Search for a Heavy Neutral Higgs Boson decaying into Tau Pairs with the CMS Experiment at the LHC

Dissertation

zur Erlangung des Doktorgrades
des Department Physik
der Universität Hamburg

vorgelegt von

Luigi Calligaris

aus Trieste (Italien)

Hamburg
2015

Gutachter/in der Dissertation:	Dr. Andreas Meyer Prof. Dr. Elisabetta Gallo
Mitglieder/in der Prüfungskommission:	Prof. Jan Louis Prof. Dr. Elisabetta Gallo Dr. Andreas Meyer Prof. Dr. Christian Sanders Dr. Roberval Walsh
Datum der Disputation:	29 Juli 2015
Vorsitzender des Prüfungsausschusses:	Prof. Jan Louis
Vorsitzender des Promotionsausschusses:	Prof. Jan Louis
Leiter des Fachbereichs Physik:	Prof. Dr. Peter Hauschildt
Dekan der Fakultät für Mathematik, Informatik und Naturwissenschaften:	Prof. Dr. Heinrich Graener

Abstract

A search for a heavy MSSM Higgs boson is performed in the $\phi \to \tau\tau \to e\mu$ decay channel, using data corresponding to an integrated luminosity of 19.7 fb^{-1} at a p-p center of mass energy of 8 TeV, collected by the CMS experiment at the Large Hadron Collider. The search is optimized for Higgs boson masses between 300 and 1000 GeV and makes the use of Boosted Decision Trees for an improved selection of signal events. The results are consistent with the predictions of the Standard Model.

Im $\phi \to \tau\tau \to e\mu$-Zerfallskanal wird eine Suche nach einem schweren MSSM Higgs-Boson durchgeführt. Die Daten hierzu entsprechen einer integrierten Luminosität von $19,7\ \text{fb}^{-1}$ und wurden vom CMS-Experiment bei einer Schwerpunktsenergie von 8 TeV am LHC aufgenommen. Die Suche ist für eine Masse des Higgs-Bosons zwischen 300 und 1000 GeV optimiert. Eine multivariate Analysetechnik wird verwendet, um die Auswahl der Signalereignisse zu verbessern. Die erzielten Ergebnisse stimmen mit Vorhersagen des Standardmodells überein.

Contents

Introduction

The development of theories describing nature stimulates experiments to test their assumptions and vice versa. Currently, matter and its interactions through forces are well described by the Standard Model of particle physics, a model that has been shown to be valid in the course of the last 40 years, and that is right now being subjected to stringent tests in a new generation of experiments. The Large Hadron Collider at CERN is today one of the largest and most demanding experimental facilities ever attempted in history, requiring the collaboration of more than ten thousand persons in the fields of physical theory, data analysis, technology development and operation, administration and services. This effort led to the discovery of the Higgs boson at a mass of 125.7 GeV/c^2, marking a milestone in elementary particle physics. The Higgs boson has a fundamental role in the Standard Model, as it is responsible for giving mass to particles. The electroweak bosons $W^{\pm}$ and Z^0 acquire mass through the mechanism of electroweak symmetry breaking mediated by the Higgs boson, and the fermions acquire their mass through Yukawa couplings with the same boson. A precise measurement of the properties of the the new discovered state will give us new insights about the inner mechanisms of the Standard Model and point to possible extensions, such as supersymmetry.

