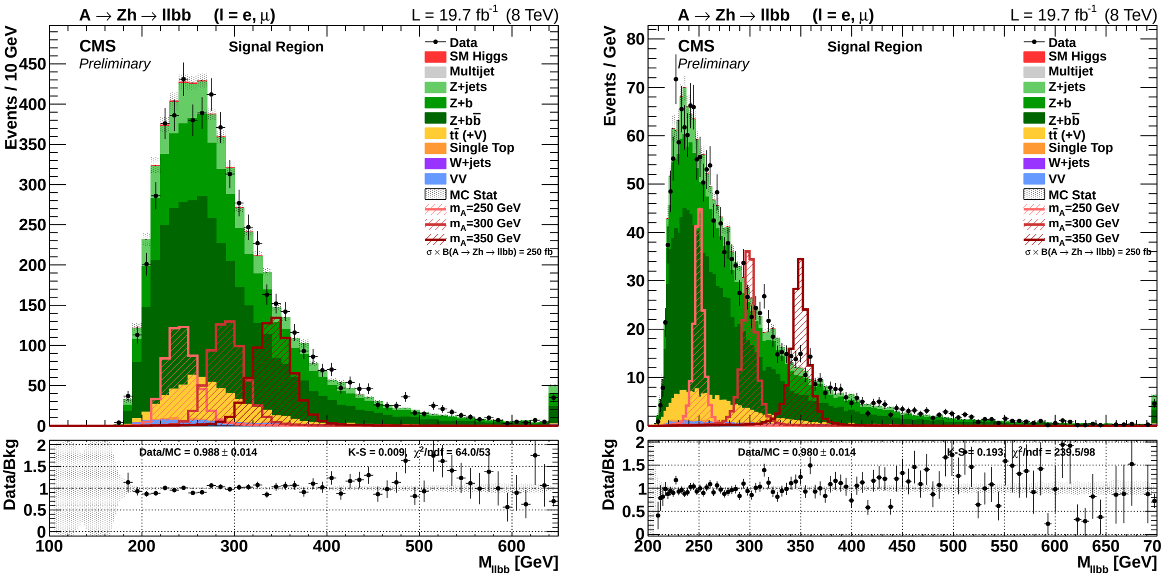

Figure 5.25: Four-body invariant mass $m_{bb\ell\ell}$ for backgrounds and signals (enhanced), for three mass points (m_{A} =250, 300, and 350 GeV) before (left) and after (right) the kinematic fit on bb system.

signal distribution are reported in Tab. 5.10. The kinematic fit improves the relative four-body invariant mass resolution from 6.3 and 4.0% to 1.2 and 1.9%, respectively, for the smallest and largest values of m_{A}, reducing thereby the mass offset, and centering the peaks around their nominal values. Although both the background and signal $m_{\ell\ell bb}$ distributions are modified by the kinematic fit, the signal significance in a mass window close to the peak increases by a factor 2 at lowest mass and by 34% at highest mass. The resulting jet three-momenta redefine all the kinematic variables in the event.

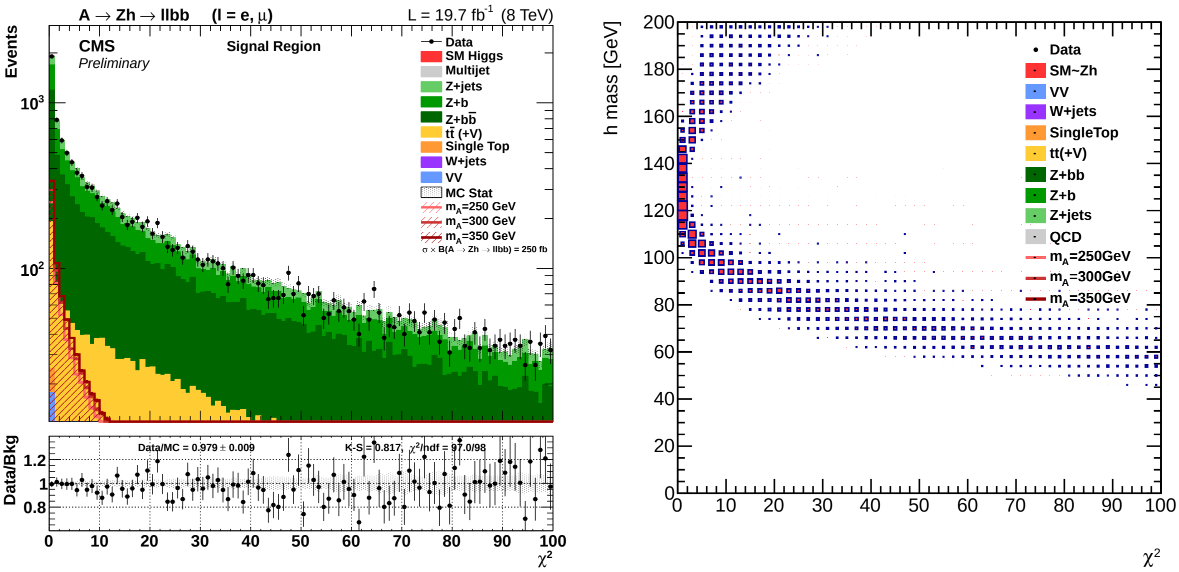

Figure 5.26: Distribution of χ^2 of kinematic fit (left) and m_{bb} unconstrained vs χ^2 (right).

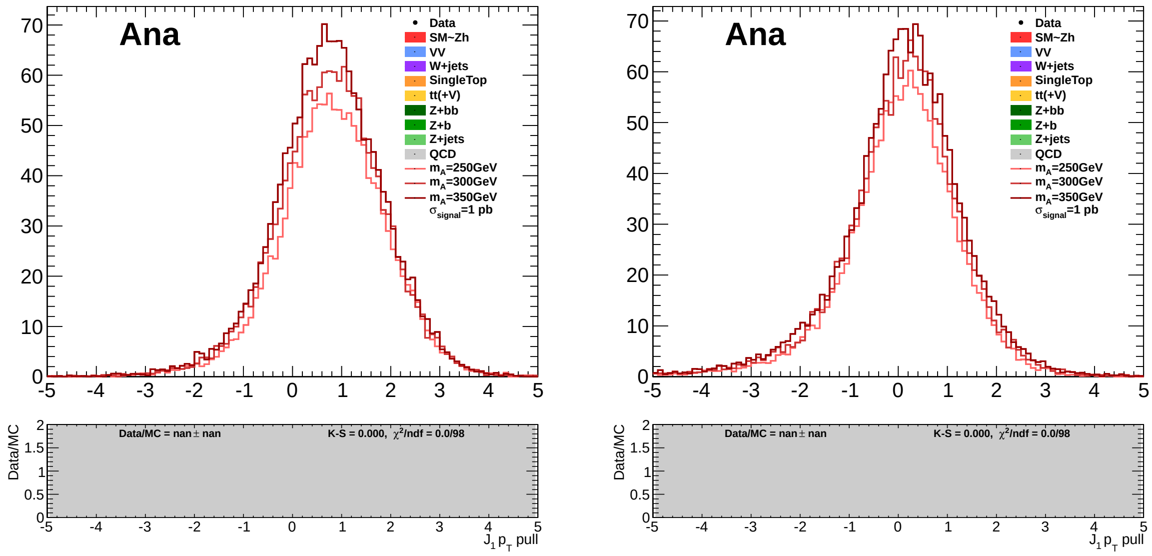

Figure 5.27: Pulls of jet p_T against the generated jet p_T (left) and generated parton p_T (right), for signal. For the signals (red lines), the generated jet pull is off-center, while the one computed against the parton is well centered, confirming that the kinematic fit recovers the original b quark p_T. The corresponding pulls for background are not meaningful, given the use of a wrong hypothesis for these samples, and are not shown.

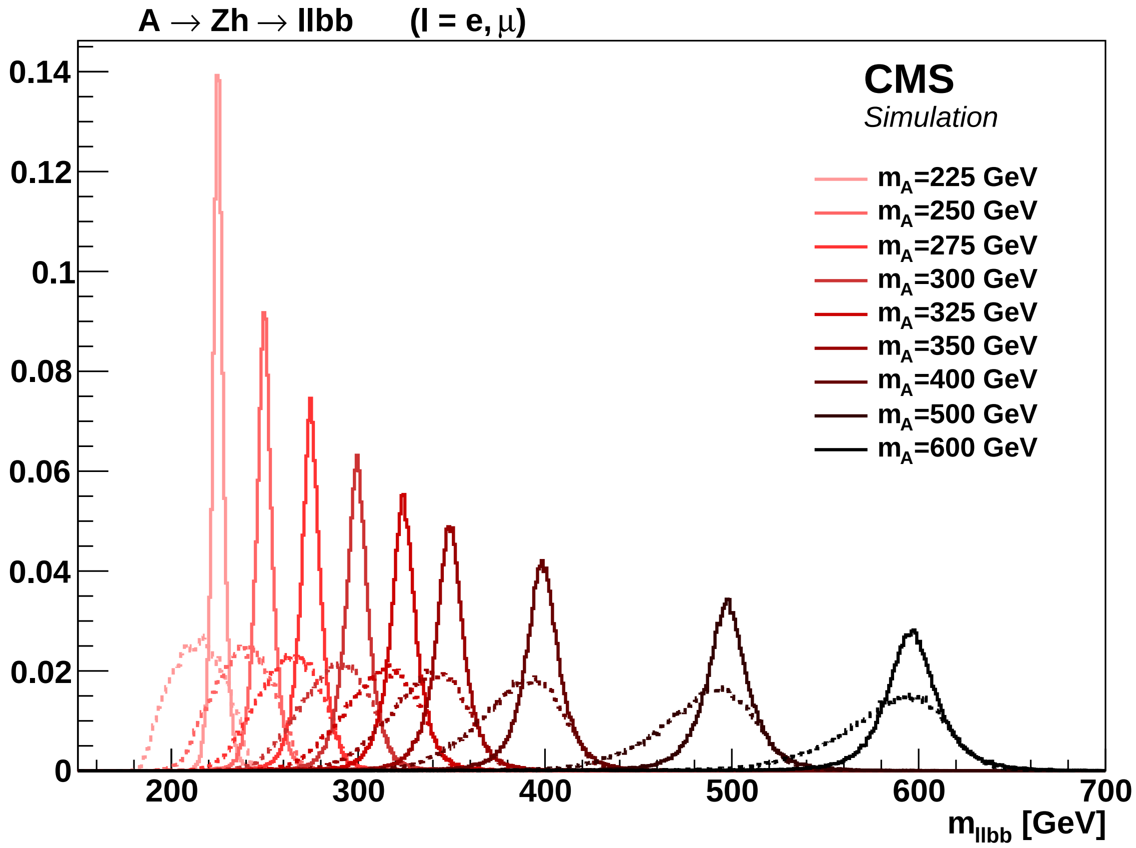

Figure 5.28: Simulated distributions for $m_{\ell\ell bb}$ before (dotted lines) and after the kinematic fits (solid lines). Histograms are normalized to unit area.

	m_A (GeV)	225	250	275	300	325	350	400	500	600
No Fit	Mean (GeV)	212.9	237.5	262.7	287.6	312.8	338.1	388.4	489.5	590.9
	Width (GeV)	13.53	14.59	15.53	16.31	17.17	17.78	18.96	21.41	23.41
	Res (%)	6.35	6.14	5.91	5.67	5.49	5.26	4.88	4.37	3.96
Kin Fit	Mean (GeV)	225.0	249.8	274.6	299.5	324.1	349.1	398.5	497.4	595.9
	Width (GeV)	2.66	3.64	4.45	4.99	5.81	6.17	7.29	8.92	11.47
	Res (%)	1.18	1.46	1.62	1.67	1.79	1.77	1.83	1.79	1.93

Table 5.10: Results of the Gaussian fits to the $m_{\ell\ell bb}$ peaks, with and without the kinematic fit.

5.6.1 RESULTS WITH THE KINEMATIC FIT

The kinematic fit can improve substantially the sensitivity, thanks to the better resolution of the $m_{\ell\ell bb}$ mass peak. The cut-based analysis, described in Sec. 5.5 is repeated using the b jets momenta after the kinematic fit instead of the original ones. Quantities derived from the b jet, as the ΔR_{bb}, S_{T}, H_{T}, $E_{\mathrm{T}}^{\mathrm{miss}}$, are recalculated as well. Since the m_{bb} selection is not effective anymore, it is replaced by a $\chi^2 < 10$ cut.

The distributions for $m_{\ell\ell bb}$ after the final selections in the three mass regions (low, intermediate, and high), together with a signal with $\sigma \times \mathcal{B}r(\mathrm{A} \to \mathrm{Zh} \to \ell\ell\mathrm{b}\bar{\mathrm{b}}) = 25\,\mathrm{fb}$, are shown in Fig. 5.30.

Upper limits at 95% CL resulting from a fit on the four-body invariant mass distributions after the kinematic fit are presented in Fig. 5.29, and numerical values are reported in Tab. 5.11. The comparison of the results without kinematic fit shows that the improvement on the expected upper limit, as well as on the expected sensitivity, goes from $\lesssim 2$ at low mass, to $\sim 10\%$ at high mass, where the $m_{\ell\ell bb}$ resolution is dominated by p_{T} of the $\mathrm{b}\bar{\mathrm{b}}$ system, and not by its invariant mass.

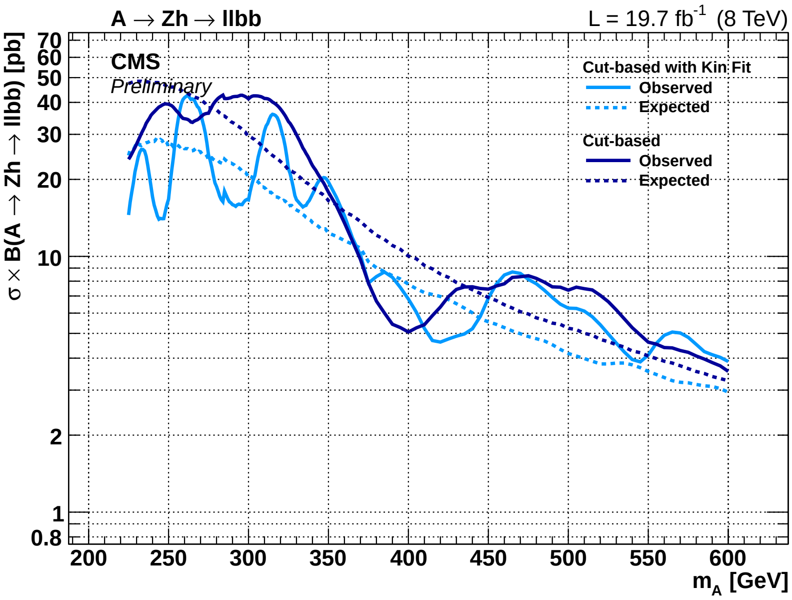

Figure 5.29: Expected and observed 95% CL upper limits with cut-based analysis using the kinematic fit, compared to the analogous result without the fit. The sensitivity improvement is clearly visible in the low mass region.

m_{A} (GeV)	225	250	275	300	325	350	400	500	600
with the kinematic fit									
Observed (fb)	14.5	16.7	24.9	16.6	22	19.5	6.81	6.27	3.88
Expected (fb)	24.8	26.8	23.9	20.3	15.7	12.3	7.78	4.17	2.95
without the kinematic fit									
Observed (fb)	24	39.4	36.3	41.3	33.7	17.8	5.06	7.38	3.55
Expected (fb)	47.4	45.4	38.1	29.4	21.7	16.5	10	5.23	3.27

Table 5.11: Expected upper limit on $\sigma \times \mathcal{B}r(\mathrm{A} \to \mathrm{Zh} \to \ell\ell\mathrm{b}\bar{\mathrm{b}})$. Previous results without the kinematic fit (Tab. 5.9) are reported in the bottom for comparison.

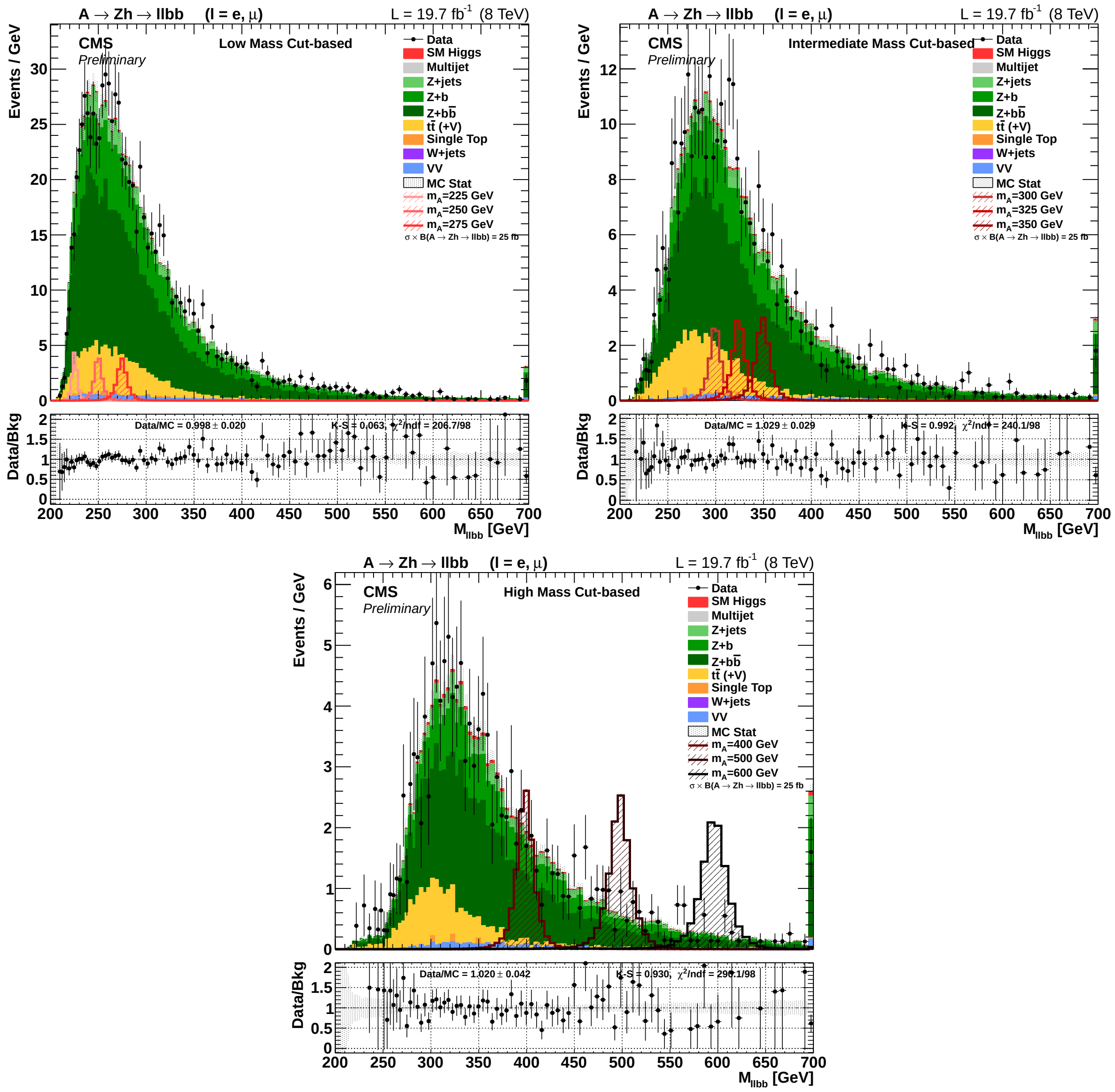

Figure 5.30: Four-body invariant mass $m_{\ell\ell bb}$ with the kinematic fit, for the low (top left), intermediate (top right), and high mass region (bottom) after the final selections, including the $\chi^2 < 10$ cut.

5.7 MULTIVARIATE ANALYSIS

The sensitivity of the analysis can be further improved by using multivariate (MVA) techniques, fully exploiting the correlations between discriminating variables in signal and background events. The selected MVA algorithm is a *Boosted Decision Tree* (BDT) [133], implemented in the TMVA package [134].

The MVA algorithm is trained and applied to events passing the signal region selections,after relaxing the cuts on the missing energy ($E_{\mathrm{T}}^{\mathrm{miss}} < 100\,\mathrm{GeV}$) and the dilepton invariant mass ($90 < m_{\mathrm{bb}} < 140\,\mathrm{GeV}$). Even if BDTs offer a relatively fast and accurate response, with similar performances than the most performing Neural Networks, their strongest point is a relative robustness to the most common flaws of multivariate methods, such as overtraining, fine parameters tuning and performance deterioration with small statistical samples. However, a systematic optimization is still required.

The first step is the variable choice. More than 40 different variables can be defined to describe the signal and background process (see Sec. 5.4). However, only a subsample of these is able to carry enough information, the others being highly correlated or weakly discriminating. Starting from the initial nine considered for the cut-based analysis (Sec. 5.5), variables are added one by one starting from the most promising ones, and checking the discriminating power of the BDT. Unlike the cut-based analysis, where the signal efficiency and background rejection can be evaluated in a single working point, MVA algorithms offer a continuous output, which results in many combinations of efficiency/rejection depending on the cut on the discriminator. These points form a curve (the ROC curve) in signal efficiency vs. background rejection plane, and the area under the curve (AUC) is proportional to the discriminating power of the multivariate method, as it is thus considered as a figure of merit for these tests.

The final 16 variable list are reported in Tab. 5.12 and shown in Fig. 5.32 with a brief description. It is also verified that assuming the same list of variables for all mass points does not negatively affect the performance. In Fig. 5.31 the correlation matrix for signal and background are shown.

Variable	Description
CSV_1	highest CSV value between the two jets
CSV_2	second-highest CSV value
$m_{\ell\ell}$	invariant mass of the lepton pair
p_{T}^{Z}	p_{T} of the Z candidate
p_{T}^{h}	p_{T} of the dijet pair (h candidate)
ΔR_{bb}	angular separation of the two jets in the $\eta - \varphi$ space
τ_{bb}	twist angle between the two jets $\tau \equiv \tan^{-1} \Delta\varphi/\Delta\eta$ (Sec. 5.4)
$E_{\mathrm{T}}^{\mathrm{miss}}$ sign	Missing energy significance (Sec. 3.2.12)
χ^2	χ^2 of the kinematic fit
S_{T}	scalar sum of the p_{T} of jets, leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$ in the event
nJets	number of jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ in the event
Centrality	Centrality of the four decay products in the A rest frame (Sec. 5.4)
Aplanarity	event Aplanarity calculated with the four decay products (Sec. 5.4)
$\cos\theta^*$	polar angle between the direction of flight of the Z boson and of the beam in the A rest frame
$\cos\theta_1$	Z decay angle w.r.t. its flight direction in the Z rest frame
Pull Angle	angle of the pull vector of the highest-p_{T} jet (Sec. 5.4)

Table 5.12: List and description of all the variables included in the BDTs.