Indeed, some observed phenomena lack a satisfactory explanation in the context of the Standard Model. One of the most important issues is the large difference in scale of energy, and absence of new physics, between the observed scale of electroweak symmetry breaking ($m_{W^{\pm}} = 80.3$ GeV) and the Planck scale ($m_{Pl} \simeq 10^{19}$ GeV). The Planck scale is the scale of energy at which new physics is expected, because gravity becomes important at quantum level. Indeed, the calculation of the mass of the Higgs boson at higher orders requires the existence of new physics to cancel the quadratic divergencies arising from quantum loops in the Higgs boson self-energy. Trying to explain these cancellations as phenomena occurring at 10^{19} GeV and leading to a cancellation with a precision of 10^{-17} poses a problem of fine tuning. This problem is referred to as the *hierarchy problem* and supersymmetry provides an elegant solution to it. Supersymmetry includes the particles described by the Standard Model and for each of them predicts the existence of superpartners. In the Minimal Supersymmetric Standard Model (MSSM), which is the simplest supersymmetric extension of the Standard Model, fermions of the Standard Model have an associated boson superpartner, and bosons of the Standard Model have associated fermion superpartner. The association of a corresponding boson to each fermion provides a solution to the hierarchy problem, as the quadratic divergencies cancel out.

In parallel to measurements being performed at particle accelerators, astrophysical observations suggest the existence of a large quantity of matter that is yet unaccounted for by our theoretical models and that has not yet been directly detected, due to its weak interaction with ordinary matter. This matter is named *dark matter* and various theories provide an explanation for its existence. Supersymmetry is one of

them, predicting the existence of stable and massive neutral particles that could act as candidates for dark matter. The explanation for dark matter is another motivation for the search of evidence for supersymmetry.

This thesis describes a search for the heavy neutral Higgs bosons of the MSSM with data collected by the CMS experiment from proton-proton collisions at $\sqrt{s} = 8$ TeV during the 2012 run of the LHC. A search is performed for events in which the bosons decay into a pair of τ leptons, with one of the τ decaying to an electron and two neutrinos, the other τ decaying to a muon and two neutrinos. The thesis consists of five chapters:

In the first chapter, the Standard Model, the Higgs boson and electroweak symmetry breaking are introduced. The chapter then continues with the description of the MSSM and briefly discusses the Higgs bosons predicted by this supersymmetric model, their production at the LHC and their decay into $\tau\tau$.

In the second chapter the experimental setup is presented: a description of the Large Hadron Collider machine and the Compact Muon Solenoid experiment is provided.

In the third chapter the reconstruction of particles from data collected by the CMS detector and the simulated samples used to model the Higgs boson signals and Standard Model backgrounds are described. In the same chapter the analysis tools that will be used in the search are presented.

In the fourth chapter the selection of events, the corrections applied to simulated data, the subdivision of data and simulated samples into event categories, the observable used for the statistical analysis, $m_{\tau\tau}$ are discussed. Then, the preliminary results on the cross section and model-dependent limits on the existence of a MSSM Higgs boson are shown.

In the last chapter, optimizations to the event selection are discussed and the improvement on the analysis results is described. In the chapter an alternative observable, $m_{e\mu}$, to be used in the statistical analysis is presented and the fact that this observable gives consistent results, while reducing the dependence on systematic fluctuations, is shown.

Acknowledgements

During my three years and nine months spent at DESY I had the opportunity to work with many people, and the results exposed in this thesis disseration would not have been possible without their help, collaboration and support. I owe them my heartfelt gratitude and will name many of them below.

I want first to spend a moment in remembering Prof. Lorenzo Foa from Scuola Normale di Pisa, who encouraged me in taking the opportunity to apply for the PhD at DESY in Spring 2011. I will always keep him as an example for his experience and passion in experimental physics and for his successful contribution to the operation of an international science collaboration as Aleph and CMS.

I want to thank my supervisor, Dr. Andreas Meyer, for his patient effort in mentoring me during these years. A heartfelt greeting goes also to Dr. Alexei Raspiareza, who has always been available to help me in the development of this Higgs analysis, and is a great person to collaborate with. I cannot avoid to mention Dr. Daniel Pitzl, whose excellent supervision during the DESY test beam campaign and the development of the Single Event Upset detection software made those experiences fun and highly formative. I want to thank Dr. Olaf Behnke for the fruitful discussions about statistical analysis methods and thank Dr. Lucien Heurtier and Dr. Peter Drechsel for the interesting discussion about supersymmetry and physics beyond the Standard Model. I want to thank Dr. Aruna Nayak for the helpful discussions we had about Higgs searches in the tau lepton channels in CMS.

Many sincere thanks go to my PhD colleagues in the DESY CMS Higgs group, Dr. Agni Bethani, Dr. Jakob Salfeld-Nebgen, Dr. Armin Burgmeier, Dr. Gregor Mittag, Dr. Roberval Walsh, Dr. Rainer Mankel. It was a pleasure to work and have fun with you in Hamburg and at CERN and wish you all the best for your future.

I want to spend a thought in thanking my brother, Riccardo, my father Mario, my mother Daniela, my grandma Fiora and my grandpa Silvano for their support. A special moment of remembrance goes to my grandma Mirella, who could not celebrate with us the end of my PhD in Hamburg.

I want to thank all my friends and colleagues at DESY and in Hamburg, Ali Harb, Dr. Francesco Costanza, Dr. Paolo Gunnellini, Dr. Pedro Cipriano, Dr. Thomas Eichhorn, Dmitrij Kiessler, Engin Eren, Jasone Garay García, Aram Bugaev, Jan Kieseler, Carmen Diez Pardos, Monica Dobre, Matteo Centis Vignali, Artur Lobanov, Özgur Sahin, Martina Pola, Maria Katsiaflaka, Samaneh Javanbakht, Marina Conovca, Birgit Breetzke, Gabriele Kalhöfer, Dr. Matthias Kasemann, Prof. Somnath Choudhury, Tigran and Aram Kalaydzhyan, Simona Bianco, Dorothea Pozzato, Dr. Sasha Bylinkin, Dr.

Samantha Dooling, Dr. Christian Contreras, Dr. Doris Eckstein, Dr. Kerstin Borras, Dr. Johannes Hauk, Prof. Hannes Jung, Prof. Elisabetta Gallo, Carsten Muhl, Dr. Eleni Ntomari, Dr. Martin Vollmann, Alessandra Picchiotti, Alex Burilkov, Marco Argento, Dr. Ye Chen, Dr. Shiraz Habib, Dr. Pooja Saxena, Simon Spannagel, Dr. Cristoph Wissing, Gabriele Bellerino, Marco Filippuzzi, James Peters, Aliona Kadlubovich, Anita Chavarbasha, Iole Neviani, Victoria Sayenko, Ena Salcinovic, Masha Pereverzeva.

I want to thank my friends and colleagues at CERN and INFN, it was great to work with you and enjoy the time in Geneva during my visits: Dr. Fabrizio Palla, Dr. Francesco Fiori, Dr. Luca Martini, Dr. Gaetano Barone, Dr. Gino Bolla, Dr. Daniel Duggan.

This list is not by any means complete and should I have overlooked anyone, please accept my apologies.

1. Theoretical background

The Universe as we know it today can be described as composed of particles. Elementary units interact with each other following a small number of basic laws, and give rise to the complexity of macroscopic phenomena such as stars, planets and - ultimately - intelligent life. Particles are physical manifestations of quantum fields, described in the framework of quantum field theory.

Four forces are known in nature: the electromagnetic force, the weak force, the strong force and gravity. A theory to describe the behavior of the first three forces was developed between the early 1960 and the late 1970. This theory is commonly referred to as the **Standard Model**, it has been tested extensively and has so far withstood all tests exceptionally well.

1.1. The Standard Model of Particles

A detailed description of the Standard Model (SM) can be found in the literature [1, 2, 3, 4]. The relevant properties of this theory will be briefly described in the following paragraphs.