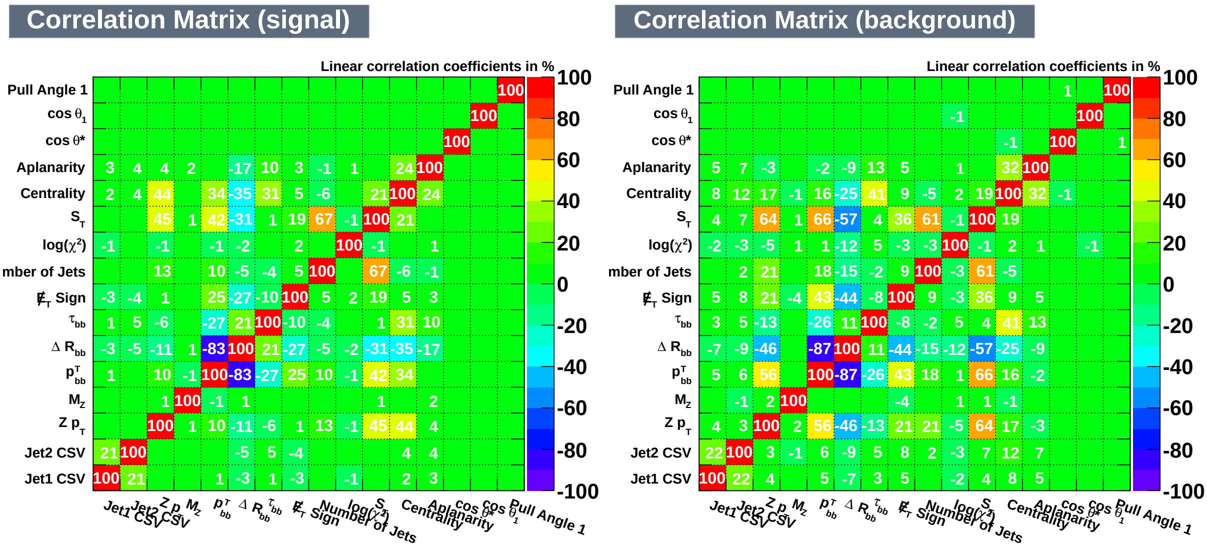

Figure 5.31: Correlation matrix for signal (left) and background (right).

A custom optimization to the specific case of the most relevant BDT parameters often leads to improvements in the BDT general behavior and discrimination power. Only four of the most important are considered:

Number of trees: a large "forest" offers better discrimination power, but an excessively large number of trees may lead up to overtraining the BDT. Ten different values are tested, ranging from 200 to 6000 trees (default is 800). While the AUC reaches saturation with a very high number of trees (> 4000), the training is stopped for a value of 1000 trees, since the error fraction vs. the tree numbers begins to show some instability if the forest is allowed to grow beyond $1100 - 1200$ trees.

Depth: analogously to the previous case, performance reaches a plateau for a high depth ($\geq 6 - 8$, default is 3). Growing trees beyond this value has the double drawback of reducing performance and of increasing training and testing time, due to the complexity of the BDT. Furthermore, it is found that letting the trees grow in depth leads very quickly to overtraining; the first signs are clearly visible for depths > 6, and can be spotted by looking for a disagreement between the training and testing outputs for signal and background. As a consequence, the maximum tree depth is limited to 5. With this setting, no signs of overtraining are present, and training and testing samples are in good agreement.

Minimum number of events: the minimum number of events in each leaf is also varied, but the BDT final result has very little dependence on this parameter. 5 different values are considered, and the best is taken (200 events).

Number of cuts: the number of steps during node cut optimization is tested as well, but as in the previous case, this parameter does not really affect the BDT power. It is verified that the default value (20) is confirmed the best choice.

Before looking at data in the signal region, the BDTs are tested in the control regions, specifically the $t\bar{t}$ and Z +$b\bar{b}$ control region (Fig. 5.33). A fair agreement is observed, especially in the region closer to one, where signal is expected. Small trends can be observed in the background region close to 0, but shape uncertainties (described in Sec. 4.6) cover most of the discrepancy.

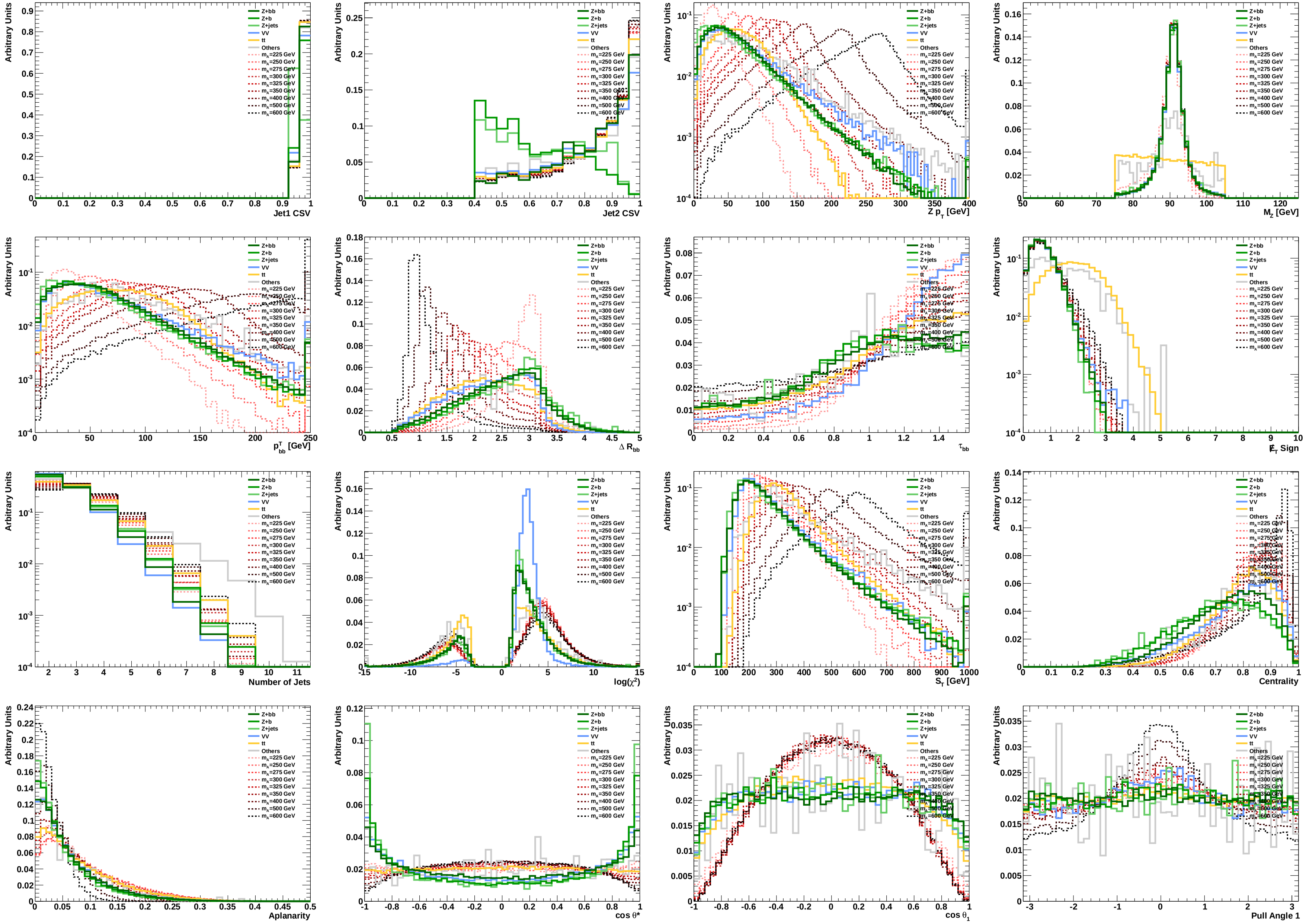

Figure 5.32: BDT variables distributions normalized to unity for the four main backgrounds (Z + *jets*, Z + *b*, Z + *bb*, $t\bar{t}$) and the nine signal samples.

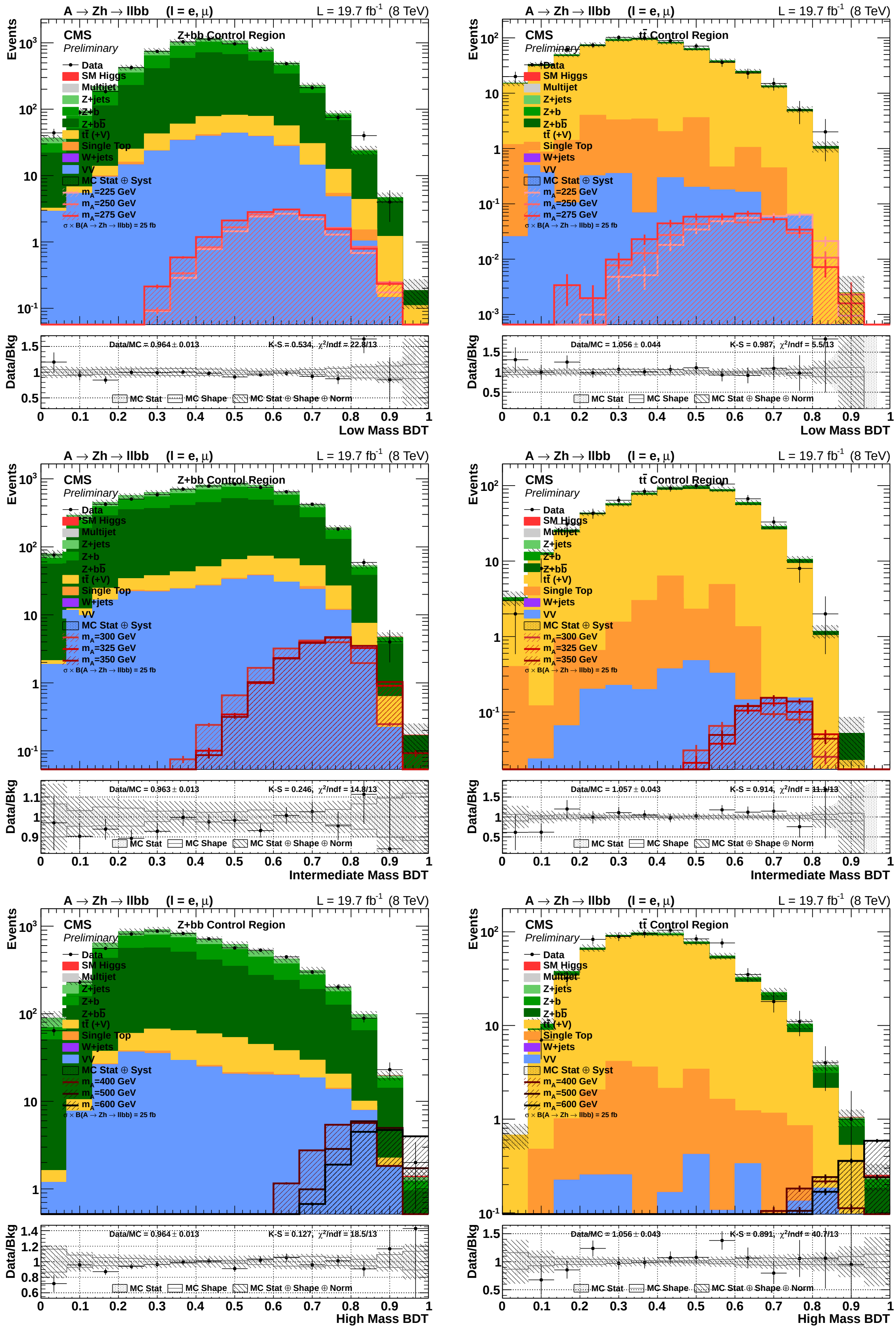

Figure 5.33: Discriminator output for the low (top), intermediate (middle) and high mass (bottom) BDT in the $Z + b\bar{b}$ (left) and $t\bar{t}$ (right) control regions.

5.8 SYSTEMATIC UNCERTAINTIES

5.8.1 BACKGROUND NORMALIZATION

An important part of the analysis resides in the data-driven determination of the four most important backgrounds, performed in appropriate control regions. The statistical uncertainty returned by the control regions fit due to the limited population is a primary source of uncertainty. However, the control regions could be quite far in the variables phase-space with respect to the signal region, and the extrapolation from the control to the signal region can be thus affected by systematic uncertainties, that should be evaluated and taken into account.

The uncertainty due to the extrapolation to the signal region is evaluated on data and simulation yields in a number of additional regions, statistically independent from the signal region, obtained by altering the selections used to define the four control regions (m_{bb}, $E_{\mathrm{T}}^{\mathrm{miss}}$, $\mathrm{CSV}_{1,2}$). Since a large amount of combinations are possible, only the 20 regions closest to the signal region are considered (listed in Tab. 5.13), if they fulfill the requirement of at least one veto on m_{bb} or $E_{\mathrm{T}}^{\mathrm{miss}}$ to avoid signal contamination.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
CSV_1	T	T	T	M	T	T	T	M	T	T	T	M	M	M	T	T	T	M	M	M
CSV_2	L	M	-	M	L	M	-	M	L	M	-	L	M	-	L	M	-	L	M	-
m_{bb}	<	<	<	<	>	>	>	>	<	<	<	<	<	<	>	>	>	>	>	>
$E_{\mathrm{T}}^{\mathrm{miss}}$	<	<	<	<	<	<	<	<	>	>	>	>	>	>	>	>	>	>	>	>

Table 5.13: Selection for the 20 auxiliary regions used to determine the extrapolation uncertainty of the four main backgrounds. Different thresholds are applied to jet CSV (T is for the tight working point, M for the medium, L for the loose), dijet invariant mass (< stands for $m_{\mathrm{bb}} < 90\,\mathrm{GeV}$, > for $m_{\mathrm{bb}} > 140\,\mathrm{GeV}$, since the region between could contain a fraction of signal) and $E_{\mathrm{T}}^{\mathrm{miss}}$ (> or < 40 GeV)

The total normalization uncertainty is evaluated through a simultaneous likelihood fit on the data and simulation yield in the auxiliary regions, fixing the four background normalization with their scale factors, but introducing four additional nuisance parameters σ_k with Gaussian constraints (one for each process k=Z +jets, Z +b, Z +b$\bar{\mathrm{b}}$, t$\bar{\mathrm{t}}$). The other backgrounds (diboson and single top) are kept fixed. The likelihood formula is:

$$\mathcal{L} = \prod_i \frac{e^{-\mu}\mu^n}{n!} \cdot \prod_k e^{\sigma_k^2}$$

$$\mu = N_0 + \sum_k (1 + \sigma_k) \cdot N_k$$

where N is the number of events for each k sample (N_0 is for the yield of non-data-driven backgrounds) in the i-th auxiliary region. The σ_k parameters returned by the fit should correspond to the relative uncertainty for each background; results and their errors are reported in Tab. 5.14. Z +jets and Z +b errors are larger than others, since they are evaluated in control regions far away the signal region, and both differ from the latter for the b tagging application. It is verified that these results do not change significantly altering the definition of the auxiliary regions. As a cross-check, the RMS of the data over MC ratios with respect to the mean value (Fig. 5.34) can give an indication of the total uncertainty on the backgrounds: a value of 3.78% is found, while a mean value 1.016 ± 0.004 gives some additional confidence about the scale factors.

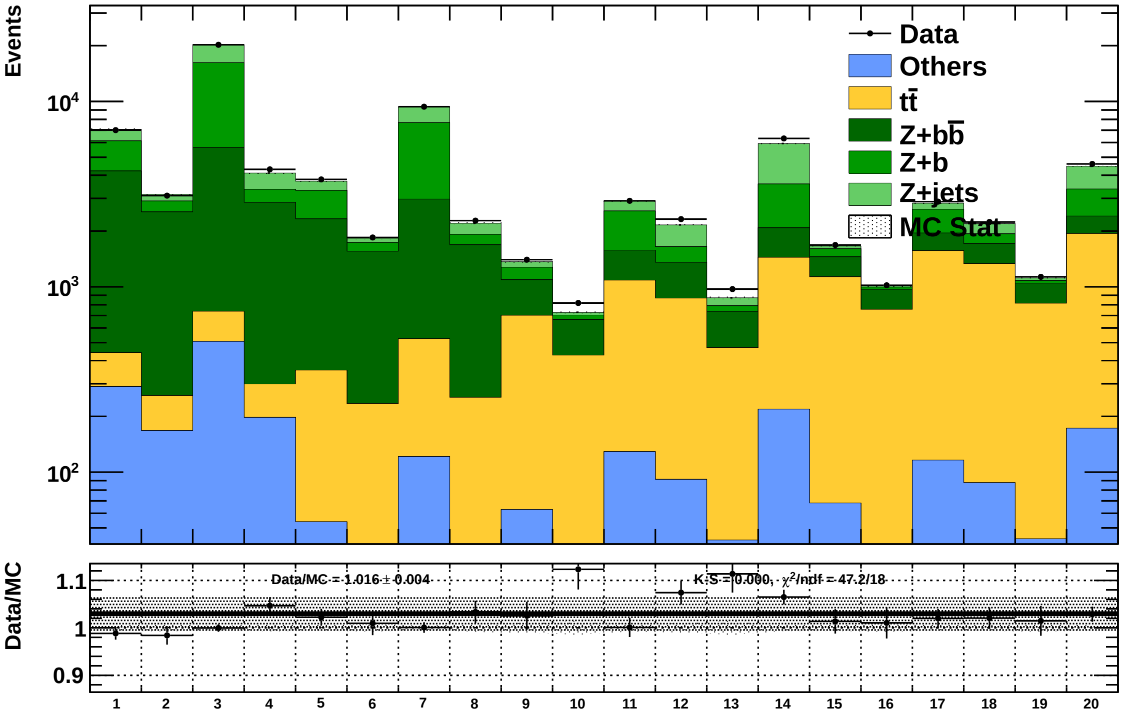

Figure 5.34: Monte Carlo and Data yield in the 20 auxiliary regions selected for evaluation of the extrapolation systematic uncertainties.

Z+jets	Z+b	Z+bb	$t\bar{t}$
$13.0 \pm 3.9\%$	$11.8 \pm 3.1\%$	$2.15 \pm 1.58\%$	$6.24 \pm 1.61\%$

Table 5.14: Results from the likelihood fit for each background extrapolating from control regions to signal region.

The other backgrounds have a 15% uncertainty on the calculated or measured cross section, except the Z Z sample which has a smaller 9.2% measured uncertainty (see Sec. 5.2). For backgrounds normalized to calculated cross sections, a further 2.6% uncertainty is assigned due to luminosity [126].

5.8.2 LEPTONS

Systematic uncertainties for momentum scale and resolution are evaluated by rescaling the corresponding quantities within their uncertainty and counting the yield difference in the signal region. The normalization is relevant only for the signal samples and for the other backgrounds than Z +jets, Z +b, Z +$b\bar{b}$, $t\bar{t}$, the latter being normalized on data. After electron regression and smearing (described in Sec. 3.2.5) and muon calibration (Sec. 3.2.10), the energy scale uncertainty is estimated to be 0.1% and 0.4% for electrons in the barrel and endcaps, respectively [135]. For muons, the momentum scale is a flat 0.1% in the considered p_T range [88]. After lepton calibrations, the energy and momentum resolution in simulation show good agreement with data, and the residual effects are negligible for the analysis.

Identification and isolation systematic uncertainties are evaluated by moving up and down the scale factors for identification and isolation by their errors, determined from specific studies on the Z peak; this effect is small $(0.1 - 0.3\%)$. Trigger uncertainty is conservatively taken as 2.0% for both double-muon and double-electron triggers. Since lepton systematic uncertainties are relatively small, a round number of 2.5% is taken regardless of the flavour. No shape uncertainty is considered for lepton energy scale and resolution.

5.8.3 JETS

The main jet uncertainty sources are the jet energy scale (JES) and resolution (JER), and they are evaluated similarly to the lepton uncertainties. The impact of b tagging uncertainty is instead evaluated by moving the CSV scale factors in the reshaping procedure (described in Sec. 3.2.8). An average systematic uncertainty of 6% per b jet, 12% per c jet, and 15% per fake tag (light quarks and gluons) is used; the proper value is assigned for each jet as a function of its p_T and η [86].

While for cut-based analyses only uncertainties on normalization are considered, for BDT and 2D analyses also shape variations due to JES, JER and CSV are taken into account. The procedure is the following: after the variation of the selected uncertainty source, a new kinematic fit is performed with the varied four-momenta, resulting in a different χ^2 value. The composite candidates Z and h are then recalculated, and the BDT output discriminant is recomputed from the shifted variables. The difference with respect the original output is taken as systematic uncertainty for that specific source. The kinematic fit sets aside most of the effect of the JES and JER uncertainties, but second-order effects affect the $m_{\ell\ell bb}$ and the BDT distributions. Variations in m_A and BDT output are considered simultaneously, taking into account the full correlation between the two variables.

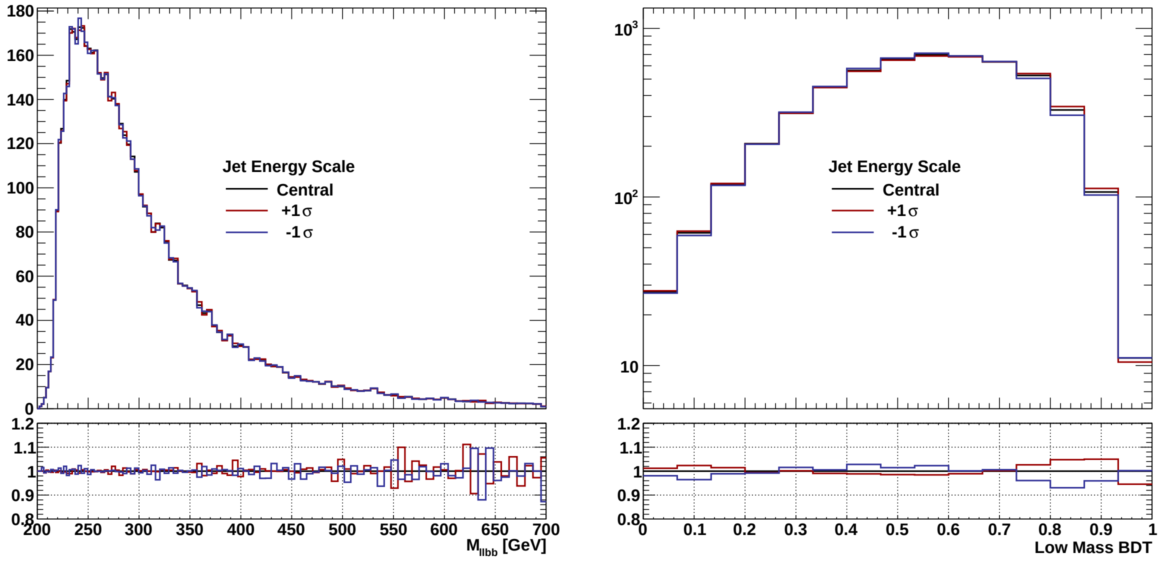

Figure 5.35: Projection of the jet energy scale shape variations on the 4-body invariant mass (left) and low mass BDT (right).

5.8.4 FACTORIZATION AND RENORMALIZATION SCALE

Uncertainties on the acceptance due to the truncated perturbative series, the factorization and renormalization scale used for the event generation are varied in a correlated way by a factor of 1/2 and 2 around the central value. For each shifted scale value Q', events are assigned the weight:

$$W^{\pm}_{Q^2}(Q,Q') = \frac{g(x_1,x_2,Q')}{g(x_1,x_2,Q)} \cdot \left(\frac{\alpha_s(Q')}{\alpha_s(Q)}\right)^n$$

where n is the number of extra partons in the process. Fig. 5.36 show the uncertainty band on m_A and the BDT for the Q^2 scale variations.

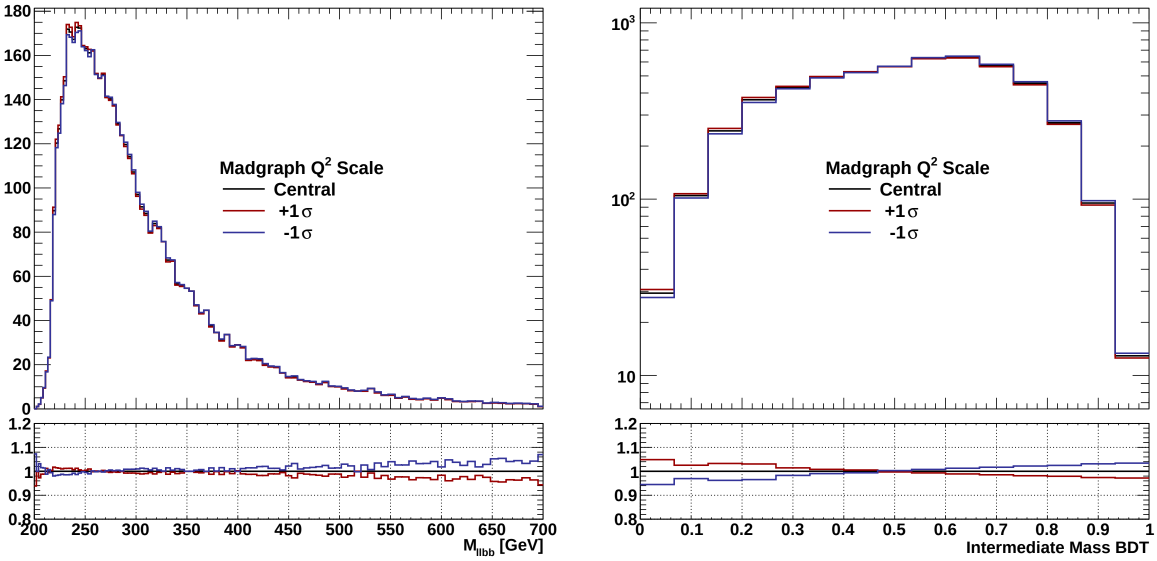

Figure 5.36: Projection of the Q^2 scale shape variations on the 4-body invariant mass (left) and intermediate Mass BDT (right).

5.8.5 MONTE CARLO MODELING

An additional systematic uncertainty has to be added due to the Monte Carlo mismodeling corrected by reweighting the centrality distribution (see Sec. 5.3.2 for details), which affects also $m_{\ell\ell bb}$ and the BDTs output. An uncertainty band is built by varying the scale factor weight from zero to two, defining a $\pm 1\sigma$ envelope centered around the corrected distribution. Fig. 5.37 shows the variation in the two-dimensional space of the two $m_{\ell\ell bb}$ and BDT variables.

5.8.6 MONTE CARLO STATISTICS

The effect of the limited background Monte Carlo statistics becomes non-negligible in the signal-enriched BDT bins, where only a small fraction of the simulated events are present. This effect is accounted by introducing additional coherent bin-by-bin shape variations for each process equal to the bin statistical uncertainty [136]. Since the bin number is very high in the 2D analysis, only the 30 most significant bins (ordered by the expected S/B) are used; adding more bins does not change the expected limit. Figure 5.38 shows the expected limit relative difference considering the Monte Carlo statistics, as a function of the number of bins considered. The larger impact is for low masses, due to the narrowness of the peaks, accounting for about 4% increase in the expected limit.

5.8.7 OTHER UNCERTAINTIES

The missing energy is a complex quantity, built from all the reconstructed physics objects, and it is consequently affected by all the previous systematic sources, plus the contribution of particles and energy deposits that are not clustered into jets. The effect on the $E_{\mathrm{T}}^{\mathrm{miss}}$ relative to

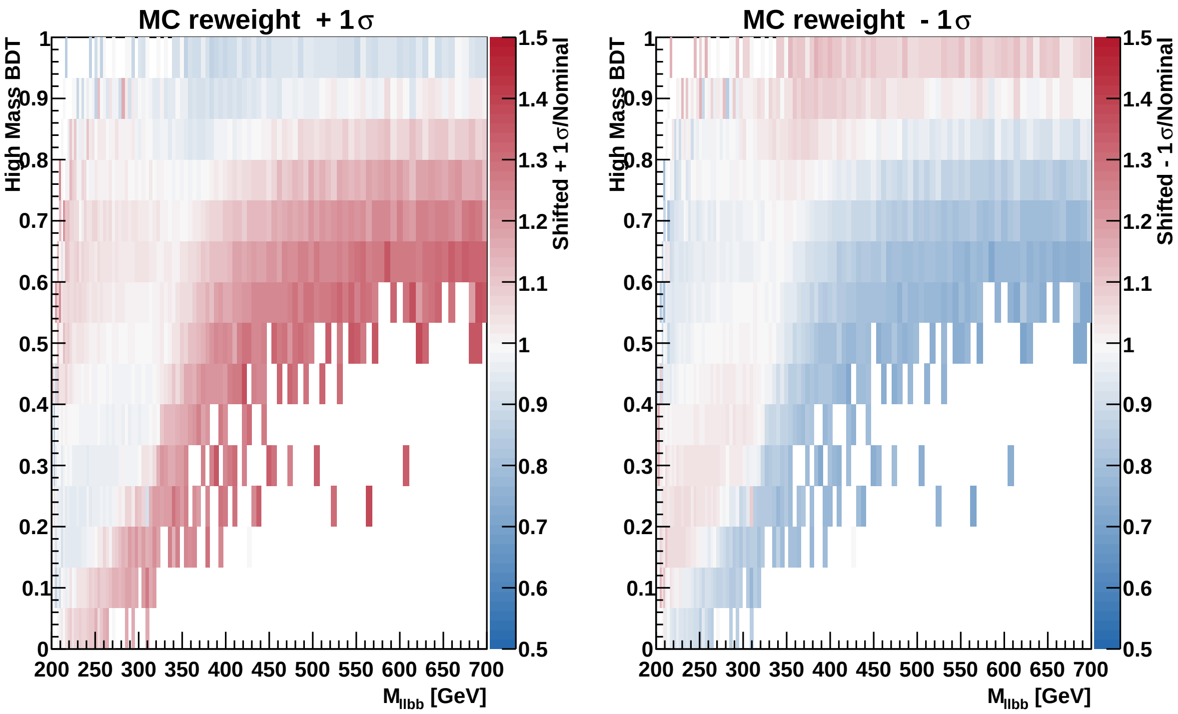

Figure 5.37: Ratio between the nominal value and the $\pm 1\sigma$ shifted template for the MC reweight systematic in the 4-body invariant mass vs low Mass BDT plane.

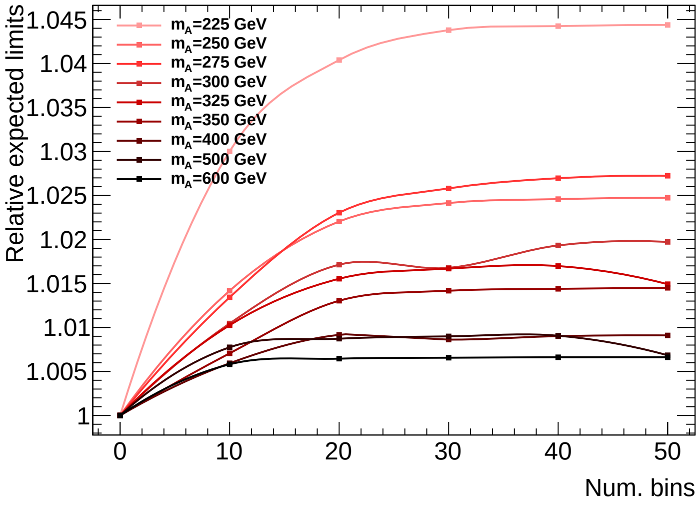

Figure 5.38: Expected limits as a function of the number of bins considered for the Monte-Carlo-statistics systematic uncertainty, relative to the expected one without considering it, for different m_{A} values.

leptons and jets are already included in the specific uncertainty evaluation; no shape uncertainty is considered.

An additional uncertainty source is the limited knowledge of the total inelastic cross-section, used to estimate the expected number of interactions for pileup reweighting. A 5% uncertainty is assumed for the default value of 69.3 mb, and the pileup weights are varied accordingly. No shape uncertainty is considered.

The impact of PDF uncertainty is evaluated following the recommendations of the PDF4LHC [103] working group. A per-event weight is calculated with different PDF sets (CTEQ66, MSTW2008nlo and NNPDF20) provided by the LHAPDF package [137]. The sum of the event weights in the signal region corresponding to the relative change between the "generation" and "evaluation" PDF is considered as systematic uncertainty.

A summary of all systematic uncertainties is listed in Tab. 5.15.

Table 5.15: Summary of systematic uncertainties for backgrounds and signals. The top table shows uncertainties concerning shape, bottom normalization. The last two columns report respectively the increase in the expected limit and the percentage decrease in the expected limit spread when each single systematic source is frozen.

	Main backgrounds (Drell-Yan, $t\bar{t}$)	Other electroweak (single-top, VV, Vh)	Signal	Effect after freezing exp. limit	Effect after freezing spread
		Shape			
Jet Energy Scale	0 − 4%		0 − 8%	0 − 1%	0 − 2%
Jet Energy Resolution	0 − 2%		0 − 4%	0 − 1%	0 − 2%
b tagging	0 − 4%		0 − 8%	< 1%	0 − 2%
Factorization and renormalization scale	0 − 6%		6 − 10%	0 − 2%	0 − 4%
Monte Carlo modeling	0 − 15%		-	0 − 6%	1 − 8%
Monte Carlo statistics	1 − 4%		-	0 − 4%	1 − 4%
		Normalization			
Control region fit	0 − 2.4%	-	-	< 1%	0 − 1%
Extrapolation	2 − 13%	-	-	0 − 1%	0 − 2%
Lepton and trigger efficiency	-	2.5%	2.5%	< 1%	< 1%
Jet Energy Scale	-	5.7%	3.8 − 0.2%	< 1%	< 1%
Jet Energy Resolution	-	3.2%	0.8 − 0.5%	< 1%	< 1%
b tagging	-	4.9%	3.6 − 3.2%	< 1%	< 1%
Unclustered $E_{\mathrm{T}}^{\mathrm{miss}}$	-	1.9%	1.4 − 1.0%	< 1%	< 1%
Pile-up	-	0.9%	1.2%	< 1%	< 1%
PDF	-	4.3%	4.0 − 7.9%	< 1%	0 − 3%
Cross Section	-	9.2 − 15 %	-	< 1%	< 1%
Luminosity	-	2.6%	2.6%	< 1%	< 1%

5.9 SIGNAL EXTRACTION STRATEGY AND RESULTS

The 4-body invariant mass peak in signal after the kinematic fit becomes very narrow, up to few GeV for the low mass points. Combining all variables (including m_{A}) to build a single MVA discriminator would require a different training and testing for the whole mass range in steps comparable to the peak resolution. Furthermore, since the discrimination power of the 4-body mass is much larger than any other variable, the is the concrete risk for the BDT to specialize only on that particular variable, assigning a very low weight to all the others. Two solutions have been considered to overcome this potential issue, and to check the potential BDT improvement over the cut-based analysis at the same time.

1. The first and easier step is to train the BDT without m_{A}. A cut on the BDT output selects a signal-enriched region, followed by a fit to the 4-body invariant mass. The discriminating power of the MVA method is reduced due to the sharp cut on the output.
2. A more complex yet effective method can be built, fully exploiting the correlation between the BDT output and m_{A} through a *two-dimensional* (2D) fit. This method has better performance than the cut-based one especially for higher masses, and henceforth will be considered as default for the signal extraction fits with the BDT.

Two-dimensional distributions are projected into subsequent 1D histograms, one each BDT bin, with BDT bin 15 being the closest to 1 and the most signal-enriched, while BDT bin 1 containing the most background-like events. Events passing base BDT selection cuts are reported in Tab. 5.16. The binning of the $m_{\ell\ell bb}$ mass histograms is variable as a function of the mass resolution. At lower masses the bin width is 2 GeV, increasing linearly to 8 GeV at the end of the mass spectrum. Signal templates are generated every 1 GeV in the low and intermediate mass regions, and 2 GeV in the high mass region, using the template morphing technique described in App. A.2. In order to avoid empty bins for the background in the $m_{\ell\ell bb}$ distributions in any BDT bin, which might cause problem in the fit procedure, a single underflow and overflow

Dataset	Exp. Events	Fraction	Signal	Exp. Events	Efficiency
Data	5356 ± 73	-	m_{A}		
VV	158.5 ± 2.2	2.9%	225 GeV	46.9 ± 0.2	9.98%
W+jets	0 ± 0	0%	250 GeV	55.2 ± 0.2	11.8%
Single Top	37 ± 4	0.68%	275 GeV	61.0 ± 0.2	13.0%
tt (+V)	1022 ± 7	19%	300 GeV	65.7 ± 0.3	14.0%
Z+bb	2460 ± 20	45%	325 GeV	70.3 ± 0.3	15.0%
Z+b	1211 ± 13	22%	350 GeV	74.1 ± 0.3	15.8%
Z+jets	516 ± 14	9.5%	400 GeV	81.0 ± 0.3	17.2%
Multijet	0 ± 0	0%	500 GeV	92.1 ± 0.3	19.5%
SM Zh	29.36 ± 0.12	0.54%	600 GeV	97.8 ± 0.3	20.7%
All Bkg.	5434 ± 29	-			

Table 5.16: Expected and observed events after the BDT selection cuts and the $90 < m_{\mathrm{bb}} < 140\,\mathrm{GeV}$ requirement. Together with the number of expected events also the event fractions are reported for backgrounds, while for signal the selection efficiency. The signal yield is referred to a $\sigma \times \mathcal{B}r(\mathrm{A} \to \mathrm{Zh} \to \ell\ell\mathrm{b}\bar{\mathrm{b}}) = 25\,\mathrm{fb}$.

bin is used at the lower and upper edges of the bulk of the $m_{\ell\ell bb}$ distribution. Since the signal is significantly present only at high BDT, and the background normalization is already constrained by the scale factors, the 2D fit is performed on the last 6 BDT bins only, namely only for BDT$>$ 0.6.

The BDT distribution for data and background are shown in Fig. 5.33. For illustration purposes, the four-body mass distribution for the last 6 BDT bins, weighted by the S/(S+B) ratio in each BDT bins, are shown in Fig. 5.41.

5.9.1 POST-FIT RESULTS

The signal extraction procedure allows the background to vary within its systematic uncertainties modifying the corresponding nuisance parameters. The results are presented as histograms (Fig. 5.39), where the bin values represent the pull values $\Delta x/\sigma_{in}$, and its error bar the fit constraining on the Gaussian width of the nuisance parameter σ_{out}/σ_{in}, for both the background-only and the signal+background fit. The pre- and post-fit histogram comparisons for the BDT and $m_{\ell\ell bb}$ distributions for each bin of BDT are reported in Fig. 5.40 for the sample mass point $m_{\rm A} = 560\,$GeV.

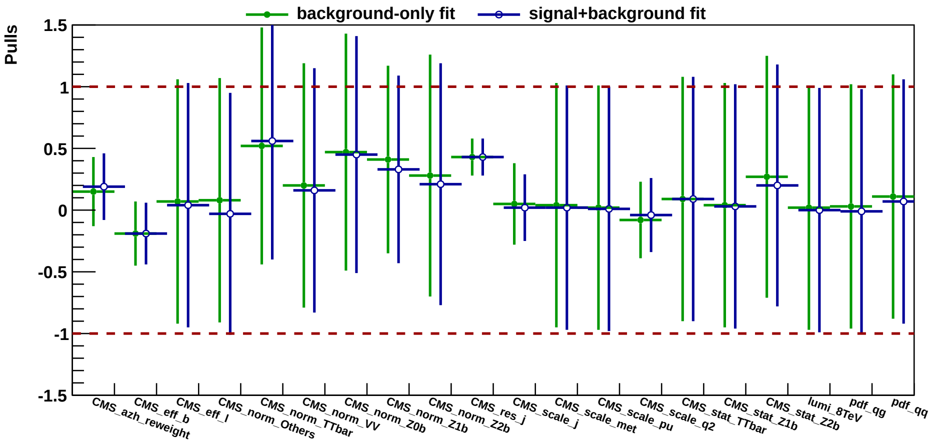

Figure 5.39: Values of nuisance parameters after the fit for $m_{\rm A}$ =560 GeV in the high mass BDT region. Uncertainties due to MC statistics are not shown.

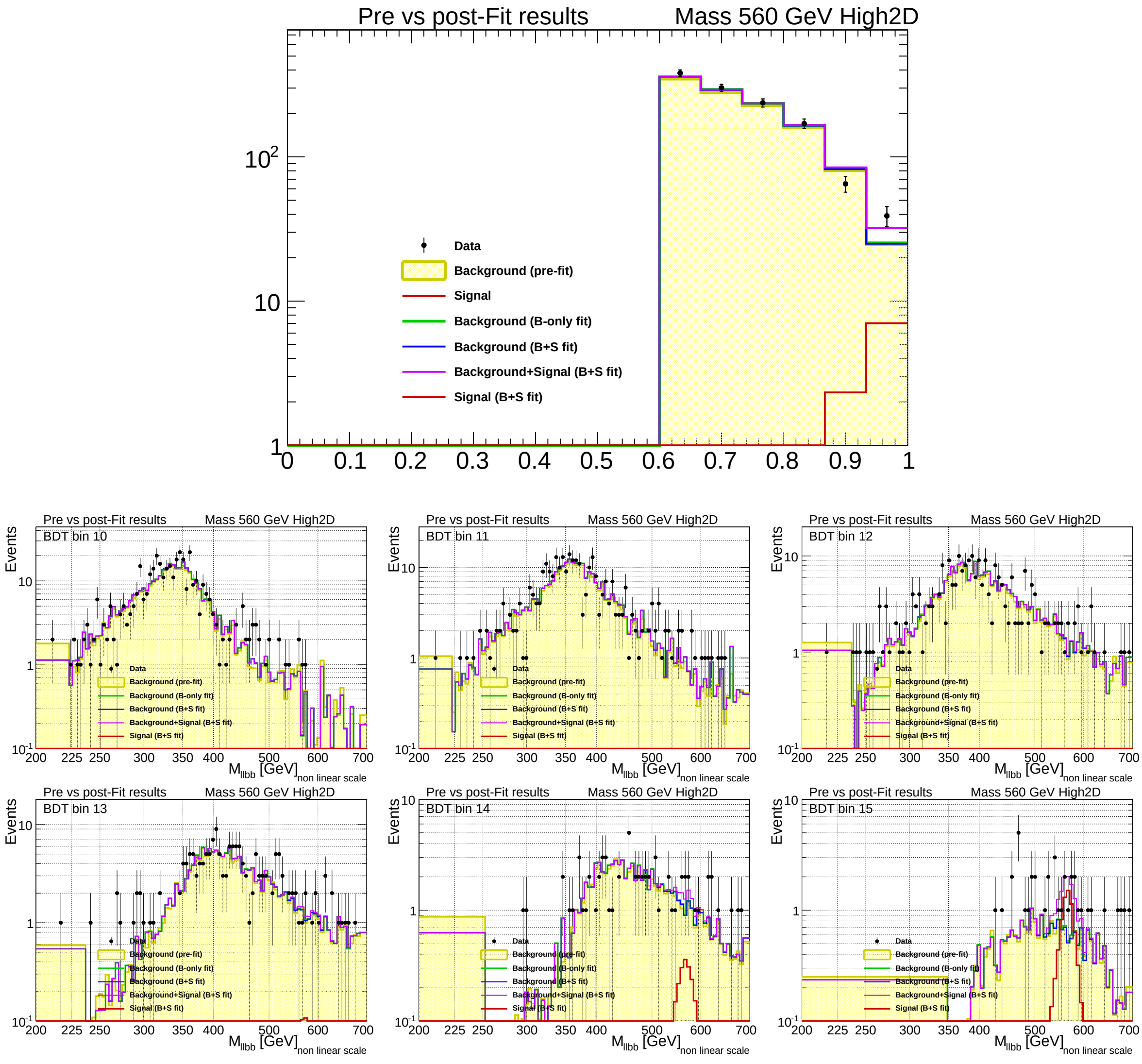

Figure 5.40: Pre- and post-fit distribution for background against data for m_{A} =560 GeV. The top plot shows the BDT projection, bottom plots the $m_{\ell\ell bb}$ in each BDT bin used in the 2D fit (BDT$>$ 0.6).

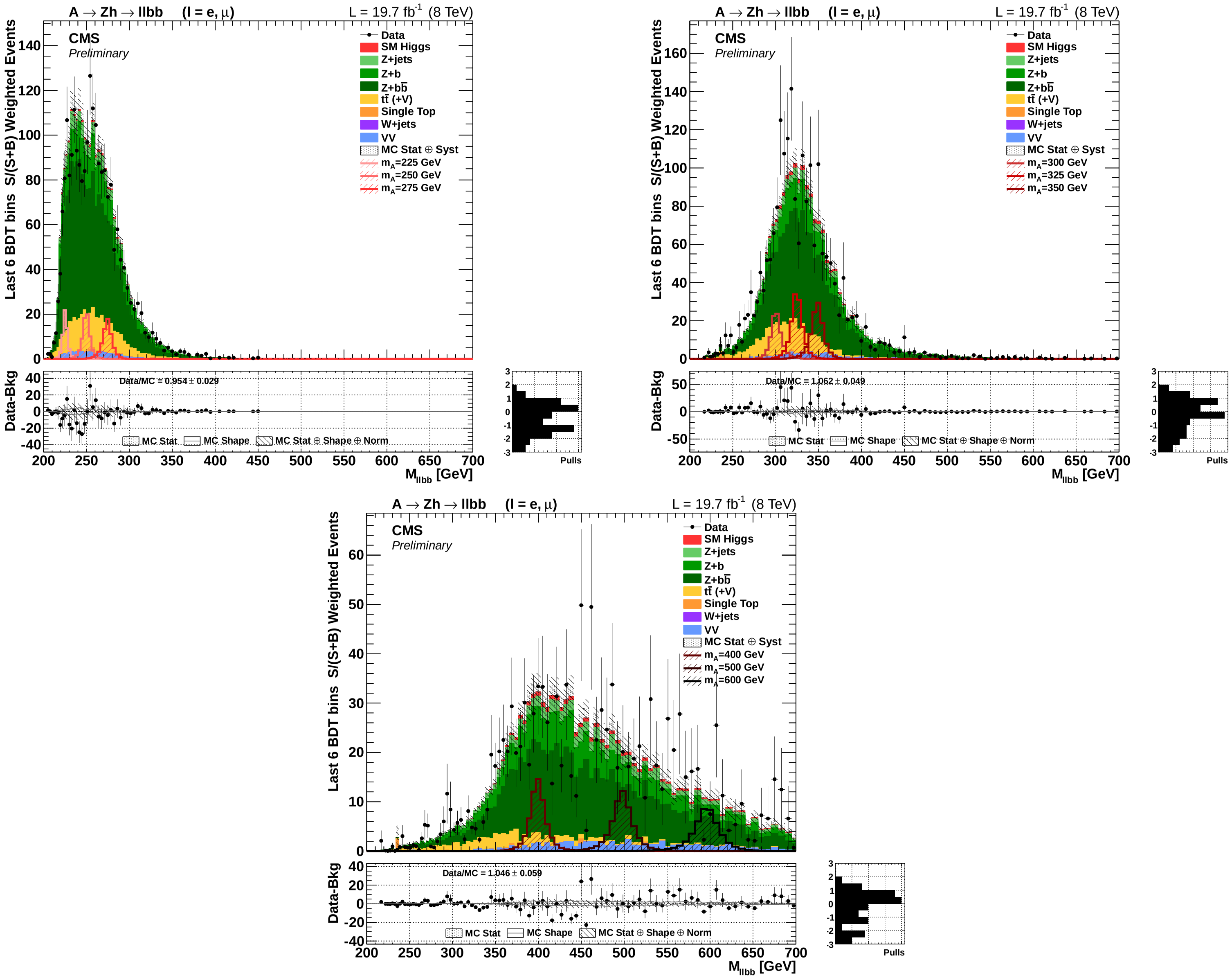

Figure 5.41: Pre-fit BDT output for data and background processes (left), and four-body invariant mass distribution (right) for the Low (top), Intermediate (center) and High (bottom) mass regions. The $m_{\ell\ell bb}$ plots are shown for the last 6 BDT bins, weighted by the signal over background ratio of each BDT bin. Signal histograms are normalized to the expected exclusion limit.