In the Standard Model, matter and its interactions are described by twelve fermions, their antifermions and six gauge bosons. The particles predicted by the Standard Model, represented graphically in figure 1.1, are the following:

- **Gauge bosons**, responsible for the electromagnetic (the photon, denoted as γ), weak (W^+, W^-, Z^0) and strong forces (the gluon, denoted as g).
- **Higgs boson**, responsible for the electroweak symmetry breaking mechanism, giving mass to the other particles in the model.
- **Quarks**: the up, charm and top quark (u, c, t) and the down, strange and bottom quark (d, s, b).
- **Charged leptons**: the electron, muon and tau lepton (e, μ, τ).
- **Neutral leptons**: the electron neutrino, muon neutrino and tau neutrino (ν_e, ν_μ, ν_τ).

Gravity is not included in the SM and its theoretical treatment in terms of quantum fields (quantum gravity) is still under study. Excluding gravity from the model has negligible effects on particles at energies at which we can currently perform observations, since in this regime the force is extremely weak compared to the others. Indeed, the gravitational interaction is not expected to have significant effects below the Planck scale, at around 10^{19} GeV. This energy is much larger than the highest energy reached

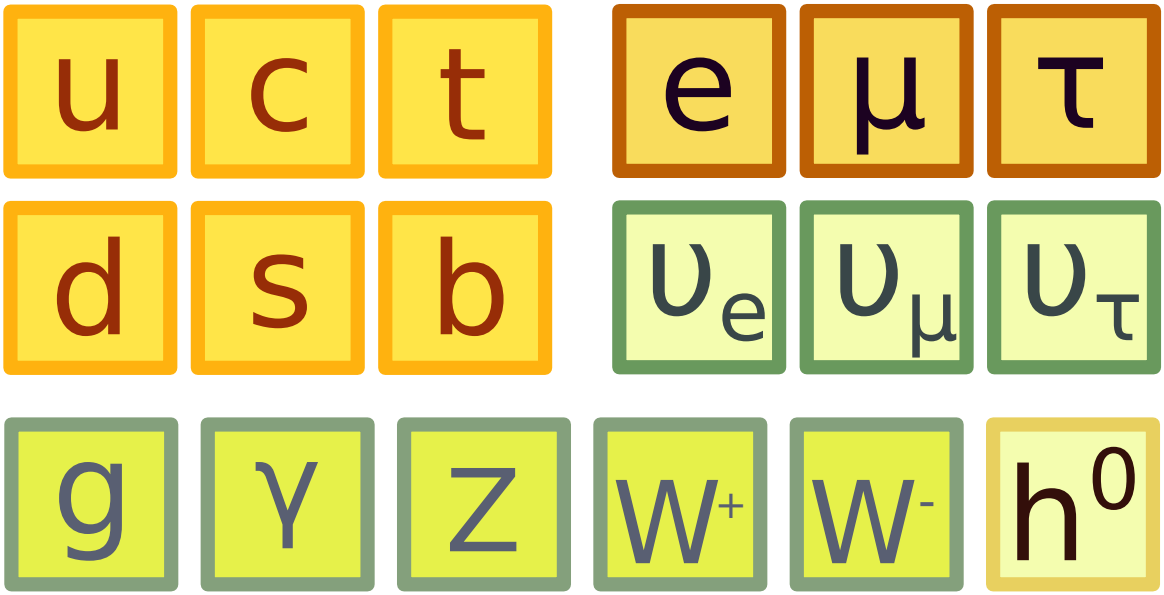


Figure 1.1.: The elementary particles of the Standard Model.

at particle colliders (the LHC, $1.3 \cdot 10^4$ GeV) and much larger than the highest energy measured, in the laboratory frame, for a cosmic ray ($3 \cdot 10^{11}$ GeV) [5].

In the Standard Model the quarks, charged leptons and neutrinos are fermions (that is, their spin is half-integer) making up the matter we observe in the universe. The forces through which these fermions interact are mediated by the elementary bosons (integer spin) of the model. In the SM the interactions are locally *gauge-invariant*, meaning that the phase of the quantum fields in each point of space-time can be rotated under certain symmetries, without affecting the outcome of the calculation on the observables.