5.9.2 RESULTS

The results for the $\mathrm{A} \to \mathrm{Zh} \to \ell\ell\mathrm{b}\bar{\mathrm{b}}$ search are computed with the $\mathrm{CL_s}$ criterion [105, 106] implemented in the ROOSTATS package [107]. The *asymptotic* approximation, described in App. A, is used to calculate preliminary 95% CL upper limits with 1σ and 2σ bands using the $\mathrm{CL_s}$ frequentist calculation currently recommended by the LHC Higgs Combination Group [138]. The *profile likelihood* method is used for significance and the background p-value; finally, a *maximum likelihood* fit allows to get the signal best fitted ratio.

Results are shown graphically in Fig. 5.42 in term of $\sigma \times \mathcal{B}r(\mathrm{A} \to \mathrm{Zh} \to \ell\ell\mathrm{b}\bar{\mathrm{b}})$, including all systematic uncertainties described in Sec. 4.6, together with expected limit with ± 1 and $\pm 2\sigma$ bands, as a function of the mass of A. Numerical values are reported in Tab. 5.17. An upper limit as low as 20 fb for low mass and 2.0 fb at high mass is expected.

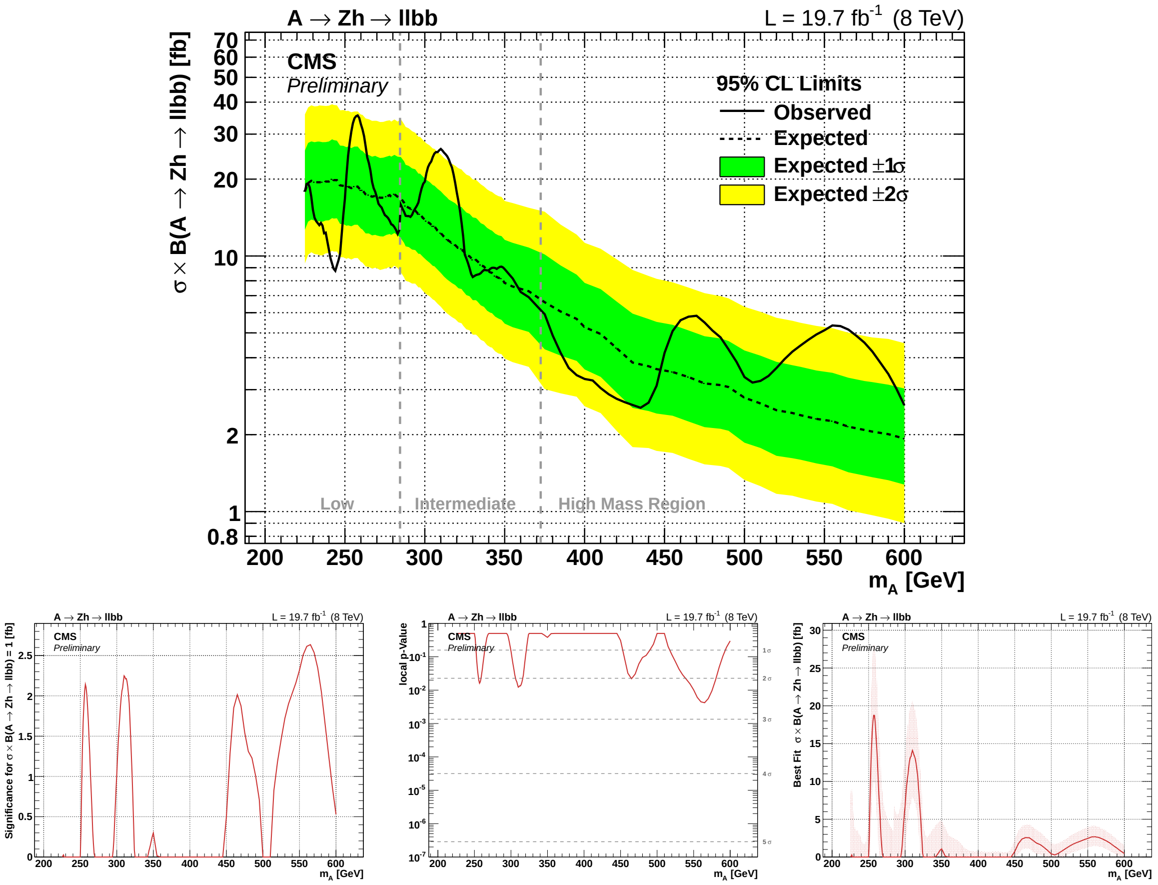

Figure 5.42: Observed and expected (with ± 1, 2σ band) 95% CL upper limit on $\sigma \times \mathcal{B}r(\mathrm{A} \to \mathrm{Zh} \to \ell\ell\mathrm{b}\bar{\mathrm{b}})$ (top), significance of the excesses (bottom left), background p-value (bottom center) and best fitted cross section (bottom right) as a function of m_{A}, including all statistical and systematic uncertainties.

The most significant excess, located at $m_{\mathrm{A}} = 560\,\mathrm{GeV}$, has a local p-value of $4.57 \cdot 10^{-3}$, corresponding to a significance of 2.6σ. Trial factors are calculated to take into account the probability of observing such an excess anywhere in the scanned mass range. In this case, the upcrossing approximation is used [139]. The trial factor is 30, thus the global p-value is 0.139, corresponding to a significance of 1.1σ. The local and global significance for the largest excess in each mass region are reported in Tab. 5.18.

m_{A} (GeV)	225	250	275	300	325	350	400	500	600
Observed (fb)	17.9	16.8	14.8	19.5	10.1	8.84	3.29	3.35	2.61
Expected (fb)	17.9	18.1	16.4	13.6	10.0	7.84	5.27	2.79	1.93
Expected -1σ (fb)	12.7	12.8	11.7	9.50	6.98	5.44	3.59	1.86	1.27
Expected -2σ (fb)	9.35	9.42	8.64	7.02	5.11	3.97	2.59	1.33	0.90
Expected $+1\sigma$ (fb)	25.9	25.9	23.6	19.7	14.7	11.6	7.89	4.27	3.03
Expected $+2\sigma$ (fb)	35.8	35.8	32.6	27.3	20.7	16.5	11.3	6.32	4.56

Table 5.17: Observed and expected (with $\pm 1(2)\sigma$ band) 95% CL upper limit on $\sigma \times \mathcal{B}r(\mathrm{A} \to \mathrm{Zh} \to \ell\ell\mathrm{b}\bar{\mathrm{b}})$ as a function of m_{A}, including all statistical and systematic uncertainties.

Region	Local		Global		Trial factors
	significance	p-value	significance	p-value	
All	2.6	0.0046	$1.1^{+0.4}_{-0.3}$	0.196	30
Low Mass	1.9	0.0290	$0.9^{+1.0}_{-0.5}$	0.196	6
Int Mass	2.2	0.0123	$0.9^{+0.9}_{-0.4}$	0.173	14
High Mass	2.6	0.0046	$1.5^{+0.5}_{-0.3}$	0.066	15

Table 5.18: Local and global p-values and significance for the whole mass spectrum and each mass region independently.

5.9.3 WIDTH SCAN

While the narrow-width approximation is generally valid for $m_{\mathrm{A}} < 2m_{top}$, the width of the A boson heavily depends on the model parameters for higher masses. Upper limits are derived also in the case the narrow-width approximation is not valid, thus looking for a resonance with a given width larger than experimental resolution. Since generating signal samples scanning a grid of A boson mass and width values would be extremely CPU expensive, an alternative approach is adopted. The signal 4-body invariant mass histograms are smeared with a Breit-Wigner function with a certain width, and the resulting templates are provided to the 2D fit instead of the original ones. This procedure assumes that the variables other than $m_{\ell\ell bb}$ are not significantly influenced by the A boson width. Expected and observed limits are obtained for the three mass points above $2m_{top}$ ($m_{\mathrm{A}} = 400, 500, 600\,\mathrm{GeV}$) for $\Gamma_{\mathrm{A}} = 0, 5, 10, 20, 30, 40\,\mathrm{GeV}$, as shown in Fig. 5.43. The region in the 2HDM parameter space with a larger width is negligible. Exclusion limits as a function of the mass in the high mass region are recomputed as well, considering the A width fixed to a typical value of $\Gamma_{\mathrm{A}} = 30\,\mathrm{GeV}$. In this case, the local significance of the $m_{\mathrm{A}} = 560\,\mathrm{GeV}$ peak increases to 2.9σ, given the broadness of the excess; the global significance, taking into account the look-elsewhere effect, is $1.5^{+0.3}_{-0.2}\sigma$.

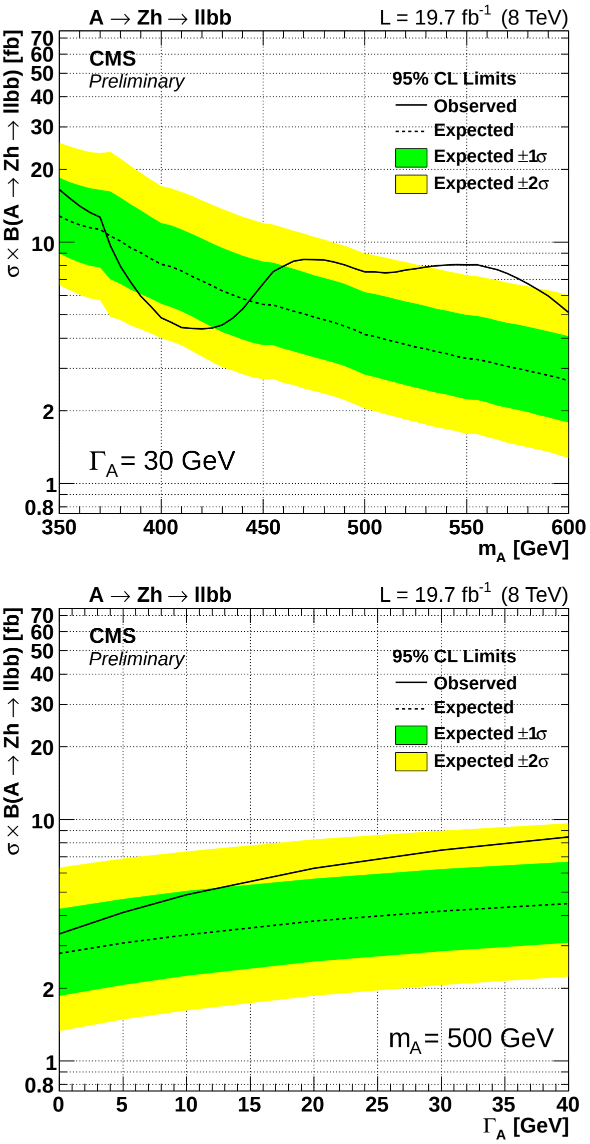

Figure 5.43: Observed and expected 95% CL upper limit as a function of m_A for $\Gamma_A = 30\,\text{GeV}$ (top left), and as a function of Γ_A for $m_A = 400\,\text{GeV}$ (top right), 500 GeV (bottom left), 600 GeV (bottom right).

5.9.4 2D FIT CROSS-CHECKS

In order to validate the 2D fit, two different but complementary methods can be pursued. The first one is to cut on the BDT output and perform a one-dimensional signal fit to the $m_{\ell\ell bb}$ shape, as described in the beginning of the section. A second alternative approach is to select events in a $m_{\ell\ell bb}$ mass window and then fit the 1D BDT shape. For the first test, a fixed cut at BDT$>$ 0.8 is applied for all mass points, while for the second one the widow width ranges from 5 GeV to 20 GeV depending on the generated mass point. The shape fits are performed with all normalization and shape systematic uncertainties included. Fit results are then compared with the expected limit obtained with the 2D fit. The maximum sensitivity is expected for the latter, since cutting on the BDT or the mass does not allow to exploit the full shape information of the two, but the expected limits should not show relevant differences. Fit results are listed in Tab. 5.19 and shown in Fig. 5.44.

m_{A} (GeV)	225	250	275	300	325	350	400	500	600
2D fit m_{A} vs BDT									
Observed (fb)	17.9	16.8	14.8	19.5	10.1	8.84	3.29	3.35	2.61
Expected (fb)	17.9	18.1	16.4	13.6	10.0	7.84	5.27	2.79	1.93
1D fit m_{A}, BDT cut									
Observed (fb)	25.4	18.0	20.9	28.5	16.0	9.26	6.32	4.72	3.59
Expected (fb)	21.7	21.0	19.6	15.1	12.7	10.3	6.78	4.36	2.99
1D fit BDT, cut m_{A}									
Observed (fb)	24.7	22.4	21.6	21.3	12.4	11.3	5.78	3.78	2.60
Expected (fb)	25.4	25.6	22.9	16.6	12.9	9.91	6.17	3.36	2.41

Table 5.19: Observed and expected limit for the 2D fit and for the cross-check 1D fits.

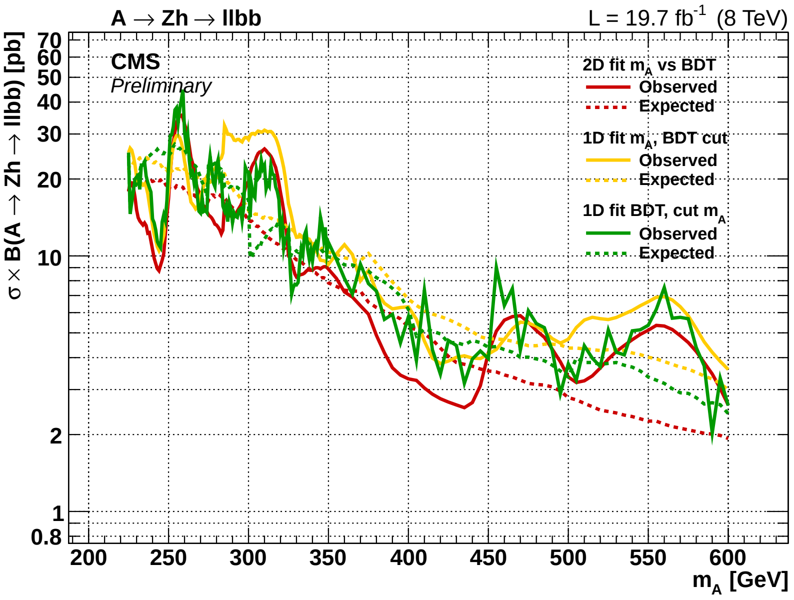

Figure 5.44: Observed and expected limit comparison for the 2D fit cross-check.

5.10 RESULTS INTERPRETATION

5.10.1 INTERPRETATION IN 2HDM

The $\mathrm{A} \to \mathrm{Zh} \to \ell\ell\mathrm{b}\bar{\mathrm{b}}$ analysis results can be interpreted in terms of the two Higgs doublet models (2HDM), described in Sec. 2.2. Only Type-I and Type-II 2HDM are considered, since Type-III and IV differs from I and II only for the coupling of the leptons, and the differences are irrelevant for the current analysis. In order to reduce the number of free parameters, the following assumptions are made:

- the lightest Higgs boson is the SM-like, recently discovered by LHC with $m_{\mathrm{h}} = 125\,\mathrm{GeV}$
- $m_{\mathrm{H}} = m_{\mathrm{H}^{\pm}} = m_{\mathrm{A}}$, since A searches are not strongly dependent on m_{H} and $m_{\mathrm{H}^{\pm}}$, and a large mass splitting would break the custodial symmetry and give large contribution to $\rho \neq 1$
- $m^2_{12} = m^2_A \frac{\tan\beta}{1+\tan^2\beta}$, so the discrete Z_2 symmetry is broken as in the MSSM
- $\lambda_{6,7} = 0$ to avoid CP-violation at tree-level.

The signal cross section, in both the gluon-gluon fusion and b quarks associated production modes, and $\mathcal{B}r(\mathrm{A} \to \mathrm{Zh})$ and $\mathcal{B}r(\mathrm{h} \to \mathrm{b}\bar{\mathrm{b}})$ branching fractions are calculated with the SUSHI [140] and 2HDMC [141] software, both at next-to-next-leading order (NNLO), using the MSTW2008LO90CL, NLO, NNLO [142] parton density functions. The b quark associated production cross section (bbA) is rescaled by the different signal acceptance relative to the gluon-gluon fusion process (ggA), as well as for the selection efficiency of the dijet pair due to the possible combinations arising from the presence of additional b quarks in the event. After the previous assumptions, the remaining free parameters are one mass, m_{A}, and the two angles, α, and β. The results will be presented a given mass point m_{A} in the plane $\tan\beta$ vs $\cos(\beta-\alpha)$. In the limit for $\cos(\beta-\alpha) \to 0$, called *alignment limit*, the light Higgs boson h behaves as the SM one, and the $\mathrm{A} \to \mathrm{Zh}$ branching ratio vanishes. The deep valley with $\mathcal{B}r(\mathrm{h} \to \mathrm{b}\bar{\mathrm{b}}) \sim 0$ present in Type-II models is due to the coupling of the down-type quarks to the Higgs boson, which is proportional to $(-\sin\alpha/\cos\beta)$, vanishing for $\alpha \sim 0$. For each mass point ($m_A = 225, 250, 275, 300, 325, 350, 400, 500, 600$ GeV), a cross section and branching fraction scan is performed for $0.1 \leq \tan\beta \leq 100$ and $-1 \leq \cos(\beta-\alpha) \leq 1$, using the convention $0 < \beta - \alpha < \pi$. The total cross section ($ggA + bbA$) multiplied by the $\mathcal{B}r(\mathrm{A} \to \mathrm{Zh})$ and the $\mathcal{B}r(\mathrm{h} \to \mathrm{b}\bar{\mathrm{b}})$ branching ratios for the mass sample $m_{\mathrm{A}} = 300\,\mathrm{GeV}$ are shown in Fig. 5.45. Finally, in Fig. 5.46, the expected and observed limit for each mass point are reported in the $\tan\beta$ vs $\cos(\beta-\alpha)$ plane.

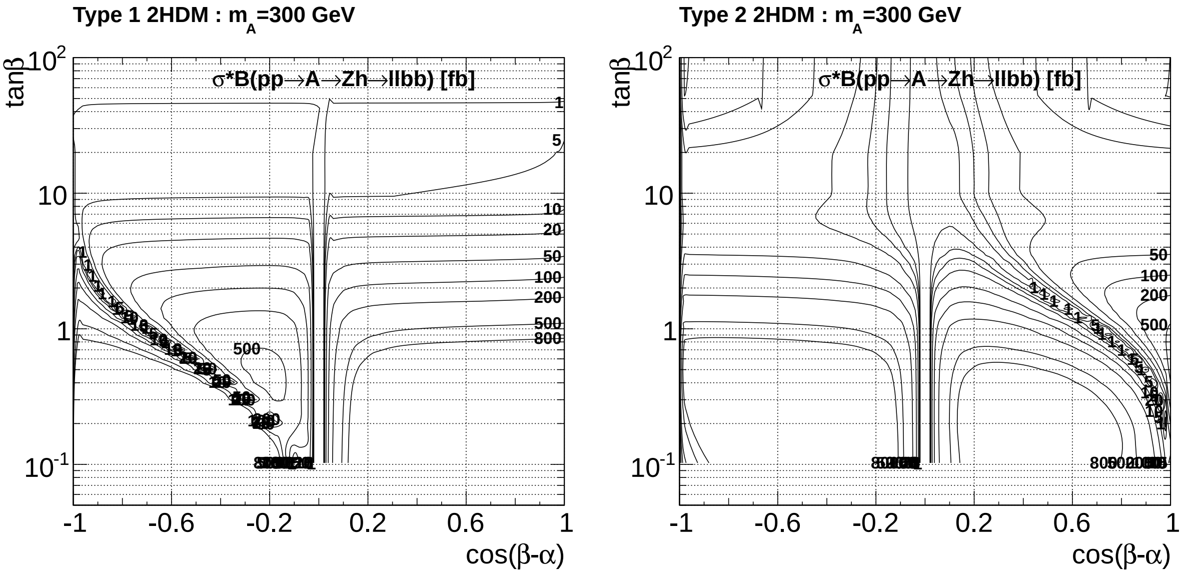

Figure 5.45: $\sigma(pp \to \text{A}) \times \mathcal{B}r(\text{A} \to \text{Zh} \to \ell\ell bb)$ for Type-I (left) and Type-II (right) model, as a function of $\tan\beta$ and $\cos(\beta-\alpha)$, for m_A =300. Both ggA and bbA processes are included.

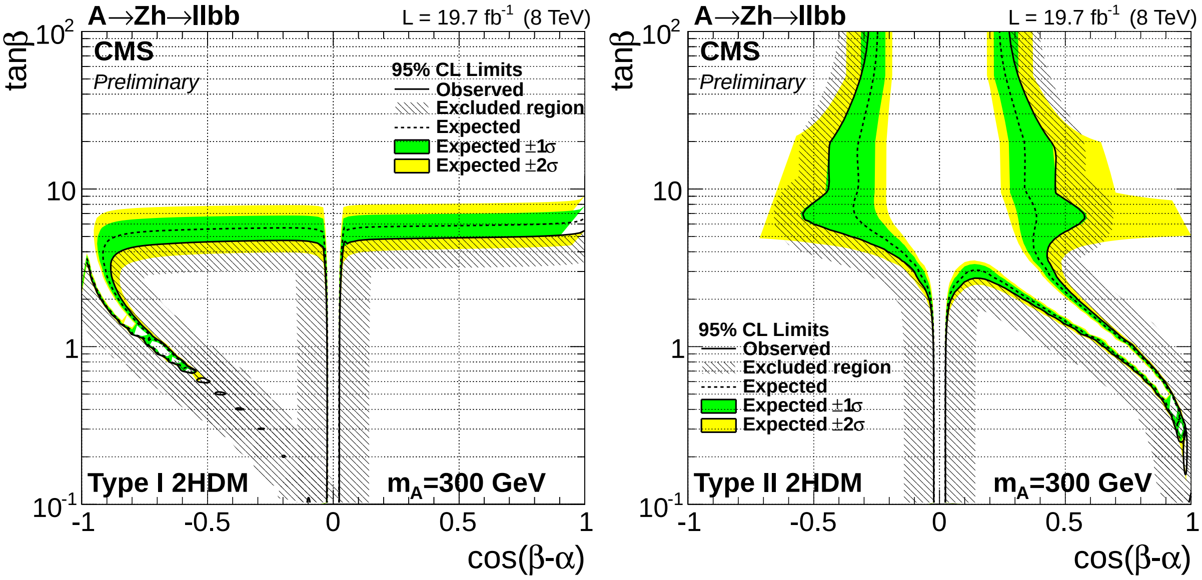

Figure 5.46: Observed and expected 95% CL upper limits, together with $\pm 1, 2\sigma$ filled bands, for Type-I (left) and Type-II (right) 2HDM models, as a function of $\tan\beta$ and $\cos(\beta-\alpha)$, for m_A =300 GeV.

5.10.2 INTERPRETATION IN MSSM

An interpretation of the results within Minimal Supersymmetric Standard Model (MSSM) is also possible, but the problem is to define a suitable benchmark point. As in 2HDM Type-II models, the A production is important only at low $\tan\beta$, so a MSSM scenario where the light Higgs mass m_{h} is close to 125 GeV should be used.

The widely used m_h^{max} scenario, for instance, is well known to allow for a light Higgs with the correct mass only for a limited region of the phase space, more precisely for a narrow band in the plane m_{A}-$\tan\beta$ with $\tan\beta \sim 5$ [143, 144]. Modification of these benchmark scenarios, such as $m_h^{mod\pm}$ extend the allowed region to higher values of $\tan\beta$, but fails to do the same also for low $\tan\beta$. Scenarios with larger M_{SUSY}, of the order of 10-100 TeV, have been studied [109], and can accommodate a proper mass for the h also for very low values of $\tan\beta$, as shown in Fig. 5.47, but no actual benchmark point has been proposed so far.

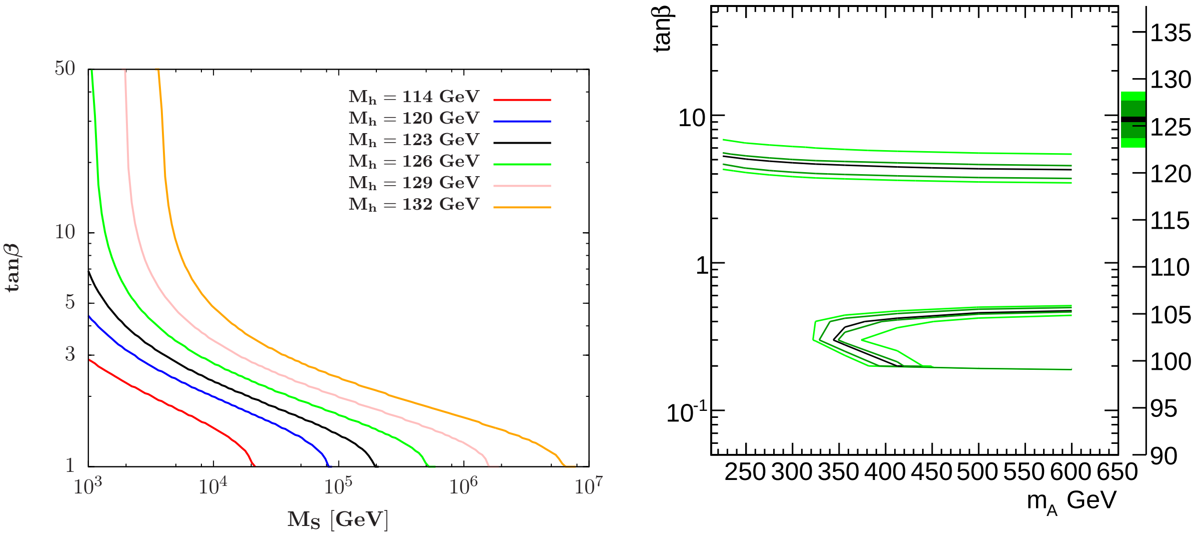

Figure 5.47: Left: dependence of the light Higgs mass in MSSM on $\tan\beta$ and M_{SUSY}, for various m_{h}. The plot comes from [109]. Low values of $\tan\beta$ can produce a correct m_{h} in case M_{SUSY} is large enough. Right: region in the m_{A}-$\tan\beta$ plane where $m_{\mathrm{h}} = 125 \pm 2(3)$ GeV in the MSSM m_h^{max} scenario. The lower $\tan\beta$ region is already excluded by LEP [145].

To quantify the problem of the interpretation of the current analysis in MSSM, the results obtained within m_h^{max} are presented even if this benchmark is not the best choice for the current analysis. The MSSM branching ratios are computed with FEYNHIGGS [146–150], using the parameters as defined in Sec. 2.3.

Fig. 5.48 shows the expected limit for this analysis, in the m_{A}-$\tan\beta$ plane. The sensitivity reaches $\tan\beta \simeq 2$, very far from the region where the benchmark scenario foresees a proper m_{h}. Since the value of m_{h} is an input of the analysis itself, the m_h^{max} scenario leads to an inconsistent interpretation of the results in the MSSM.

In Fig. 5.49 expected limits for two MSSM scenarios alternatives to m_h^{max} are presented: m_h^{mod+} and *light stop*, as described in [143, 144], respectively. The m_h^{mod+} scenario allows for a light Higgs boson of the right mass also for values of $\tan\beta$ higher than those permitted in m_h^{max}, but not lower. The *light stop* foresees a correct m_{h} for even higher regions of $\tan\beta$, equal or larger than $\tan\beta \gtrsim 10$.

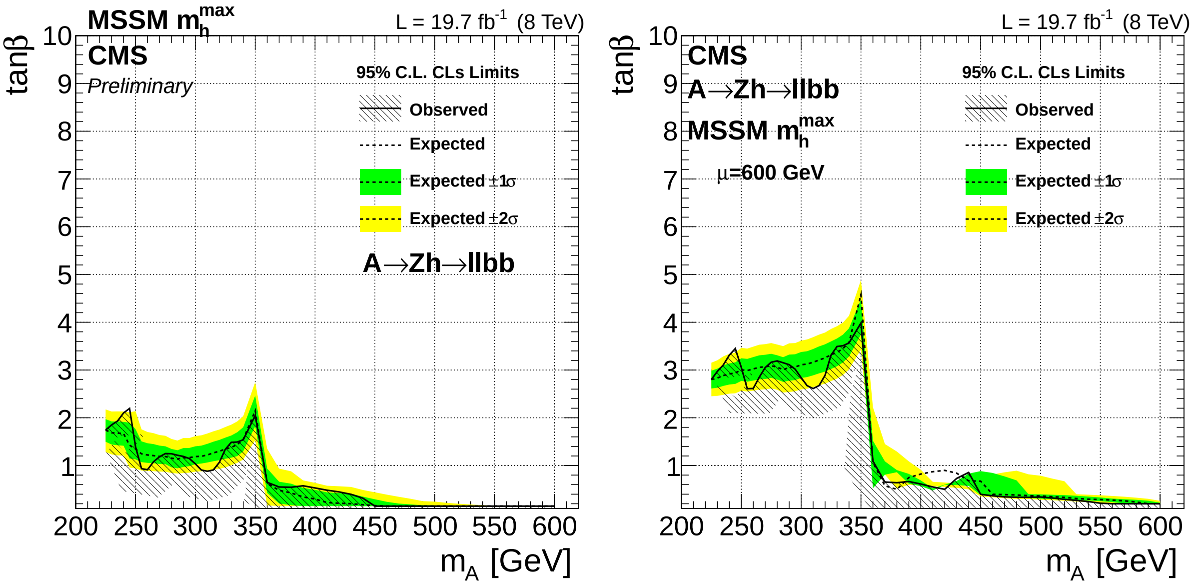

Figure 5.48: Expected limit in the m_A-tan β plane for this analysis, in the MSSM m_h^{max} scenario, with the parameter μ set to the default value of 200 GeV (left), and with $\mu = 600$ GeV (right). The region below the limit is excluded.

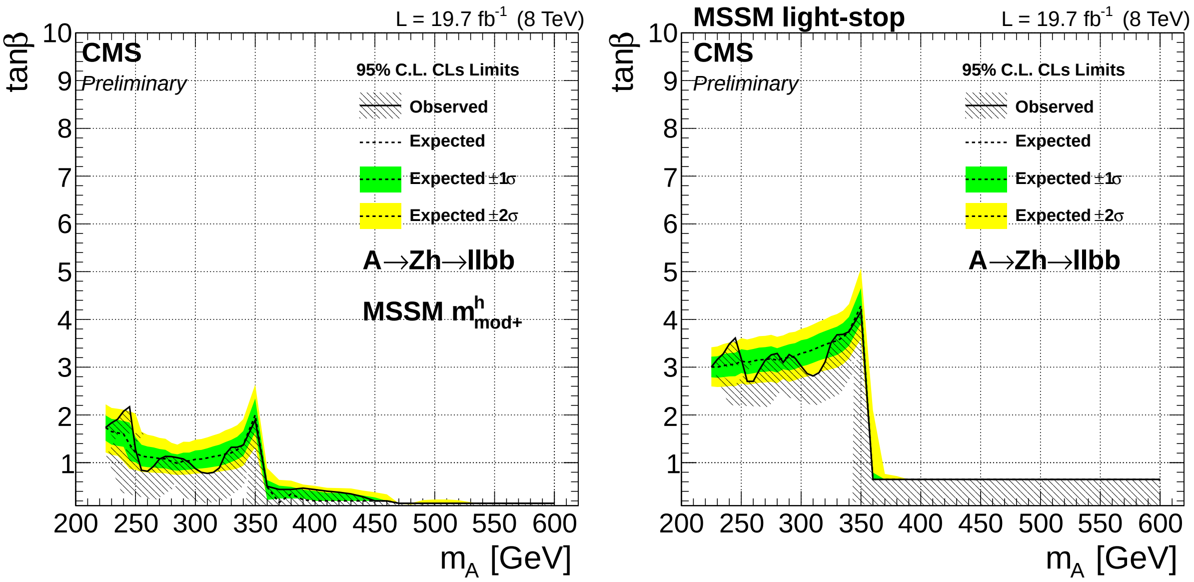

Figure 5.49: Expected limit in the m_A-tan β plane for this analysis, in the MSSM m_h^{mod+} (left) and *light stop* scenario (right). The region below the limit is excluded.

In these benchmark scenarios, proposed before the start of LHC, the M_{SUSY} is rather low, in the order of the TeV, leading to light sparticles below the experimental limit of direct and indirect searches [151]. The mass of charginos and neutralinos, and their couplings to the Higgs bosons, depend on the MSSM parameters μ and M_2 (the Higgs-Higgsino and $SU(2)$ gaugino mass parameters) and the decay A $\rightarrow \chi\chi$ can be dominant, reducing the sensitivity of the analysis. Large values of μ and M_2 are more favorable for the observation of the A $\rightarrow$ Zh channel. The analysis results are then evaluated both in the m_h^{max} benchmark scenario (where μ is set to 200 GeV, see Sec. 2.3) and in a slightly modified scenario with $\mu = 600$ GeV, where sparticles are heavier and $\mathcal{B}r(\mathrm{A} \rightarrow \chi\chi)$ is small. This solution, as shown in Fig. 5.48, allows to exclude a region in the m_A-tan β plane where a light Higgs with mass $m_h = 125$ GeV is accessible.

5.11 FUTURE OUTLOOK

In the first months of 2015, after the Long Shutdown 1, LHC is scheduled to restart with an increased center-of-mass energy of 13 TeV and instantaneous luminosity up to $2 \cdot 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, probing an unexplored phase space and beginning a new era for particle physics. The energy boost and the increased luminosity will be extremely profitable for searches of heavy resonances, due to the sharp increase of the partonic luminosity, especially for gluon-initiated processes. The production cross section of heavy particles via gluon-gluon fusion, like the $gg \to \mathrm{A} \to \mathrm{Zh}$, will be enhanced with respect to $\sqrt{s} = 8\,\mathrm{TeV}$ collisions by approximately a factor 3 and 4 for a pseudoscalar boson with mass 300 GeV or 600 GeV, respectively.

The main backgrounds, namely Drell-Yan processes, are expected to increase only by a factor 1.7, due to the fact that their initial state is $q\bar{q}$. Instead, the top quark cross section will be 3.2 times larger, implying a change in the expected background composition, and possibly the need to adjust the analysis strategy to have a better $\mathrm{t\bar{t}}$ discrimination. The top quark background fraction will become $\approx 1/3$ instead of the present $\approx 20\%$, while the Drell-Yan will drop from 80% to about 60%. The S/B ratio at 13 TeV is then expected to be 1.5 times larger for $m_{\mathrm{A}} = 300\,\mathrm{GeV}$, and ≈ 2 times larger for $m_{\mathrm{A}} = 600\,\mathrm{GeV}$. The sensitivity of the current analysis, assuming that systematic uncertainties scale with $\sqrt{\mathcal{L}}$, will be overcome with less than $10\,\mathrm{fb}^{-1}$ for a 300 GeV pseudoscalar, and only $7\,\mathrm{fb}^{-1}$ for a 600 GeV A boson. In the prospect of collecting $300\,\mathrm{fb}^{-1}$ of data during LHC Run II, this channel can exclude a total $\sigma \times \mathcal{B}r$ up to $\approx 2\,\mathrm{fb}$ for $m_{\mathrm{A}} = 300\,\mathrm{GeV}$, and ≈ 0.2 for $m_{\mathrm{A}} = 600\,\mathrm{GeV}$. Interpreting these projections in the MSSM parameter space, the expected exclusion for m_{A} near the top quark threshold reaches $\tan\beta \approx 3$ in the m_h^{max} scenario, or $\tan\beta \approx 5 \sim 6$ if sparticles are heavy enough (e.g., with $\mu = 600\,\mathrm{GeV}$).

In the future outlook, an extension towards higher masses is worth to be considered. Beyond the 600 GeV threshold, the A boson decay products are emitted with increasingly larger p_{T}, entering in the boosted regime. On the hadronic side, the two b quarks emerging from the h decay are so close in angle that requiring two separately-clustered jets could result in an important loss of efficiency. Jet substructure techniques allow to overcome this issue: a single, "fat"-jet of charged and neutral particles is reconstructed in the detector, and inspection of the internal jet substructure can lead to the identification of the two separate sub-jets. To reject the backgrounds, the combined invariant mass of the fat-jet has to be compatible with the mass of the Higgs boson, within resolution. Algorithms for sub-jet b tagging can be used to further discriminate against the background.

LHC Run II will soon provide a large amount of collision at an energy never reached before. The better S/B ratio, and the fact that the analysis is limited by the available amount of data and not by systematic uncertainties, will make the $\mathrm{A} \to \mathrm{Zh} \to \ell\ell\mathrm{b\bar{b}}$ channel much more sensitive to new BSM particles in the very next years.

Conclusions

The first years of LHC operation have been extremely profitable. The main objective of the Run I, the discovery of the Higgs boson, has been fulfilled observing a new resonance with mass $m_{\mathrm{h}} = 125\,\mathrm{GeV}$ and properties fully compatible with those expected for a standard model Higgs boson. The announcement was given simultaneously by the ATLAS and CMS experiments, claiming the discovery with a statistical significance of 5 standard deviations for each experiment.

The searches for an extended Higgs sector, already underway since the first collision data, received a new boost from the discovery. This thesis focuses on two specific analyses aiming to detect signals of beyond the standard model Higgs bosons. In the first one, performed with 2011 data before the Higgs boson discovery, we search for neutral BSM Higgs boson decays into b a quark pair and produced is association with additional b quarks. The channel is predicted by the MSSM to be sensitive to a certain region of the parameter space, where the $\tan\beta$ parameter is large. Even though the best discovery chance in this phase-space belong to the tau-decay channel, the search with b quarks was motivated by some excess observed by Tevatron searches in the same channel. We developed a characteristic analysis strategy based on the presence of a muon from semileptonic b quark decays, relying on background prediction techniques based on data. The results were combined with a separate analysis targeting the same final state but carried out with a different strategy, greatly improving the Tevatron upper limits, reducing sizably the allowed parameter space. For $m_{\Phi} \approx 120\,\mathrm{GeV}$ and $m_{\Phi} \approx 300\,\mathrm{GeV}$, the upper limits on $\tan\beta$ set by Tevatron were brought from 60 and 90 to 20 and 30, respectively.

After the Higgs boson discovery, we used the knowledge of the Higgs mass to increase our reach for new physics searches. Among the largest recipients of this valuable information, we chose to pursue a heavy pseudoscalar Higgs boson decaying into a cascade of a lighter scalar h boson and a Z boson. The Z boson has been detected in the decay into a pair of electrons or muons, and the light Higgs into b quarks. This analysis is quite complementary to the multi-b Higgs search, being sensitive to low-$\tan\beta$ scenarios. Interpreting the light scalar as the SM Higgs boson, the kinematic phase-space of the decay was considerably reduced. We employed several analysis techniques to increase the sensitivity to signal, exploiting the presence of three reconstructible resonances in the final state to improve the reconstructed peak resolution and to discriminate against standard model backgrounds. With no significant excess observed over expected backgrounds, we set model-independent 95% CL upper limits on the product of the production cross section and branching fraction, excluding 10 to 30 fb for m_{A} near the kinematic threshold $m_{\mathrm{A}} \approx 225\,\mathrm{GeV}$, $\approx 8\,\mathrm{fb}$ for $m_{\mathrm{A}} \approx 2m_{top}$, and up to $\approx 3\,\mathrm{fb}$ at the high end (600 GeV) of the considered mass range. We presented the results also as a function of the A boson width, and we interpreted them in the context of Type-I and Type-II 2HDM formulations, reducing thereby the phase-space for extensions of the standard model.

Appendix A

Statistical methods

A.1 STATISTICAL ANALYSIS

A.1.1 THE MODIFIED FREQUENTIST CLS LIMITS

The method chosen for extracting limits for the analyses described in this thesis is the modified frequentist construction ($\mathrm{CL_s}$) [105, 106]. The method is defined by the choice of the test statistics and the treatment of the nuisance parameters. A summary of the $\mathrm{CL_s}$ method is reported in this appendix, and a complete description is provided by the LHC Higgs Combination Group [152].

A preliminar step is the reinterpretation of the systematic error pdfs $\rho(\theta|\tilde{\theta})$ as posteriors arising from a measurement $\tilde{\theta}$, as given by the Bayes' theorem:

$$\rho(\theta|\tilde{\theta}) \sim p(\tilde{\theta}|\theta) \cdot \pi_{\theta}(\theta)$$

where $\pi_{\theta}(\theta)$ functions are hyper-priors for those measurements, chosen to be uniform distributions. This choice entails that if $p(\tilde{\theta}|\theta)$ is Poisson, $\rho(\theta|\tilde{\theta})$ is a gamma distribution, and if $p(\tilde{\theta}|\theta)$ is normal, ρ is normal or log-normal. The assumption on the types of priors $\rho(\theta|\tilde{\theta})$ depend on the source of the systematic uncertainty:

Flat priors : parameters unconstrained by any a-priori consideration or measurement

Gaussian distributions : parameters that can have both positive and negative values

Log-normal distributions : observables that can take only positive values (e.g., cross sections, luminosity, selection efficiencies).

Gamma distributions : uncertainties of statistical nature (e.g., statistical error associated with a number of events simulated in MC or a number of observed events in a control region).

This approach allows the construction in a frequentist manner of sampling distributions for the test statistic. Given the signal strength modifier μ and the set of the nuisance parameters θ, the likelihood $\mathcal{L}(data|\mu, \theta)$ used constructing the test statistic is:

$$\mathcal{L}(data|\mu, \theta) = \text{Poisson}\left(data|\mu \cdot s(\theta) + b(\theta)\right) \cdot p(\tilde{\theta}|\theta)$$

where $\text{Poisson}\left(data|\mu s(\theta) + b(\theta)\right)$ is the Poissonian probability to observe a certain value in *data*, assuming signal $s(\theta)$ and background models $b(\theta)$ that depend on the nuisance parameters θ.

The test statistic is then defined as the profile likelihood ratio:

$$q_\mu = -2 \ln \frac{\mathcal{L}(\text{data}|\mu, \hat{\theta}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu}, \hat{\theta})}$$

where *data* can be the actual observation or pseudo-data, and with the constraint $0 \leq \hat{\mu} \leq \mu$. Both the denominator and numerator are maximized. In the numerator, μ remains fixed and only the nuisance parameters θ are allowed to float; $\hat{\theta}_\mu$ are the values at which $\mathcal{L}$ is maximized. In the denominator, both μ and θ are allowed to float in the fit, and $\hat{\mu}$ and $\hat{\theta}$ are the values that maximize $\mathcal{L}$. The lower constraint on $\hat{\mu}$ ($0 \leq \hat{\mu}$) is imposed as the signal rate cannot be negative, and the upper constraint $\hat{\mu} \leq \mu$ forces the limit to be one-sided. The value of the test statistic for the actual observation is denoted as q_μ^{obs}.

The values of nuisance parameters $\hat{\theta}_0^{obs}$ and $\hat{\theta}_\mu^{obs}$ are the best describing experimental data for the *background-only* and *signal+background* hypotheses, respectively. Toy Monte Carlo pseudo-data are generated to construct the test statistic *pdf*s for a signal with strength μ, including also the *background-only* hypothesis ($\mu = 0$). The measurements $\tilde{\theta}$ are also randomized in each pseudo-data, using the *pdf*s $p(\tilde{\theta}|\theta)$. For the purposes of *pseudo-data generation*, the nuisance parameters are fixed to their data-driven best-fit values $\hat{\theta}_\mu^{\text{obs}}$ or $\hat{\theta}_0^{\text{obs}}$, but are allowed to float in fits needed to evaluate the test statistic.

After building the two pdfs, the two p-values p_μ and p_0 relative to the observation for the *signal+background* and *background-only* hypotheses are:

$$\begin{aligned} p_\mu &= P\left(q_\mu \geq q_\mu^{obs} | \mu s(\hat{\theta}_\mu^{obs}) + b(\hat{\theta}_\mu^{obs})\right) \\ p_0 &= P\left(q_\mu \geq q_\mu^{obs} | b(\hat{\theta}_0^{obs})\right) \end{aligned}$$

and the $\mathrm{CL_s}(\mu)$ is simply ratio of these two p-values:

$$CL_s(\mu) = \frac{p_\mu}{p_0}$$

The presence of an hypothetical particle is excluded at the $(1-\alpha)$ Confidence Level (CL) if for $\mu = 1$ then $\mathrm{CL_s} = \alpha$. To quote the 95% CL upper limit, the μ value is varied until $\mathrm{CL_s} = 0.05$.

The numerical differences between the three $\mathrm{CL_s}$ formulations employed by LEP [105], Tevatron, and LHC experiments are quite small. The main is that both LEP and Tevatron used a Bayesian-frequentist approach to handle systematic errors generating sampling distributions, while LHC uses a pure frequentist approach. Additionally, LEP did not profile systematic errors in the test statistic, which does not allow to take advantage of the constraints arising from the data used in the statistical analysis. Both LEP and Tevatron used the test statistic with fixed $\mu = 0$ in the denominator, which does not guarantee the desired asymptotic behavior, forcing to generate large amounts of pseudo-data.

A more detailed discussion can be found in Ref. [152].

A.1.2 THE PROFILE LIKELIHOOD ASYMPTOTIC APPROXIMATION

If the requirement $\hat{\mu}$ from the test statistics $\tilde{q}_\mu$ is relaxed, as a consequence of Wilks' theorem [153] q_μ is expected to have half a χ^2 distribution for one degree of freedom in the asymptotic regime. The value of μ that makes $\frac{1}{2}q_\mu = 1.35$ corresponds to a one-sided $CL_{s+b} = 0.05$

probability. Another common choice is $\frac{1}{2}q_\mu = 1.92$, that corresponds to $CL_{s+b} = 0.025$ and, hence, matches $CL_s = 0.05$.

However, with the physical requirement $\hat{\mu} > 0$, the asymptotic behavior of $f(\tilde{q}_\mu|s+b)$ does not follow half the χ^2 distribution, but the formula [154]:

$$f(\tilde{q}_\mu|\mu) = \frac{1}{2}\delta(\tilde{q}_\mu) + \begin{cases} \frac{e^{-\tilde{q}_\mu/2}}{2\sqrt{2\pi\tilde{q}_\mu}} & \text{if } 0 < \tilde{q}_\mu \leq \mu^2/\sigma^2 \\ \frac{1}{\sqrt{2\pi}(2\mu/\sigma)} \exp\left[-\frac{1}{2}\frac{(\tilde{q}_\mu+\mu^2/\sigma^2)^2}{(2\mu/\sigma)^2}\right] & \text{if } \tilde{q}_\mu > \mu^2/\sigma^2 \end{cases}$$

given $\sigma^2 = \frac{\mu^2}{q_{\mu,A}}$, with $q_{\mu,A}$ being the test statistics evaluated with the Asimov dataset, namely the expected background and the nominal nuisance parameters without statistical fluctuations.