Before introducing the electroweak symmetry breaking, the gauge symmetries of the fields in the SM are described by the group $U(1)_Y \times SU(2)_L \times SU(3)_C$. The first two terms are responsible for the unification of electromagnetic and weak forces into the electroweak force, and the third term is responsible for the strong force. Following Noether's theorem [6] there are conserved charges associated to each of these symmetries. These charges are, respectively, the weak hypercharge Y, the weak isospin L and the color charge C.

The group $U(1)$ has one generator, the group $SU(2)$ has three and $SU(3)$ has eight. We expect to have as many gauge boson states as generators for the corresponding symmetry group. The gauge boson corresponding to the $U(1)_Y$ symmetry is referred to as B, the ones for the $SU(2)_L$ symmetry are referred to as W^1, W^2, W^3, and the gauge bosons of $SU(3)$ are the eight gluons, indicated as g. It is important to note that, in this picture, all the bosons responsible for carrying the forces are massless states.

Experimental evidence was formed first in indirect experiments (observation of charged and neutral weak currents) and later in direct observation that the weak gauge bosons are not massless, and that they are instead very heavy compared to the other particles.

Discovered in 1983-1984 [7, 8, 9, 10], the $W^{\pm}$ and Z^0 bosons are today known to have a mass of 80.385 GeV and 91.1876 GeV [11], respectively. At the same time, the mass of the boson responsible for the electromagnetic force, the photon, has been proven to have a mass compatible with zero in all experimental tests.

Calculations show that it is not possible to account for the masses of the weak bosons with simple mass terms of the form $m_W^2 W_\mu W^\mu$ in the Lagrangian because they lead to non-renormalizable divergencies. The electroweak spontaneous symmetry breaking mechanism, instead, accounts at the same time for the very large masses of the weak bosons and the masslessness of the photon.

Spontaneous symmetry breaking as conceptualized in the SM adds a complex scalar field Φ. In the minimal construction proposed by its authors [12, 13, 14, 15, 16, 17], the field is a scalar $SU(2)$ doublet with weak hypercharge $Y = 1$. The doublet can be written as:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} (\phi_1 + i\phi_2) \\ \frac{1}{\sqrt{2}} (\phi_3 + i\phi_4) \end{pmatrix}, \tag{1.1}$$

where the complex fields are written as the sums of four independent real fields ϕ_i. In the SM, this field is supposed to have a self-potential $V(\Phi)$,

$$V(\Phi) = \lambda \left(\Phi^\dagger \Phi\right)^2 + \mu^2 \, \Phi^\dagger \Phi, \tag{1.2}$$

where the parameter is assumed to be $\lambda > 0$. Treating this expression as a second degree polynomial in $\Phi^\dagger\Phi$, and solving for the minimum $\frac{\partial V}{\partial(\Phi^\dagger\Phi)} = 0$, the solutions show that the shape of the potential depends on the sign of μ^2. If $\mu^2 \geq 0$, the solution is $\Phi = 0$, i.e. the field describes a massless boson doublet. In the case in which $\mu^2 < 0$, the potential reaches a minimum for values of the field that satisfy:

$$\Phi^\dagger \Phi = \frac{\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2}{2} = -\frac{\mu^2}{2\lambda}. \tag{1.3}$$

The field is said to acquire a *vacuum expectation value* (VEV) of $-\frac{\mu^2}{2\lambda}$. The choice on $\phi_1, \phi_2, \phi_3, \phi_4$ is arbitrary in the region defined by the condition above. Choosing $\phi_1 = \phi_2 = \phi_4 = 0$ we have $\phi_3^2 = -\frac{\mu^2}{\lambda} \equiv v^2$. Expanding the field in perturbation theory around the (displaced from zero) minimum, one finds three massless scalar bosons (Goldstone bosons), and a massive one (the Higgs boson, H^0). The three Goldstone bosons can be absorbed into the terms for the electroweak bosons W^1, W^2, W^3 and B to obtain four new gauge bosons states: three are the massive W^+, W^+ and Z^0 and the fourth is the photon γ, which is massless. The fact that the photon is massless is related to the fact that one symmetry remains unbroken. This symmetry is the $U(1)_{\text{em}}$ gauge symmetry, and its corresponding charge is the electric one. The $SU(3)$ part of the SM is not affected by the Higgs mechanism.