From the similar asymptotic formula for $f(\tilde{q}_\mu|b)$ the median expected limits and their 1 and 2σ bands can be derived without performing any toy Monte Carlo experiment. It can also be demonstrated [154] that in the asymptotic limit the test statistics $\tilde{q}_\mu$ and q_μ are equivalent, leading to the same p-values. These formulae can be used to derive the asymptotic relations for the $\mathrm{CL_s}$ upper limit:

$$CL_s = 0.05 = \frac{1 - \Phi(\sqrt{q_\mu})}{\Phi(\sqrt{q_{\mu,A}} - \sqrt{q_\mu})}$$

where Φ^{-1} is the quantile (inverse of the cumulative distribution) of the standard Gaussian. The median and the error bands are given by

$$\mu_{up+N} = \sigma(\Phi^{-1}(1 - \alpha\Phi(N)) + N)$$

with $\alpha = 0.05$ (μ can be taken as μ_{up}^{med} in the calculation of σ). For $N = 0$, the median expected is the same of the $\mathrm{CL_s}$ limit:

$$\mu_{up}^{med} = \sigma\Phi^{-1}(1 - 0.5\alpha)$$

The asymptotic is a good approximation of the full $\mathrm{CL_s}$ method, but possible biases can arise in application cases with a small number of events. In the analyses described in this thesis, both the full $\mathrm{CL_s}$ and the asymptotic approximation are tested for a certain number of mass points, yielding very similar results, with differences usually smaller than 5%.

A.1.3 SIGNIFICANCE OF AN EXCESS

The approximate $\tilde{p}$-value is derived from the asymptotic properties of the test statistic based on the profile likelihood ratio [153]:

$$\tilde{p} = \frac{1}{2}\left[1 - \mathrm{erf}\left(\sqrt{q_0^{\mathrm{obs}}/2}\right)\right]$$

where q_0^{obs} is the observed test statistic calculated for $\mu = 0$ and with only one constraint $0 \leq \hat{\mu}$, which ensures that data deficits are treated differently than data excesses. The significance value is associated to a p-value using the "one-sided Gaussian" convention:

$$p = \int_Z^\infty \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)dx$$

The so-built significance is relative to only one counting-experiment channel, or in a local region in a continuous variable like the invariant mass of a particle. A more accurate characterization of the local p-value should take into account the effect of the choice in possible values of invariant mass, known as *look-elsewhere effect* [139];.

A.2 HISTOGRAM INTERPOLATION

The use of the kinematic fit in the $\mathrm{A} \to \mathrm{Zh} \to \ell\ell\mathrm{b}\bar{\mathrm{b}}$ analysis (described in Sec. 5.6) improves the 4-body mass resolution with invariant mass peaks narrower than 3 GeV in the low mass region. An accurate scan of the whole mass spectrum considered (225 − 600 GeV) would require signal templates separated by less than the peak width, but generated signal mass points are only available every 25 GeV up to $m_{\mathrm{A}} = 350$ GeV, and every 100 GeV for higher masses.

The solution to this problem relies in the *template morphing* technique, consisting in a linear interpolation of the signal template histograms to intermediate points. The implemented procedure is similar to the one used at LEP [155]. Given a generic variable x (m_{A} in the specific case), the interpolated value m and the two starting distributions, $f_1(x)$ and $f_2(x)$ and their corresponding mass values m_1 and m_2, cumulative distributions are built as $F_1(x) = \int_{-\infty}^{x} f_1(x')dx'$ and $F_2(x) = \int_{-\infty}^{x} f_2(x')dx'$. The goal of the interpolation procedure is to obtain a final distribution $\hat{f}(x)$ with a cumulative distribution function $\hat{F}(x) = \int_{-\infty}^{x} \hat{f}(x')dx'$, that has to fulfill the requirement $F_1(x_1) = F_2(x_2) = \hat{F}(x)$ for all the x, x_1, x_2 values such that $x = ax_1 + bx_2$; a and b are parameters between 0 and 1 defined as $a = \frac{m_2 - m}{m_2 - m_1}$ and $b = \frac{m - m_1}{m_2 - m_1}$. This is achieved finding the x position corresponding to a general $y = F_{1,2}(x)$ value by interpolating linearly adjacent points in the $F_1(x)$ and $F_2(x)$ distributions. Once $\hat{F}(x)$ is built, finding the final interpolated distribution $\hat{f}(x)$ is straightforward.

In order to validate the method, a template for a known mass point (e.g. $m_{\mathrm{A}} = 300$ GeV) is predicted starting from the adjacent mass points, and the morphed histogram is compared with the generated one. The result is shown in Fig. A.1-A.3: even if small differences are visible, the Kolmogorov-Smirnov compatibility test p-value returned is close to 1, confirming the good agreement with the interpolated shape. Also the values of the peak mean value, width and integral derived with linear interpolation show full compatibility.

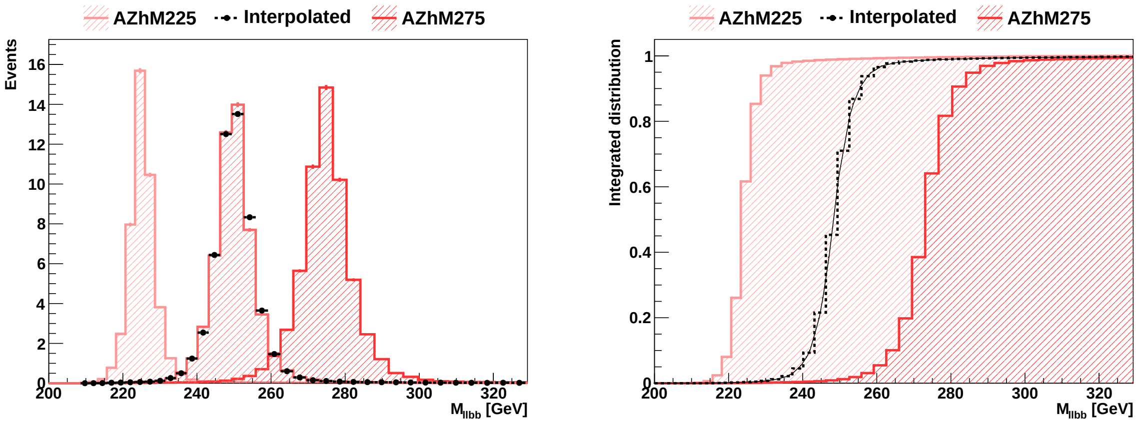

Figure A.1: Morphing test, low mass: a known mass point ($m_{\mathrm{A}} = 250$ GeV) is predicted by template interpolation starting from the two adjacent mass points, $m_{\mathrm{A}} = 225$ GeV and 275 GeV.

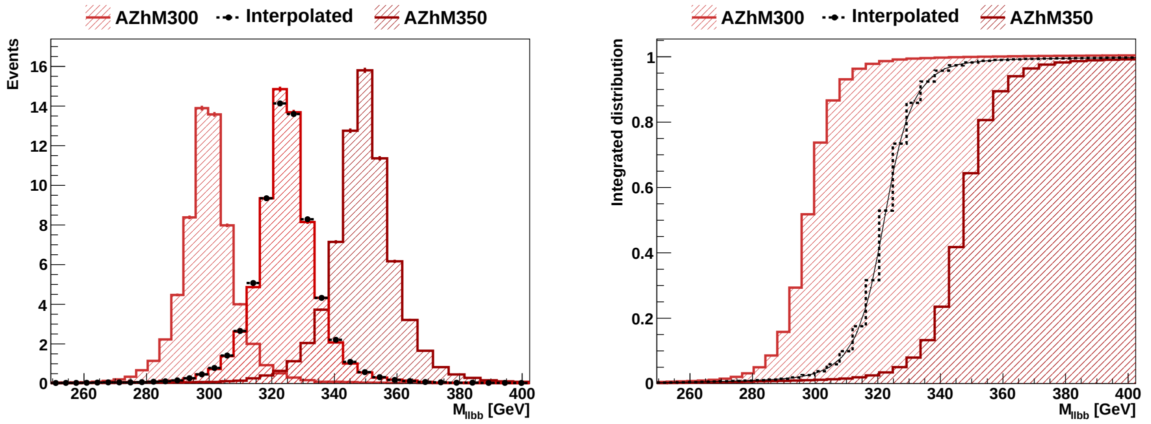

Figure A.2: Morphing test, intermediate mass: a known mass point ($m_\mathrm{A} = 300\,\mathrm{GeV}$) is predicted by template interpolation starting from the two adjacent mass points, $m_\mathrm{A} = 275\,\mathrm{GeV}$ and $325\,\mathrm{GeV}$.

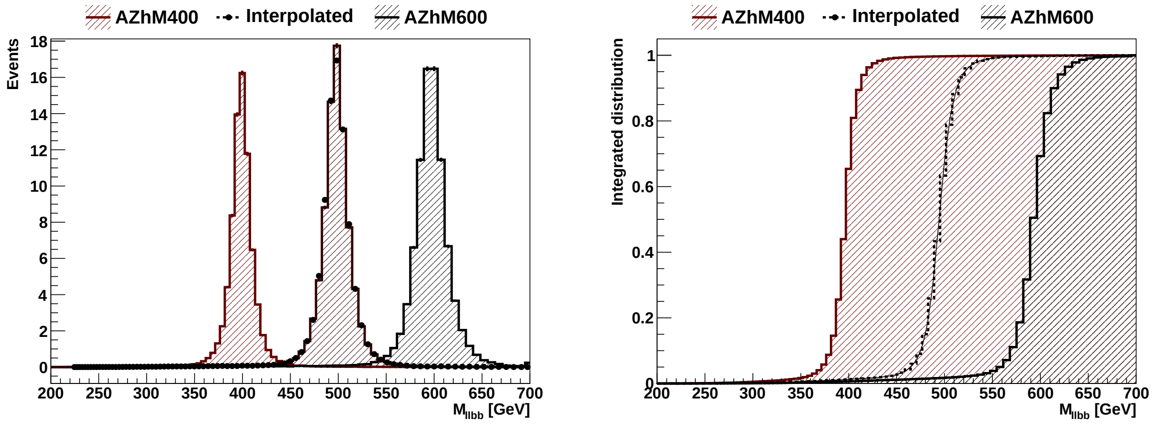

Figure A.3: Morphing test, high mass: a known mass point ($m_\mathrm{A} = 500\,\mathrm{GeV}$) is predicted by template interpolation starting from the two adjacent mass points, $m_\mathrm{A} = 400\,\mathrm{GeV}$ and $600\,\mathrm{GeV}$.

Appendix B

Validations in a semi-leptonic top sample

An important process in data analysis is the validation of the techniques and objects used directly on data, to check against the introduction of potential bias sources. Many tests can be performed directly on the Z boson decaying to leptons, but a known hadronic resonance decaying into jets is really useful for jet testing. The top quark is suitable for this kind of studies, but requiring two isolated leptons in the final state as in the analysis selections does not allow to study hadronic resonances. To overcome this problem, an almost-pure $\mathrm{t\bar{t}}$ sample is selected by reversing the isolation cut (Iso>0.12) on the sub-leading muon, exploiting the presence of a soft muon coming from b quarks semi-leptonic decays. This procedure cannot be done with the electrons, since isolation is required already at trigger level. The selected sample is enriched by semi-leptonic top quark decays, where the isolated, high-p_{T} muon comes from the leptonic $\mathrm{W} \to \mu\nu$ decay, while the other W decays hadronically into a pair of light jets, making possible the full reconstruction of one W and one top quark candidate. This also results in a statistically independent sample with respect to the main analysis.

To enhance $\mathrm{t\bar{t}}$ composition, the p_{T} cut on the leading muon is tightened to $p_{\mathrm{T}}^{\mu} > 24\,\mathrm{GeV}$ to better discriminate against QCD events, while the sub-leading threshold is reduced to 8 GeV, and if the two muons have opposite sign, a Z veto ($75 < m_{\ell\ell} < 105\,\mathrm{GeV}$) and a cut to avoid hadronic resonances ($m_{\ell\ell} > 12\,\mathrm{GeV}$) is applied. As final selection, two of the jets should pass the CSV tight and loose working point, respectively, while light ones should have CSV values not exceeding the loose working point. In order to select events with the correct jet choice from the W, a mass window around its nominal value is imposed ($60 < m_{jj} < 100\,\mathrm{GeV}$). The final selection includes the requirement of exactly 4 jets to avoid ambiguities in the jet choice, a b jet veto for the two W jets, and large missing energy ($E_{\mathrm{T}}^{\mathrm{miss}} > 40\,\mathrm{GeV}$). After each cut, scale factors are applied to simulated events when needed.

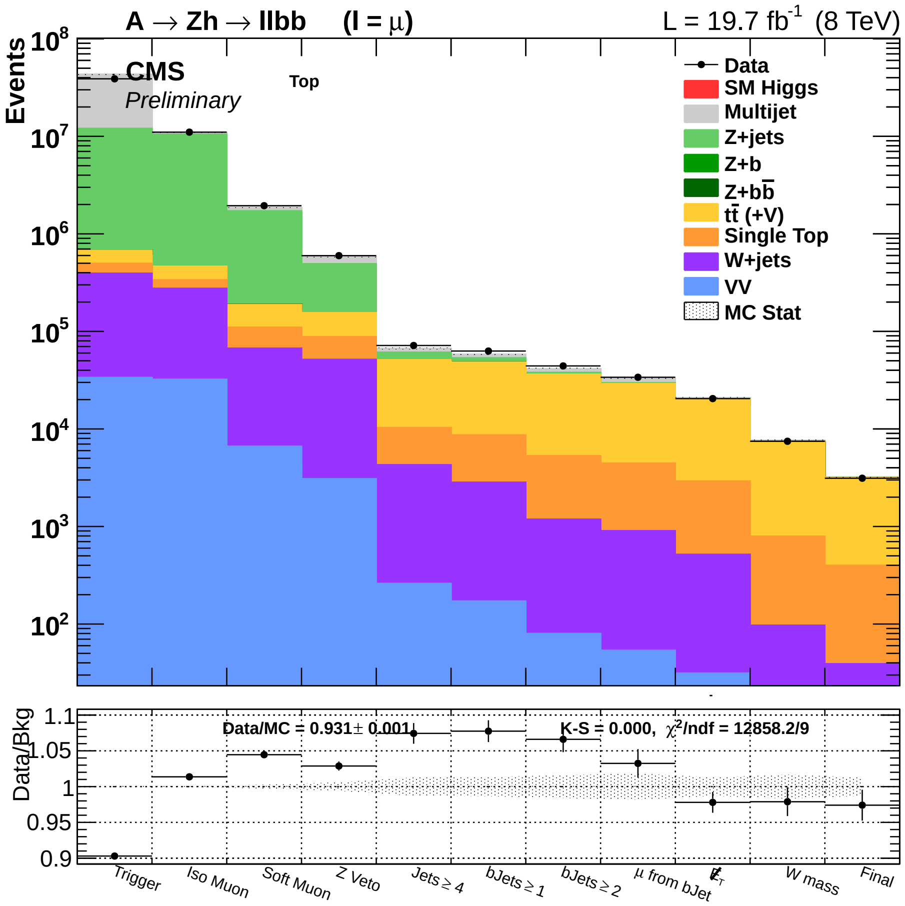

Figure B.1: Number of events for the top quark test sample after each cut.

B.1 JET ENERGY SCALE AND RESOLUTION

The mass peak of the W boson and the final top quark mass are clearly visible in Fig. B.2, showing a good agreement between data and simulation. A Gaussian fit is performed on the bulk of the distributions for both data and MC, giving the compatible results reported in Tab. B.1. Peak means and widths agree within their error with data for both the W boson and the top quark, confirming the good simulation of the jet reconstruction and measurement performances of the detector. The fitted hadronic top quark mass slightly lower than the expected value, due to the semi-leptonic B-hadron decay which entails the presence of a neutrino inside one of the two b-tagged jets. This implies an undermeasured p_T for this jet, shifting the mass peak to lower values.

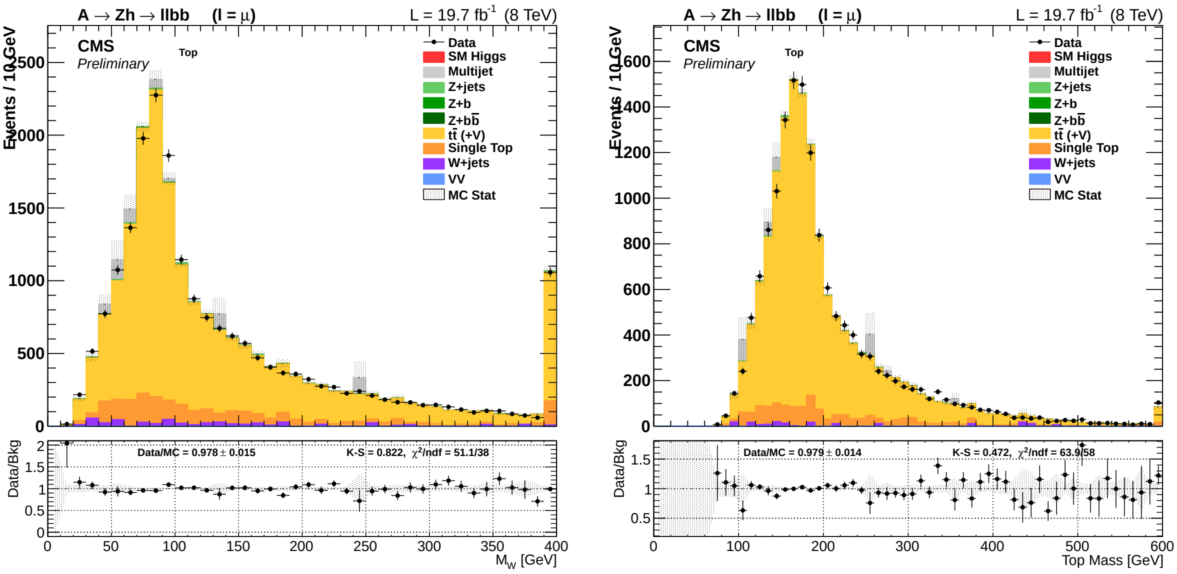

Figure B.2: W boson mass peak (left) and top quark mass (right) mass peaks in the selected top quark test sample; Gaussian fits to data and MC are superimposed to the histograms, and fit results are listed in table B.1. The right plot includes multiple combinations.

	W boson mass		Top quark mass	
	Mean (GeV)	Sigma (GeV)	Mean (GeV)	Sigma (GeV)
Data	83.6 ± 0.3	18.4 ± 0.4	167.6 ± 0.5	25.0 ± 0.7
Monte Carlo	82.7 ± 0.2	18.1 ± 0.3	166.9 ± 0.5	25.7 ± 0.6

Table B.1: Results of the Gaussian fit to the W and top quark mass peaks, as shown in Fig. B.2.

B.2 KINEMATIC FIT

The A → Zh → $\ell\ell b\bar{b}$ analysis performs a kinematic fit on the two jets originating from the light Higgs h, in order to enhance the signal significance by increasing the peak resolution. There is interest to demonstrate that this technique is also valid for a generic resonance. The kinematic fit is applied to the two jets identified as coming from the W boson to force the dijet invariant mass of the $m_W = 80.1\,\mathrm{GeV}$. The parametrization of the p_T, η and φ resolution is the same used in the A → Zh → $\ell\ell b\bar{b}$ analysis. In Fig. B.3, the hadronic top quark mass is shown before and after the kinematic fit; the increased peak resolution is noticeable, with the peak width reduced from 25 GeV to 20 GeV, while the mean remains almost unchanged. The agreement with data also improves, given the smaller χ^2 and larger Kolmogorov Test p-value.

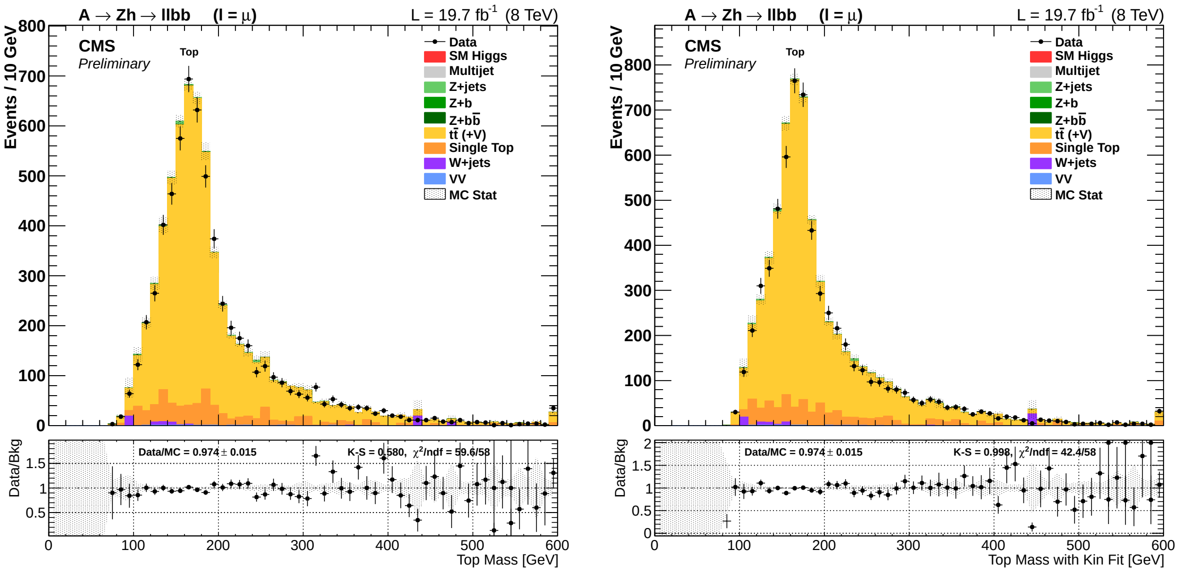

Figure B.3: W boson mass peak (left) and top quark mass (right) mass peaks in the selected top quark test sample; Gaussian fits to data and MC are superimposed to the histograms, and fit results are listed table B.2

	Before Kinematic Fit		After Kinematic Fit	
	Mean (GeV)	Sigma (GeV)	Mean (GeV)	Sigma (GeV)
Data	167.0 ± 0.8	25.5 ± 1.3	165.7 ± 0.6	20.9 ± 0.7
Monte Carlo	166.3 ± 0.5	25.2 ± 0.7	165.3 ± 0.3	20.8 ± 0.4

Table B.2: Results of the Gaussian fit to the W and top quark mass peaks, as shown in Fig. B.3.

B.3 COLOUR VARIABLES

The BDT used in the main analysis is built on 16 variables, and some are used seldom in physics analyses. One of them is the pull angle, a variable sensitive to the colour flow between the two quarks originating from a colourless objects and described in detail in Sec. 5.4.4. Even if the colour connection has been experimentally observed at LEP [156] [157], an additional cross check is useful to check for the presence of possible biases. Furthermore, there is the need to compare the Monte Carlo distribution with data before giving the variable as input to the analysis discriminator. Considering the two jets identified as originating from the W boson, the pull angle (the pull vector angle between the two jets in the $y - \varphi$ plane) is calculated for both, as well as the so-called beam angle (the pull vector angle between a single jet and the closest beam in the $y - \varphi$ plane). The pull angle distribution is expected to peak at 0, wile the beam angle should have the opposite behavior. Plots in Fig. B.4 display this trend and show also a good agreement between data and simulation. The leading jet has the most distinctive peak, while the same effect in the sub-leading one is less discriminating; this is the reason because only this first one is provided to the analysis discriminator.

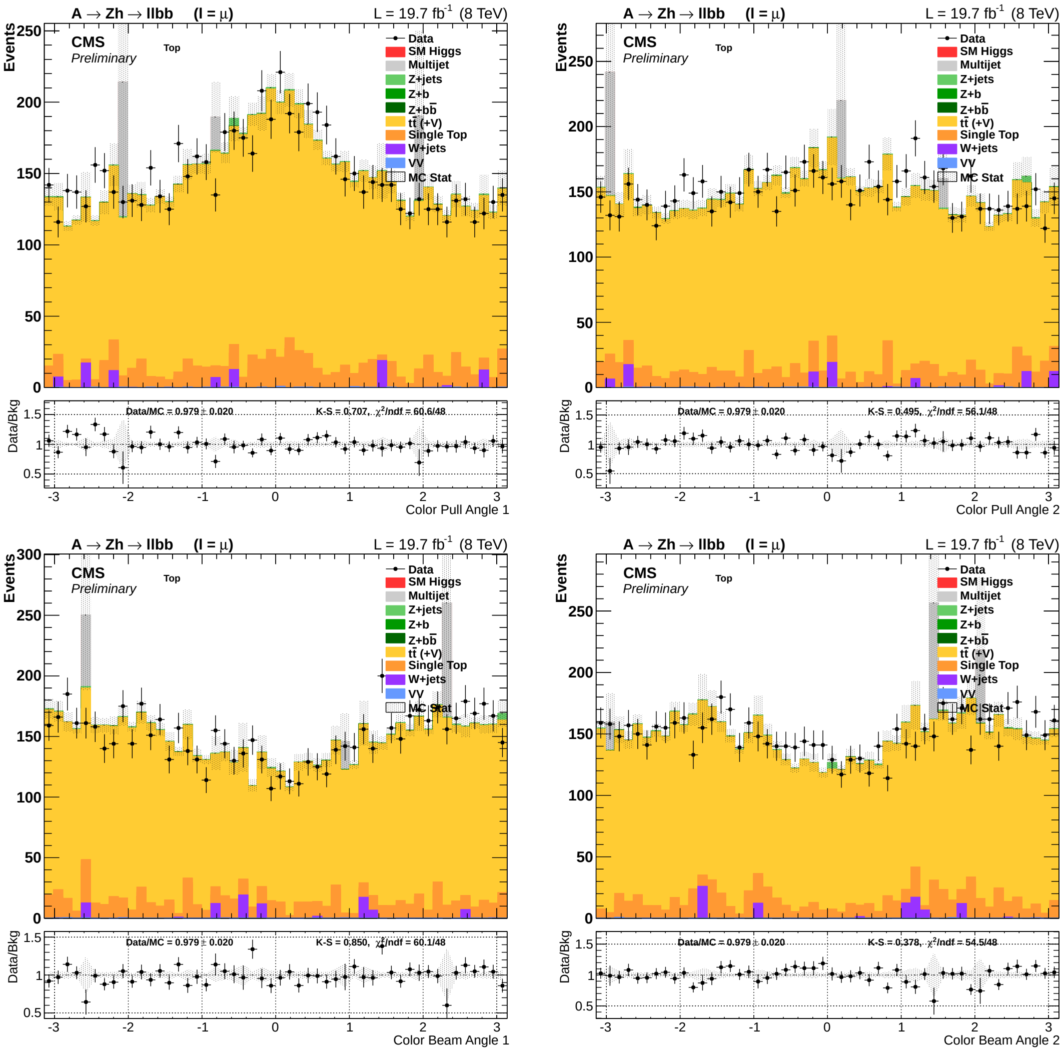

Figure B.4: Colour pull angles (top) and beam angles (bottom) for the leading (left) and sub-leading (right) b-tagged jet in the event.

Appendix C

Search for a standard model diboson signal

C.1 EVIDENCE OF THE ZZ DECAY INTO LEPTONS AND B QUARKS

The Boosted Decision Tree is a powerful yet complex multivariate technique (MVA), and as such has to be tested and validated on a dedicated and separate sample. The simultaneous production of two Z bosons, with one of them decaying in a lepton pair and the other one in a pair of b quarks, is an useful benchmark to extract a signal of a known resonance ($\mathrm{Z} \rightarrow \mathrm{b}\bar{\mathrm{b}}$) with a larger cross section but the same final state of the $\mathrm{A} \rightarrow \mathrm{Zh} \rightarrow \ell\ell\mathrm{b}\bar{\mathrm{b}}$ analysis. This search has already been performed by CMS, adding the $\mathrm{Z} \rightarrow \nu\nu$ and the W Z channels in the boosted regime, yielding a significance of over 6σ and a cross section in agreement with the standard model expectations [158]. Since the present analysis is not performed in the boosted regime, there is the interest to extract a signature of the Z Z production also in a low-p_{T} range, in spite the smaller expected sensitivity.

The BDT used for this purpose is the same of the main analysis, sharing the same preselections and variables, except for the training, where the diboson process is considered as the signal. Systematic uncertainties are computed as in the main analysis, and added as nuisances parameters in the signal fit. It is important to note that no cut on the Z p_{T} (leptonic or hadronic) is applied. The cut on the dijet invariant mass is set around the Z boson mass (60 GeV and 110 GeV). Figure C.1 shows the BDT output, and Table C.1 lists the expected and observed fit results on the BDT shape. The results are in good agreement with the standard model expected cross section, and the absence of the $\mathrm{ZZ} \rightarrow \ell\ell\mathrm{b}\bar{\mathrm{b}}$ signal is highly disfavored.

Even if the Z Z total cross section is several times larger than the expected $\mathrm{A} \rightarrow \mathrm{Zh} \rightarrow \ell\ell\mathrm{b}\bar{\mathrm{b}}$ signal, diboson events lie in a quite different phase space, being a process initiated by a $q\bar{q}$ state while the A is produced via gluon-gluon fusion. The consequence is that background is more difficult to discriminate, due to its similarity to the Z Z signal in several distributions but $\mathrm{Z}p_{\mathrm{T}}$ and m_{jj}. The variables chosen are not optimized for the diboson search, and only moving to the boosted regime can improve the signal over background ratio and increase the significance. The kinematic fit is also applied to the h mass, which is different from the Z mass, and this a source of bias; the fit χ^2 is not as useful as m_{jj} in this frame. However, this test does not represent a proper Z Z cross section measurement, which is out of the scope on this analysis, being interested mainly in the BDT validation.

Even if in a boosted regime, the $\mathrm{ZZ} \rightarrow \ell\ell\mathrm{b\bar{b}}$ signal can be extracted through a cut-based analysis. The event selection is the same as the signal region, defined in Sec. 5.3.1, but without the m_{bb} cut and with the addition of a cut on the Z $p_T > 100\,\mathrm{GeV}$. This kind of event selection selects a region of the phase space quite far from the analysis interest, but the simplest way to look for such a signal. Without the boosting requirement, a cut-based approach could be not feasible. Another important difference is that the cut selection does not introduce a bias on the dijet invariant mass spectrum, and the $\mathrm{Z} \rightarrow \mathrm{b\bar{b}}$ peak should be observable. Fig. C.1 shows the dijet invariant mass after the selection cuts, and Tab. C.1 the numerical values returned by the fit on the m_{jj} mass.

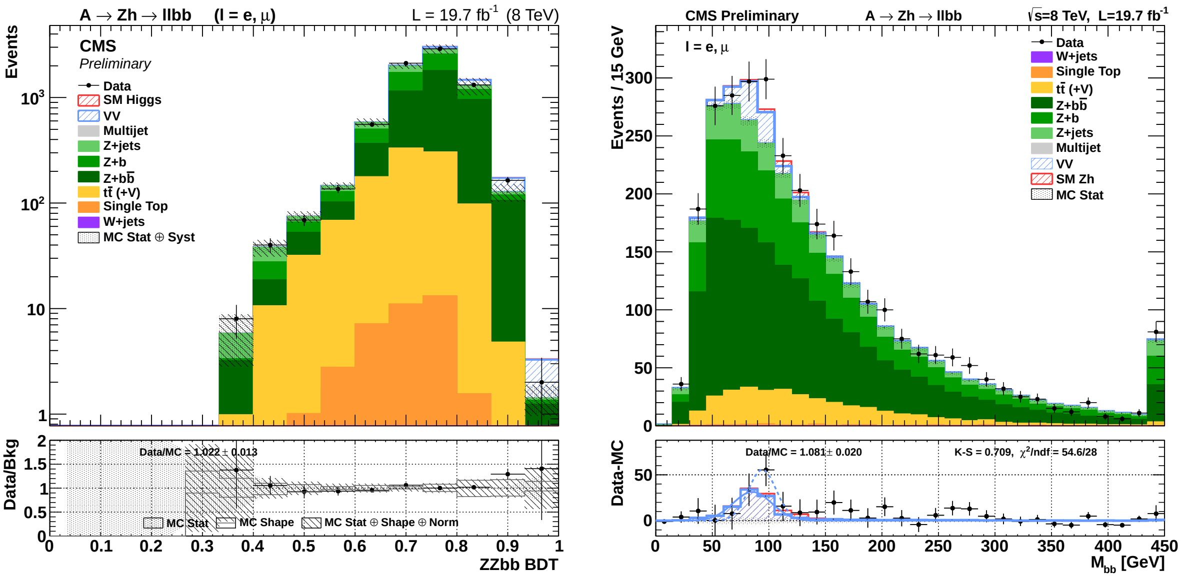

Figure C.1: Left: $ZZ \rightarrow \ell\ell\mathrm{b\bar{b}}$ BDT output. The diboson process is trained as signal, and the m_{bb} mass cut applied is $60\,\mathrm{GeV} < m_{\mathrm{bb}} < 110\,\mathrm{GeV}$. The diboson process is displayed in light blue. Right: dijet invariant mass spectrum in the signal region, without the m_{bb} cut and with Z $p_T > 100\,\mathrm{GeV}$. The diboson process is displayed in light blue, and the SM Higgs in red.

	BDT analysis		Cut-based analysis	
	Significance	Best Fit μ	Significance	Best Fit μ
Observed	2.94σ	$1.05^{+0.49}_{-0.40}$	2.29σ	$1.01^{+0.47}_{-0.44}$
Expected	2.96σ	$1\ ^{+0.45}_{-0.38}$	2.31σ	$1\ ^{+0.47}_{-0.44}$

Table C.1: Expected and observed significance and best fit for the $\mathrm{ZZ} \rightarrow \ell\ell\mathrm{b\bar{b}}$ signal.

C.2 SENSITIVITY TO THE SM HIGGS BOSON

The same strategy can be applied also to the search of the standard model Higgs boson produced in association to a Z boson. The kinematic fit forces the dijet invariant mass to the nominal h value, while the fit χ^2 can still discriminate against the other backgrounds. Also this test search is performed in the non-boosted regime, while the standard analysis [159] starts from a Z $p_T > 50\,\mathrm{GeV}$ cut, and can an useful cross-check. The lower sensitivity with respect to the SM search is mainly due to the following factors:

- The present search is performed in the non-boosted regime, where the S/B is significantly worse than the boosted case
- The kinematic fit does not allow to exploit the power of the m_{bb} variable, which is the most discriminating for this kind of search
- There is no b jet energy regression
- The Z $\to \nu\nu$ and W $\to \ell\nu$ channels are not considered
- The variable choice and the BDT is not optimized for the SM search

The BDT is trained with the SM Zh $\to \ell\ell \mathrm{b}\bar{\mathrm{b}}$ as signal, and a shape fit is performed on the BDT output; variables and preselections are the same as the main analysis. The BDT shape is shown in Fig. C.2, with signal fit results are reported in Tab. C.2. In spite a large statistical uncertainty, some sensitivity to the SM Higgs boson is present. The observed significance exceeds 1σ, an excess not statistically significant, favoring the presence of the Higgs with respect to the null hypothesis.

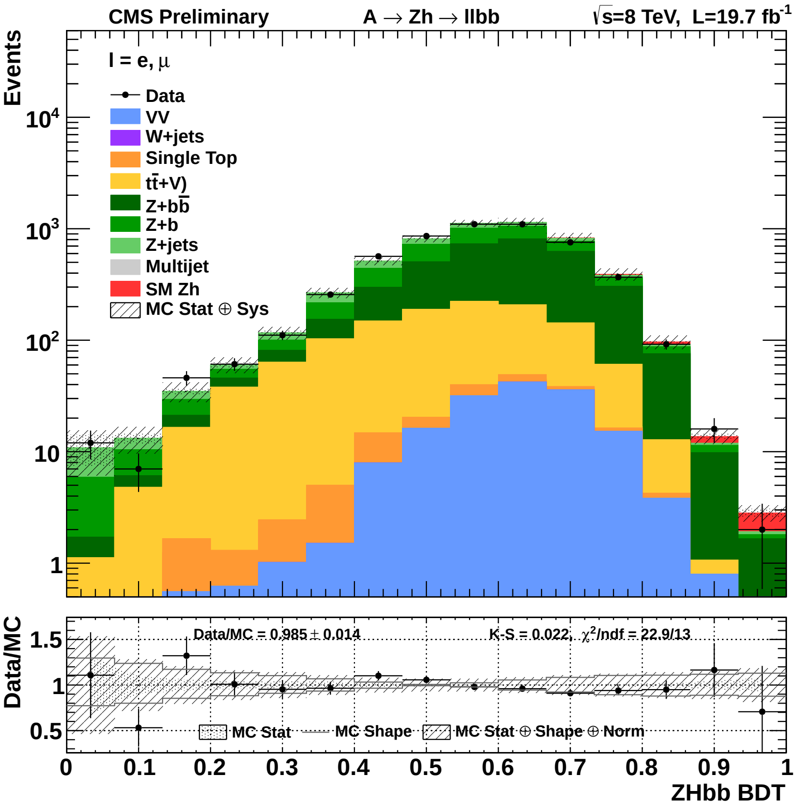

Figure C.2: Zh $\to \ell\ell \mathrm{b}\bar{\mathrm{b}}$ BDT output. The BDT is the same as the main analysis, but trained with the Vh process as signal.

	Exclusion	Significance	Best Fit μ
Observed	4.11	1.28σ	$1.60^{+1.43}_{-1.27}$
Expected	$2.51^{+3.66\ +5.13}_{-1.74\ -1.28}$	0.85σ	$1^{+1.34}_{-1.00}$

Table C.2: Results of the maximum likelihood fit on the Z h BDT output. Expected exclusion upper limits are reported with the ±1 and $\pm 2\sigma$ band uncertainties.

References

[1] S. L. Glashow, "Partial-symmetries of weak interactions", *Nuclear Physics* **22** (1961), no. 4, 579 – 588, `doi:http://dx.doi.org/10.1016/0029-5582(61)90469-2`. 9

[2] S. Weinberg, "A Model of Leptons", *Phys. Rev. Lett.* **19** (Nov, 1967) 1264–1266, `doi:10.1103/PhysRevLett.19.1264`. 9

[3] A. Salam, "Weak and Electromagnetic Interactions", *Conf.Proc.* **C680519** (1968) 367–377. 9

[4] G. S. Franz Mandl, "Quantum Field Theory". John Wiley and Sons, 1984. 9

[5] M. Gell-Mann, "A schematic model of baryons and mesons", *Physics Letters* **8** (1964), no. 3, 214 – 215, `doi:http://dx.doi.org/10.1016/S0031-9163(64)92001-3`. 9

[6] G. Zweig, "An SU_3 model for strong interaction symmetry and its breaking; Version 1", Technical Report CERN-TH-401, CERN, Geneva, (Jan, 1964). 9

[7] C. Rubbia, "Experimental Observation of the Intermediate Vector Bosons W+, W-, and Z0", *Rev.Mod.Phys.* **57** (1985) 699–722, `doi:10.1103/RevModPhys.57.699`. 10

[8] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons", *Phys. Rev. Lett.* **13** (Aug, 1964) 321–323, `doi:10.1103/PhysRevLett.13.321`. 11

[9] P. Higgs, "Broken Symmetries and the Masses of Gauge Bosons", *Phys. Rev. Lett.* **13** (Oct, 1964) 508–509, `doi:10.1103/PhysRevLett.13.508`.

[10] G. Guralnik, C. Hagen, and T. Kibble, "Global Conservation Laws and Massless Particles", *Phys. Rev. Lett.* **13** (Nov, 1964) 585–587, `doi:10.1103/PhysRevLett.13.585`. 11

[11] J. Goldstone, A. Salam, and S. Weinberg, "Broken Symmetries", *Phys. Rev.* **127** (Aug, 1962) 965–970, `doi:10.1103/PhysRev.127.965`. 11

[12] LHC Higgs Cross Section Working Group, S. Heinemeyer, C. Mariotti et al., "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties", *CERN-2013-004* (CERN, Geneva, 2013) `arXiv:1307.1347`. 13, 14, 82, 83, 84

[13] S. L. Adler, R. F. Dashen, and S. Treiman, "Comments on Proposed Explanations for the mu - Mesic Atom x-Ray Discrepancy", *Phys.Rev.* **D10** (1974) 3728, `doi:10.1103/PhysRevD.10.3728`. 15

[14] K. Sato and H. Sato, "Higgs Meson Emission from a Star and a Constraint on Its Mass", *Prog.Theor.Phys.* **54** (1975) 1564–1565, `doi:10.1143/PTP.54.1564`. 15

[15] L. Resnick, M. Sundaresan, and P. Watson, "Is There a Light Scalar Boson?", *Phys. Rev. D* **8** (Jul, 1973) 172–178, `doi:10.1103/PhysRevD.8.172`. 15

[16] D. Kohler, B. Watson, and J. Becker, "Experimental Search for a Low-Mass Scalar Boson", *Phys. Rev. Lett.* **33** (Dec, 1974) 1628–1631, `doi:10.1103/PhysRevLett.33.1628`. 15

[17] R. Barbieri and T. Ericson, "Evidence against the existence of a low mass scalar boson from neutron-nucleus scattering", *Physics Letters B* **57** (1975), no. 3, 270 – 272, `doi:http://dx.doi.org/10.1016/0370-2693(75)90073-8`. 15

[18] R. Assmann, M. Lamont, and S. Myers, "A brief history of the LEP collider", *Nuclear Physics B - Proceedings Supplements* **109** (2002), no. 2–3, 17 – 31, `doi:http://dx.doi.org/10.1016/S0920-5632(02)90005-8`. Proceedings of the 7th Topical Seminar. 15

[19] LEP Working Group for Higgs boson searches, ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, "Search for the standard model Higgs boson at LEP", *Phys.Lett.* **B565** (2003) 61–75, `doi:10.1016/S0370-2693(03)00614-2`, `arXiv:hep-ex/0306033`. 15, 16, 29, 32

[20] ALEPH Collaboration, "Observation of an excess in the search for the standard model Higgs boson at ALEPH", *Phys.Lett.* **B495** (2000) 1–17, `doi:10.1016/S0370-2693(00)01269-7`, `arXiv:hep-ex/0011045`. 15

[21] LEP Higgs Working for Higgs boson searches, ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, "Searches for invisible Higgs bosons: Preliminary combined results using LEP data collected at energies up to 209-GeV", `arXiv:hep-ex/0107032`. 15

[22] M. S. Carena, P. M. Zerwas, E. Accomando et al., "Higgs physics at LEP-2", `arXiv:hep-ph/9602250`. 15

[23] CDF Collaboration, "Observation of Top Quark Production in $\overline{p}p$ Collisions with the Collider Detector at Fermilab", *Phys. Rev. Lett.* **74** (Apr, 1995) 2626–2631, `doi:10.1103/PhysRevLett.74.2626`. 15

[24] J. Beringer and others, "Particle Data Group", *Phys. Rev.* **D86** (2012). 15, 70

[25] CDF and D0 Collaboration, "Higgs boson studies at the Tevatron", *Phys. Rev. D* **88** (Sep, 2013) 052014, `doi:10.1103/PhysRevD.88.052014`. 16, 17, 32

[26] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys.Lett.* **B716** (2012) 30–61, `doi:10.1016/j.physletb.2012.08.021`, `arXiv:1207.7235`. 17

[27] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV", *JHEP* **1306** (2013) 081, `doi:10.1007/JHEP06(2013)081`, `arXiv:1303.4571`. 17

[28] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys.Lett.* **B716** (2012) 1–29, `doi:10.1016/j.physletb.2012.08.020`, `arXiv:1207.7214`. 17

[29] CMS Collaboration, "Measurement of the properties of a Higgs boson in the four-lepton final state", *Phys. Rev. D* **89** (May, 2014) 092007, `doi:10.1103/PhysRevD.89.092007`. 17

[30] CMS Collaboration, "Observation of the diphoton decay of the Higgs boson and measurement of its properties", *The European Physical Journal C* **74** (2014), no. 10, `doi:10.1140/epjc/s10052-014-3076-z`. 18

[31] CMS Collaboration, "Measurement of Higgs boson production and properties in the WW decay channel with leptonic final states", *Journal of High Energy Physics* **2014** (2014), no. 1, `doi:10.1007/JHEP01(2014)096`. 18

[32] CMS Collaboration, "Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons", *Journal of High Energy Physics* **2014** (2014), no. 5, `doi:10.1007/JHEP05(2014)104`. 19

[33] CMS Collaboration, "Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks", *Phys. Rev. D* **89** (Jan, 2014) 012003, `doi:10.1103/PhysRevD.89.012003`. 20, 50, 102

[34] CMS Collaboration, "Precise determination of the mass of the Higgs boson and studies of the compatibility of its couplings with the standard model", Technical Report CMS-PAS-HIG-14-009, CERN, Geneva, (2014). 21, 22

[35] CMS Collaboration, "Constraints on the Higgs boson width from off-shell production and decay to Z-boson pairs", *Physics Letters B* **736** (2014), no. 0, 64 – 85, `doi:http://dx.doi.org/10.1016/j.physletb.2014.06.077`. 22

[36] CMS Collaboration, "Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV", `arXiv:1411.3441`. 22

[37] Planck Collaboration, "Planck 2013 results. XVI. Cosmological parameters", *Astron.Astrophys.* **571** (2014) A16, `doi:10.1051/0004-6361/201321591`, `arXiv:1303.5076`. 23, 25

[38] B. S. Acharya, G. Kane, S. Watson et al., "Nonthermal "WIMP miracle"", *Phys. Rev. D* **80** (Oct, 2009) 083529, `doi:10.1103/PhysRevD.80.083529`. 24

[39] Y. Fukuda, T. Hayakawa, E. Ichihara et al., "Evidence for oscillation of atmospheric neutrinos", *Physical Review Letters* **81** (1998), no. 8, 1562. 24

[40] LSND Collaboration, "Evidence for neutrino oscillations from the observation of $\bar{\nu}_e$ appearance in a $\bar{\nu}_\mu$ beam", *Phys. Rev. D* **64** (Nov, 2001) 112007, `doi:10.1103/PhysRevD.64.112007`. 24

[41] S. Fraser, E. Ma, and O. Popov, "Scotogenic Inverse Seesaw Model of Neutrino Mass", *Phys.Lett.* **B737** (2014) 280–282, `arXiv:1408.4785`. 24

[42] J. Martin, "Everything you always wanted to know about the cosmological constant problem (but were afraid to ask)", *Comptes Rendus Physique* **13** (July, 2012) 566–665, `doi:10.1016/j.crhy.2012.04.008`, `arXiv:1205.3365`. 25

[43] J. Elias-Miro, J. R. Espinosa, G. F. Giudice et al., "Higgs mass implications on the stability of the electroweak vacuum", *Phys.Lett.* **B709** (2012) 222–228, `doi:10.1016/j.physletb.2012.02.013`, `arXiv:1112.3022`. 25

[44] G. Branco, P. Ferreira, L. Lavoura et al., "Theory and phenomenology of two-Higgs-doublet models", *Physics Reports* **516** (2012), no. 1-2, 1 – 102, `doi:http://dx.doi.org/10.1016/j.physrep.2012.02.002`. Theory and phenomenology of two-Higgs-doublet models. 25

[45] N. Craig and S. Thomas, "Exclusive Signals of an Extended Higgs Sector", *JHEP* **1211** (2012) 083, `doi:10.1007/JHEP11(2012)083`, `arXiv:1207.4835`. 25

[46] G. Costa, J. Ellis, G. Fogli et al., "Neutral currents within beyond the standard model", *Nuclear Physics B* **297** (1988), no. 2, 244 – 286, `doi:http://dx.doi.org/10.1016/0550-3213(88)90020-X`. 25

[47] A. G. Cohen, D. Kaplan, and A. Nelson, "Spontaneous baryogenesis at the weak phase transition", *Phys.Lett.* **B263** (1991) 86–92, `doi:10.1016/0370-2693(91)91711-4`. 27