Many fermions of the Standard Model are experimentally confirmed to have a mass larger than zero, and in certain cases this mass can be significant (the top quark, for example, has a mass of 173.21 GeV/c^2 [11]). To generate this mass, Yukawa couplings are added to the theory between their fields and the just-introduced scalar field Φ.

The Standard Model has been subject, since the time of its inception, to a large variety of experimental tests. The discovery of the $W^{\pm}$ and Z^0 was followed by more and more stringent tests on their properties, such as cross section, mass and decay asymmetries, that confirmed the reliability of its predictions. The discovery of the top quark in 1995 at the Tevatron collider at Fermilab [18, 19], completed the third family of quarks and provided further confirmation of the soundness of the model. Studies of WW pair production at LEP and at the Tevatron were found to be in line with the predictions.

In July 2012, the ATLAS and CMS collaborations at the LHC announced the discovery of a boson [20, 21] at a mass around 125 GeV/c^2 compatible with the SM Higgs boson, the massive field of the electroweak symmetry breaking mechanism. This observation was confirmed independently in each of the decay channels that contributed to the initial observation ($H \to ZZ$ [22] and $H \to \gamma\gamma$ [23]) and it is supported by measurements in the other decay channels, such as WW [24], bb [25, 26] and $\tau\tau$ [27], and in a final combination of the results from the two collaborations [28].

Despite deviations from the expected outcome of a small number of experiments (e.g. the g-2 experiment at BNL [29]), the SM still stands as the widely accepted model to describe particles and their interactions.

From a theoretical point of view, though, the Standard Model fails in justifying why its parameters assume the values we observe in the experiments. More specifically, certain combinations of parameter values look rather improbable without the existence of a mechanism from which these values would emerge. Also, there are astrophysical arguments that justify attempts to extend the SM. They will be exposed in the next section.

1.2. Motivations for Supersymmetry

Supersymmetry (SUSY) is a family of theories that have a general assumption in common: every particle described by the Standard Model has one or more "superpartners". Supersymmetry is an elegant solution for a variety of reasons, mainly because superpartners stabilize the higher order corrections of known physical processes. Some of the reasons are discussed below:

Unification of the forces

The running of the coupling constants of the weak, electromagnetic and strong forces as the energy scale increases suggest the possibility that these forces unify at a very high energy, manifesting themselves as different forces at lower energies as a result of an undiscovered mechanism. The energy scale at which this unification takes place can be estimated to be of the order of 10^{16} GeV. Theories contemplating the unification of the forces are referred to as Grand Unification Theories (GUT), and predict the decay of the proton. The current lower limit for the mean life of the proton ($\tau_p > 2.9 \cdot 10^{29}$ years [11]) already excludes the unification in a scheme assuming only the particles of the Standard Model. In a GUT scenario, supersymmetry brings the expected proton lifetime to a much larger value, compatible with observations.

The Hierarchy problem
One of the most important aspects that the Standard Model fails to explain is the absence of new physics between the electroweak scale ($m_W = 80.385$ GeV/c^2) and the Planck scale ($M_{\mathrm{Pl}} = 10^{19}$ GeV/c^2), at which new physics is expected due to the rising importance of gravity at quantum level. The presence of a "desert" across 17 orders of magnitude in energy introduces problems related to the naturalness of the parameters of the Standard Model.
For example, higher order corrections to the Higgs mass in the SM have a quadratic term, mainly due to the top quark, in the ultraviolet cutoff scale Λ up to which the SM is considered valid [30]. If this cutoff scale is chosen to be the Planck scale, the sums of the higher order corrections at this scale are required to cancel to a precision of 10^{-33} [31, 32, 33] in order to be able to account for the relatively light mass of 125.7 GeV/c^2 of the discovered Higgs boson. This *fine-tuning* is considered unnatural, and points to a mechanism that stabilizes the corrections at much lower scale. Supersymmetry provides a candidate for this mechanism.

Candidates for Dark Matter
It is an accepted fact that a significant fraction of the total energy of the universe takes the form of a matter which is different from known stable baryonic matter, such as the prevailing hydrogen and helium that compose the galaxies [34, 35, 36, 37]. This non-baryonic matter is not visible directly in current observations, but can be inferred from the gravitational attraction that it exerts on visible mass. Since it is dark with respect to observation, this matter is denominated *dark matter*.
Many supersymmetric models provide a description for a massive, stable and non-interacting superpartner, that can be a candidate for the observed dark matter.

1.3. The Minimal Supersymmetric Standard Model

Among the various supersymmetric extensions of the Standard Model, the Minimal Supersymmetric Standard Model (MSSM) [**Fayet:1976etm**, 38, 39, 40, 41] is the simplest. It is a $N = 1$ supersymmetric model, meaning that there exists 1 supersymmetry under which particles can transform. The particles described by any supersymmetric theory are grouped in supermultiplets, that is, particles whose state is changed between one another under a supersymmetry transformation $\mathcal{Q}$:

$$\mathcal{Q}\,|Fermion\rangle = |Boson\rangle \quad \text{and} \quad \mathcal{Q}\,|Boson\rangle = |Fermion\rangle \tag{1.4}$$

The association between a particle and its superpartner of different spin is key to resolve the hierarchy problem [42], since among the higher order correction terms to the Higgs mass, the quadratic divergencies introduced by fermion loops are canceled by those from the superpartner boson loops.

In the scheme of MSSM, each fermion multiplet/singlet in the Standard Model has an associated SUSY boson multiplet/singlet, the union of which constitutes a *chiral supermultiplet*. Vice versa, each multiplet/singlet of gauge bosons of the SM has an associated SUSY multiplet/singlet of fermions, the union of which constitutes a *gauge*

supermultiplet. The chiral and gauge supermultiplets of the MSSM are shown in tables 1.1 and 1.2.

Particles	Supermultiplet	spin 1/2	spin 0
quarks and squarks	Q	(u_L, d_L)	$(\tilde{u}_L, \tilde{d}_L)$
(same for cs and tb families)	$\bar{u}$	$u_R^\dagger$	$\tilde{u}_R^*$
	$\bar{d}$	$d_R^\dagger$	$\tilde{d}_R^*$
leptons and sleptons	L	(ν, e_L)	$(\tilde{\nu}, \tilde{e}_L)$
(same for μ and τ families)	$\bar{e}$	$e_R^\dagger$	$\tilde{e}_R^*$
Higgs and higgsinos	H_u	(H_u^+, H_u^0)	$(\tilde{H}_u^+, \tilde{H}_u^0)$
	H_d	(H_d^0, H_d^-)	$(\tilde{H}_d^0, \tilde{H}_d^-)$

Table 1.1.: Chiral supermultiplets of the MSSM.

In general, supersymmetric particles are indicated with a tilde character (˜) above their symbol. The superpartners of SM fermions are spin-0 particles and are identified by prefixing an "s" (for scalar) to the name of their SM superpartner. The companion of the electron is the selectron $\tilde{e}$ (a slepton), the companion of the top quark is the stop $\tilde{t}$ (a squark), and so on. With a similar convention, the superpartners of the Standard Model bosons have an "-ino" appended to the name of their SM superpartners. The companions of the gauge bosons are called gauginos and are spin 1/2 particles. As an example, the model predicts the winos $\tilde{W}^1, \tilde{W}^2, \tilde{W}^3$ and the bino $\tilde{B}$ which are, respectively, the superpartners of the (massless) weak eigenstates of the Standard Model weak bosons W^1, W^2, W^3 and the superpartner of the B boson (which were described previously in section 1.1).