[48] Z. Maki, M. Nakagawa, and S. Sakata, "Remarks on the Unified Model of Elementary Particles", *Progress of Theoretical Physics* **28** (1962), no. 5, 870–880, `doi:10.1143/PTP.28.870`, `arXiv:http://ptp.oxfordjournals.org/content/28/5/870.full.pdf+html`. 27

[49] U. Amaldi, A. Bohm, L. Durkin et al., "A Comprehensive Analysis of Data Pertaining to the Weak Neutral Current and the Intermediate Vector Boson Masses", *Phys.Rev.* **D36** (1987) 1385, `doi:10.1103/PhysRevD.36.1385`. 27

[50] G. Costa, J. R. Ellis, G. L. Fogli et al., "Neutral Currents Within and Beyond the Standard Model", *Nucl.Phys.* **B297** (1988) 244–286, `doi:10.1016/0550-3213(88)90020-X`. 27

[51] S. Glashow and S. Weinberg, "Natural conservation laws for neutral currents", *Phys. Rev. D* **15** (Apr, 1977) 1958–1965, `doi:10.1103/PhysRevD.15.1958`. 27

[52] J. F. Gunion, H. E. Haber, G. L. Kane et al., "The Higgs Hunter's Guide", *Front.Phys.* **80** (2000) 1–448. 27

[53] S. P. Martin, "A Supersymmetry primer", *Adv.Ser.Direct.High Energy Phys.* **21** (2010) 1–153, `doi:10.1142/9789814307505_0001`, `arXiv:hep-ph/9709356`. 28

[54] M. S. Carena, S. Heinemeyer, C. Wagner et al., "Suggestions for improved benchmark scenarios for Higgs boson searches at LEP-2", `arXiv:hep-ph/9912223`. 29

[55] M. Carena, S. Heinemeyer, O. Stål et al., "MSSM Higgs boson searches at the LHC: benchmark scenarios after the discovery of a Higgs-like particle", *The European Physical Journal C* **73** (2013), no. 9, `doi:10.1140/epjc/s10052-013-2552-1`. 29, 31

[56] ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, LEP Working Group for Higgs Boson Searches Collaboration, "Search for neutral MSSM Higgs bosons at LEP", *Eur.Phys.J.* **C47** (2006) 547–587, `doi:10.1140/epjc/s2006-02569-7`, `arXiv:hep-ex/0602042`. 31, 75, 76

[57] D0 Collaboration, "Search for neutral Minimal Supersymmetric Standard Model Higgs bosons decaying to tau pairs produced in association with b quarks in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV", *Phys.Rev.Lett.* **107** (2011) 121801, `doi:10.1103/PhysRevLett.107.121801`, `arXiv:1106.4885`. 32

[58] CDF Collaboration, "Search for Neutral MSSM Higgs Bosons Decaying to Tau Pairs with 1.8 fb^{-1} of data", CDF Note 9071, FNAL, (2007). 32

[59] CDF Collaboration, "Search for Higgs bosons produced in association with b quarks", *Phys. Rev. D* **85** (Feb, 2012) 032005, `doi:10.1103/PhysRevD.85.032005`. 32, 74, 75, 76

[60] D0 Collaboration, "Search for neutral Higgs bosons in the multi-b-jet topology in collisions at $\sqrt{s} = 1.96$ TeV", *Physics Letters B* **698** (2011), no. 2, 97 – 104, `doi:http://dx.doi.org/10.1016/j.physletb.2011.02.062`. 32

[61] CMS Collaboration, "Search for neutral MSSM Higgs bosons decaying to a pair of tau leptons in pp collisions", *Journal of High Energy Physics* **2014** (2014), no. 10, `doi:10.1007/JHEP10(2014)160`. 33, 34

[62] O. S. Brüning, P. Collier, P. Lebrun et al., "LHC Design Report". CERN, Geneva, 2004. 35

[63] LEP TDR Working Group, "LEP design report". CERN, Geneva, 1984. Copies shelved as reports in LEP, PS and SPS libraries. 35

[64] TOTEM Collaboration, "Luminosity-Independent Measurement of the Proton-Proton Total Cross Section at $\sqrt{s} = 8$ TeV", *Phys. Rev. Lett.* **111** (Jul, 2013) 012001, `doi:10.1103/PhysRevLett.111.012001`. 36

[65] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, `doi:10.1088/1748-0221/3/08/S08004`. 37, 38

[66] CMS Collaboration, G. L. Bayatian, S. Chatrchyan, et al., "CMS Physics: Technical Design Report Volume 1: Detector Performance and Software". Technical Design Report CMS. CERN, Geneva, 2006. 37

[67] CMS Collaboration, "Precise mapping of the magnetic field in the CMS barrel yoke using cosmic rays", *Journal of Instrumentation* **5** (2010), no. 03, T03021. 37

[68] CMS Collaboration, S. Cittolin, A. Rácz, and P. Sphicas, "CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project". Technical Design Report CMS. CERN, Geneva, 2002. 38

[69] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and E_T^{miss}", *CMS-PAS-PFT-09-001* (2009). 39, 53

[70] CMS Collaboration, "Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector", *CMS-PAS-PFT-10-001* (2010). 39

[71] CMS Collaboration, V. Karimäki, M. Mannelli, P. Siegrist et al., "The CMS tracker system project: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 1997. 39

[72] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST* **9** (2014), no. 10, P10009, `doi:10.1088/1748-0221/9/10/P10009`, `arXiv:1405.6569`. 40, 41

[73] R. Frühwirth, "Application of Kalman filtering to track and vertex fitting", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **262** (1987), no. 2–3, 444 – 450, `doi:http://dx.doi.org/10.1016/0168-9002(87)90887-4`. 40, 51

[74] T. Speer, K. Prokofiev, R. Frühwirth et al., "Vertex Fitting in the CMS Tracker", Technical Report CMS-NOTE-2006-032, CERN, Geneva, (Feb, 2006). 41

[75] K. Rose, "Deterministic Annealing for Clustering, Compression, Classification, Regression, and Related Optimization Problems", in *Proceedings of the IEEE*, pp. 2210–2239. 1998. 41

[76] W. Waltenberger, R. Frühwirth, and P. Vanlaer, "Adaptive vertex fitting", *Journal of Physics G: Nuclear and Particle Physics* **34** (2007), no. 12, N343. 41

[77] CMS Collaboration, "Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp Collisions at $\sqrt{s} = 7$ TeV", *JINST* **8** (2013) P09009, `doi:10.1088/1748-0221/8/09/P09009`, `arXiv:1306.2016`. 42, 43

[78] W. Adam, R. Frühwirth, A. Strandlie et al., "Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC", *Journal of Physics G: Nuclear and Particle Physics* **31** (2005), no. 9, N9. 43

[79] CMS Collaboration, C. Collaboration, "The CMS hadron calorimeter project: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 1997. 44

[80] M. Cacciari, G. Salam, and G. Soyez, "FastJet user manual", *The European Physical Journal C* **72** (2012), no. 3, `doi:10.1140/epjc/s10052-012-1896-2`. 45

[81] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-kt jet clustering algorithm", *Journal of High Energy Physics* **2008** (2008), no. 04, 063. 45

[82] M. Cacciari and G. P. Salam, "Dispelling the myth for the jet-finder", *Physics Letters B* **641** (2006), no. 1, 57 – 61, `doi:http://dx.doi.org/10.1016/j.physletb.2006.08.037`. 45

[83] S. Ellis and D. Soper, "Successive combination jet algorithm for hadron collisions", *Phys. Rev. D* **48** (Oct, 1993) 3160–3166, `doi:10.1103/PhysRevD.48.3160`. 45

[84] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", *Journal of Instrumentation* **6** (2011), no. 11, P11002. 46

[85] R. Salam, "FastJet package", technical report, CNRS, (2011). 46

[86] CMS Collaboration, "Performance of b tagging at sqrt(s)=8 TeV in multijet, ttbar and boosted topology events", Technical Report CMS-PAS-BTV-13-001, CERN, Geneva, (2013). 49, 50, 114

[87] CMS Collaboration, CMS, "The CMS muon project: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 1997. 50

[88] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at sqrt(s) = 7 TeV", *Journal of Instrumentation* **7** (2012), no. 10, P10002. 51, 52, 53, 72, 113

[89] CMS Collaboration, "Calibration of track momentum using dimuon resonances in CMS", CMS Analysis Note CMS AN-2010/059, CMS, (2010). 52

[90] CMS Collaboration, "Missing transverse energy performance of the CMS detector", *Journal of Instrumentation* **6** (2011), no. 09, P09001. 53, 54

[91] CMS Collaboration, "MET performance in 8 TeV data", Technical Report CMS-PAS-JME-12-002, CERN, Geneva, (2013). 54

[92] CMS Collaboration Collaboration, "Search for a Higgs boson decaying into a b-quark pair and produced in association with b quarks in proton-proton collisions at 7 TeV", *Phys.Lett.* **B722** (2013) 207–232, `doi:10.1016/j.physletb.2013.04.017`, `arXiv:1302.2892`. 55

[93] CMS Collaboration Collaboration, "MSSM Higgs production in association with b-quarks - semi leptonic", Technical Report CMS-PAS-HIG-12-027, CERN, Geneva, (2012). 55

[94] R. V. Harlander and W. B. Kilgore, "Higgs boson production in bottom quark fusion at next-to-next-to-leading order", *Phys. Rev. D* **68** (Jul, 2003) 013001, `doi:10.1103/PhysRevD.68.013001`. 55

[95] S. Heinemeyer, W. Hollik, and G. Weiglein, "FeynHiggs: a program for the calculation of the masses of the neutral CP-even Higgs bosons in the MSSM", *Computer Physics Communications* **124** (2000), no. 1, 76 – 89, `doi:10.1016/S0010-4655(99)00364-1`. 55

[96] S. Heinemeyer, W. Hollik, and G. Weiglein, "The masses of the neutral CP-even Higgs bosons in the MSSM: Accurate analysis at the two-loop level", *The European Physical Journal C - Particles and Fields* **9** (1999) 343–366. 10.1007/s100529900006.

[97] G. Degrassi, S. Heinemeyer, W. Hollik et al., "Towards high-precision predictions for the MSSM Higgs sector", *The European Physical Journal C - Particles and Fields* **28** (2003) 133–143. 10.1140/epjc/s2003-01152-2.

[98] M. Frank, T. Hahn, S. Heinemeyer et al., "The Higgs boson masses and mixings of the complex MSSM in the Feynman-diagrammatic approach", *Journal of High Energy Physics* **2007** (2007), no. 02, 047. 55

[99] CMS Collaboration, "Identification of b-quark jets with the CMS experiment", *Journal of Instrumentation* **8** (2013), no. 04, P04013. 64, 72

[100] CMS Collaboration, "Jet Performance in pp Collisions at $\sqrt{s}$=7 TeV", CMS Physics Analysis Summary CMS-PAS-JME-10-003, CERN, (2010). 72

[101] CMS Collaboration, "Absolute Calibration of the Luminosity Measurement at CMS: Winter 2012 Update", Technical Report CMS-PAS-SMP-12-008, CERN, Geneva, (2012). 72

[102] R. Ellis, W. Stirling, and B. Webber, “QCD and Collider Physics”. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 2003. 73

[103] PDF4LHC Working Group, “LHAPDF: the Les Houches Accord PDF Interface”, technical report, HepForge, (2013). 73, 117

[104] CMS Collaboration Collaboration, “MSSM Higgs production in association with b quarks - all hadronic”, Technical Report CMS-PAS-HIG-12-026, CERN, Geneva, (2012). 73

[105] A. L. Read, “Presentation of search results: the CLs technique”, *J. Phys.* **G28** (2002) 2693, `doi:10.1088/0954-3899/28/10/313`. 75, 97, 122, 133, 134

[106] T. Junk, “Confidence level computation for combining searches with small statistics”, *Nucl. Instrum. Meth.* **A434** (1999) 435, `arXiv:9902006`. 75, 97, 122, 133

[107] L. Moneta, K. Belasco, K. Cranmer et al., “The RooStats Project”, *PoS(ACAT2010)057* (2010). 75, 122

[108] CMS Collaboration Collaboration, “Search for a pseudoscalar boson A decaying into a Z and an h boson in the llbb final state”, Technical Report CMS-PAS-HIG-14-011, CERN, Geneva, (2014). 77

[109] A. Djouadi and J. Quevillon, “The MSSM Higgs sector at a high M_{SUSY}: reopening the low tanβ regime and heavy Higgs searches”, `arXiv:1304.1787`. 78, 128

[110] CMS Collaboration, “Search for the MSSM A→Zh decay with Z→ $\ell\ell$, h→bb”, CMS Note 2006/063, CERN, (2006). 78

[111] J. Alwall, M. Herquet, F. Maltoni et al., “MadGraph 5 : Going Beyond”, *JHEP* **1106** (2011) 128, `arXiv:1106.0522`. 80, 83

[112] CMS Collaboration, “Charged particle multiplicities in pp interactions at $\sqrt{s} = 0.9, 2.36$, and 7 TeV”, *J. High Energy Phys* **2013** (2013), no. 4, `doi:10.1007/JHEP04(2013)072`. 80

[113] J. Pumplin, D. R. Stump, J. Huston et al., “New Generation of Parton Distributions with Uncertainties from Global QCD Analysis”, *Journal of High Energy Physics* **2002** (2002), no. 07, 012. 80

[114] S. Agostinelli et al., “Geant4—a simulation toolkit”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003), no. 3, 250 – 303, `doi:10.1016/S0168-9002(03)01368-8`. 80, 83

[115] CMS Collaboration, “Inclusive W/Z cross section at 8 TeV”, CMS-PAS-SMP-12-011, CERN, (2012). 82, 83, 84

[116] J. M. Campbell and R. Ellis, “MCFM for the Tevatron and the LHC”, *Nuclear Physics B - Proceedings Supplements* **205–206** (2010), no. 0, 10 – 15, `doi:http://dx.doi.org/10.1016/j.nuclphysbps.2010.08.011`. 82

[117] CMS Collaboration, “Combination of single top-quark cross-sections measurements in the t-channel at sqrt[s]=8 TeV with the ATLAS and CMS experiments”, CMS-PAS-TOP-12-002, CERN, (2012). 82, 84

[118] CMS Collaboration, “Measurement of WW production rate”, CMS-PAS-SMP-12-013, CERN, (2012). 83

[119] CMS Collaboration, “Measurement of the ZZ production cross section and anomalous trilinear gauge couplings in lll’l’ decays at sqrt(s) = 8 TeV at the LHC”, CMS-PAS-SMP-13-005, CERN, (2012). 84

[120] CMS Collaboration, “Measurement of WZ production rate”, CMS-PAS-SMP-12-006, CERN, (2012). 84

[121] CMS Collaboration, “Measurement of W+W- and ZZ production cross sections in pp collisions at $\sqrt{s} = 8$ TeV”, CMS-SMP-12-024, CERN, (2012). 83, 84

[122] CMS Collaboration, “Measurement of the tt production cross section in the dilepton channel in pp collisions at sqrt(s)=8 TeV”, CMS-PAS-TOP-12-007, CERN, (2012). 84

[123] CMS Collaboration, “Observation of Associated Production of Single Top Quark and a W Boson in pp Collisions at $\sqrt{s} = 8$ TeV”, CMS-PAS-TOP-12-040, CERN, (2012). 84

[124] T. Sjöstrand, S. Mrenna, and P. Skands, “PYTHIA 6.4 physics and manual”, *Journal of High Energy Physics* **2006** (2006), no. 05, 026. 83

[125] S. Alioli, P. Nason, C. Oleari et al., “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”, *JHEP* **1006** (2010) 043, `doi:10.1007/JHEP06(2010)043`, `arXiv:1002.2581`. 83

[126] CMS Collaboration, “CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update”, cms analysis note, CERN, (2013). 84, 113

[127] J. Gallicchio, J. Huth, M. Kagan et al., “Multivariate discrimination and the Higgs + W/Z search”, *JHEP* **1104** (2011) 069, `doi:10.1007/JHEP04(2011)069`, `arXiv:1010.3698`. 92, 94

[128] A. Bonato, A. V. Gritsan, Z. J. Guo, A. Withbeck, “Angular Analysis of Resonances $pp \rightarrow X \rightarrow ZZ$”, CMS Analysis Note 2010/351, John Hopkins University, (2010). 93

[129] J. D. Bjorken and S. J. Brodsky, “Theory of Deep-Inelastic Lepton-Nucleon Scattering and Lepton-Pair Annihilation Processes.”, *Phys. Rev. D* **1416** (1970), no. 1,. 95

[130] J. Gallicchio and M. D. Schwartz, “Seeing in Color: Jet Superstructure”, *Phys.Rev.Lett.* **105** (2010) 022001, `doi:10.1103/PhysRevLett.105.022001`, `arXiv:1001.5027`. 96

[131] CMS Collaboration, “Expected Signal Observability at Future Experiments”, Technical Report CMS Note 2005/004, CERN, (2005). 97

[132] CMS Collaboration, “KinFitter: A Kinematic Fit with Constraints”, 2010. 102

[133] B. P. Roe, H.-J. Yang, J. Zhu et al., “Boosted decision trees as an alternative to artificial neural networks for particle identification”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **543** (2005), no. 2–3, 577 – 584, `doi:http://dx.doi.org/10.1016/j.nima.2004.12.018`. 108

[134] A. Hoecker, P. Speckmayer, J. Stelzer et al., “TMVA: Toolkit for Multivariate Data Analysis”, *PoS* **ACAT** (2007) 040, `arXiv:physics/0703039`. 108

[135] CMS Collaboration, “Electron performance with 19.6fb^{-1} of data collected at $\sqrt{s} = 8$ TeV with the CMS detector”, (2013). Cern Document Server. 113

[136] R. Barlow and C. Beeston, “Fitting using finite Monte Carlo samples”, *Computer Physics Communications* **77** (1993), no. 2, 219–228, `doi:http://dx.doi.org/10.1016/0010-4655(93)90005-W`. 115

[137] Michiel Botje, Jon Butterworth, Amanda Cooper-Sarkar, Albert de Roeck, Joel Feltesse, Stefano Forte, Alexander Glazov, Joey Huston, Ronan McNulty, Torbjorn Sjostrand, Robert Thorne, “The PDF4LHC Working Group Interim Recommendations”, technical report, CERN, (2011). 117

[138] CMS Collaboration, “Search for standard model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV”, Physics Analysis Summary CMS-PAS-HIG-11-011, CERN, (2011). 122

[139] E. Gross and O. Vitells, “Trial factors for the look elsewhere effect in high energy physics”, *European Physical Journal C* **70** (November, 2010) 525–530, `doi:10.1140/epjc/s10052-010-1470-8`, `arXiv:1005.1891`. 122, 135

[140] R. V. Harlander, S. Liebler, and H. Mantler, “SusHi: A program for the calculation of Higgs production in gluon fusion and bottom-quark annihilation in the Standard Model and the MSSM”, *Computer Physics Communications* **184** (2013), no. 6, 1605 – 1617, `doi:http://dx.doi.org/10.1016/j.cpc.2013.02.006`. 126

[141] D. Eriksson, J. Rathsman, and O. Stål, “2HDMC – two-Higgs-doublet model calculator”, *Computer Physics Communications* **181** (2010), no. 1, 189 – 205, `doi:http://dx.doi.org/10.1016/j.cpc.2009.09.011`. 126

[142] A. Martin, W. Stirling, R. Thorne et al., “Parton distributions for the LHC”, *The European Physical Journal C* **63** (2009), no. 2, 189–285, `doi:10.1140/epjc/s10052-009-1072-5`. 126

[143] E. Bagnaschi, R. Harlander, S. Liebler et al., “Towards precise predictions for Higgs-boson production in the MSSM”, `arXiv:1404.0327`. 128

[144] M. Carena, S. Heinemeyer, O. Stål et al., “MSSM Higgs boson searches at the LHC: benchmark scenarios after the discovery of a Higgs-like particle”, *The European Physical Journal C* **73** (2013), no. 9, 1–20, `doi:10.1140/epjc/s10052-013-2552-1`. 128

[145] ALEPH, DELPHI, L3, OPAL Collaborations, and the LEP Working Group for Higgs Boson Searches Collaboration, “Search for the standard model Higgs boson at LEP”, *Phys. Lett. B* **565** (2003) 61, `doi:10.1016/S0370-2693(03)00614-2`, `arXiv:hep-ex/0306033`. 128

[146] T. Hahn, S. Heinemeyer, W. Hollik et al., “High-Precision Predictions for the Light CP-Even Higgs Boson Mass of the Minimal Supersymmetric Standard Model”, *Phys. Rev. Lett.* **112** (Apr, 2014) 141801, `doi:10.1103/PhysRevLett.112.141801`. 128

[147] S. Heinemeyer, W. Hollik, and G. Weiglein, “FeynHiggs: a program for the calculation of the masses of the neutral CP-even Higgs bosons in the {MSSM}”, *Computer Physics Communications* **124** (2000), no. 1, 76 – 89, `doi:http://dx.doi.org/10.1016/S0010-4655(99)00364-1`.

[148] S. Heinemeyer, W. Hollik, and G. Weiglein, “The masses of the neutral CP-even Higgs bosons in the MSSM: Accurate analysis at the two-loop level”, *The European Physical Journal C - Particles and Fields* **9** (1999), no. 2, 343–366, `doi:10.1007/s100529900006`.

[149] G. Degrassi, S. Heinemeyer, W. Hollik et al., “Towards high-precision predictions for the MSSM Higgs sector”, *The European Physical Journal C - Particles and Fields* **28** (2003), no. 1, 133–143, `doi:10.1140/epjc/s2003-01152-2`.

[150] M. Frank, T. Hahn, S. Heinemeyer et al., “The Higgs boson masses and mixings of the complex MSSM in the Feynman-diagrammatic approach”, *Journal of High Energy Physics* **2007** (2007), no. 02, 047. 128

[151] N. Craig, “The State of Supersymmetry after Run I of the LHC”, `arXiv:1309.0528`. 129

[152] The ATLAS Collaboration, The CMS Collaboration, The LHC Higgs Combination Group Collaboration, “Procedure for the LHC Higgs boson search combination in Summer 2011”, Technical Report CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, (Aug, 2011). 133, 134

[153] S. S. Wilks, “The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses”, *The Annals of Mathematical Statistics* **9** (03, 1938) 60–62, `doi:10.1214/aoms/1177732360`. 134, 135

[154] G. Cowan, K. Cranmer, E. Gross et al., “Asymptotic formulae for likelihood-based tests of new physics”, *The European Physical Journal C* **71** (2011), no. 2, `doi:10.1140/epjc/s10052-011-1554-0`. 135

[155] A. L. Read, “Linear interpolation of histograms”, *Nuclear Instruments and Methods in Physics Research Section A* (1999) `doi:http://dx.doi.org/10.1016/S0168-9002(98)01347-3`. 136

[156] L3 Collaboration, “Search for colour reconnection effects in $e^+e^- \to W^+W^- \to$ hadrons through particle-flow studies at LEP”, Technical Report Phys. Lett. B 561, 202, CERN, (2003). 143

[157] DELPHI Collaboration, “Investigation of Colour Reconnection in WW Events with the DELPHI detector at LEP-2”, Technical Report Eur. Phys. J. C 51, 249, CERN, (2007). 143

[158] CMS Collaboration, “Measurement of VZ production cross sections in VZ to Vbbbar decay channels in pp collisions at 8 TeV”, Technical Report CMS-PAS-SMP-13-011, CERN, Geneva, (2013). 145

[159] CMS Collaboration, “Search for the standard model Higgs boson produced in association with W or Z bosons, and decaying to bottom quarks for ICHEP 2012”, Technical Report CMS-PAS-HIG-12-019, CERN, Geneva, (2012). 146