The neutral states of the Higgs doublet superpartners ($\tilde{H}_u^0$ and $\tilde{H}_d^0$) mix with the winos and the bino to give birth to four neutralino mass eigenstates $\tilde{\chi}_1^0, \tilde{\chi}_2^0, \tilde{\chi}_3^0, \tilde{\chi}_4^0$. The latter are Majorana fermions, so they are their own antiparticle. At the same time, the charged states of the Higgs doublet superpartners ($\tilde{H}_u^+$ and $\tilde{H}_d^-$) mix with the winos to give birth to two chargino mass eigenstates χ_1^+, χ_2^+ (with antiparticles χ_1^-, χ_2^-). The MSSM Higgs sector and spontaneous symmetry breaking will be discussed in the next section. The physical states of the MSSM are shown in figure 1.2.

In the MSSM a discrete symmetry, called R-parity, is assumed to be conserved in all interactions. By calling B the baryon number, L the lepton number and s the spin, R-parity is defined as follows:

$$P_R = (-1)^{3\cdot(B-L)+2s}. \tag{1.5}$$

R-parity assumes the value $+1$ for the particles of the SM, and -1 for the superpartners. Its conservation provides an explanation for the very large upper limit on the lifetime of the proton, since without R-parity conservation some supersymmetric particles could mediate its decay into other particles at a much faster rate than what is observed.

Particles	Supermultiplet	spin 1	spin 1/2
gluon and gluino	$G_{a\ (a=1,\ldots,8)}$	g_a	$\tilde{g}_a$
W boson and wino	$W_{i\ (i=1,2,3)}$	W_i	$\tilde{W}_i$
B boson and bino	B	B	$\tilde{B}$

Table 1.2.: Gauge supermultiplets of the MSSM.

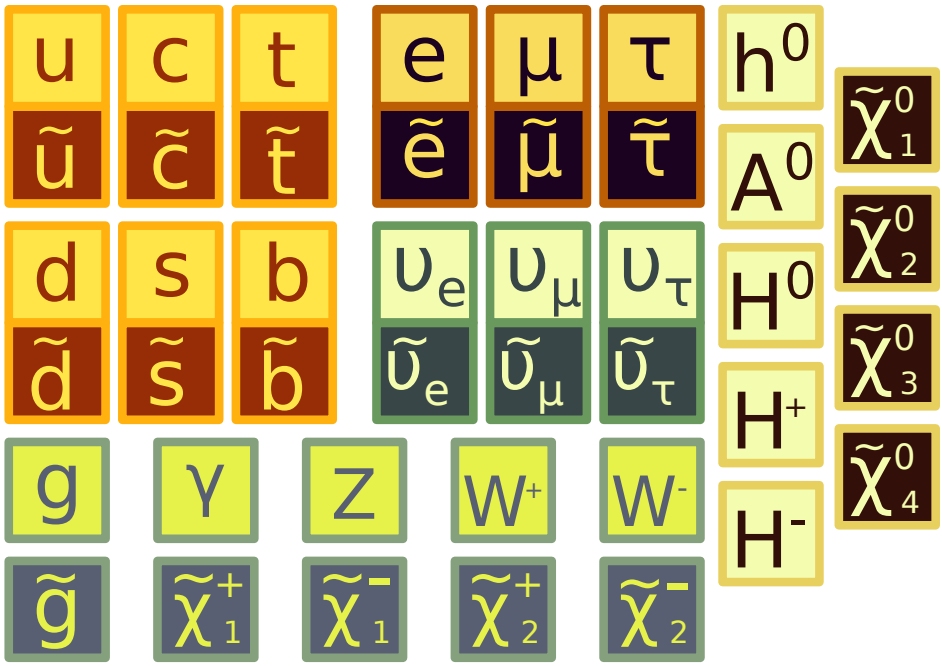


Figure 1.2.: The elementary particles corresponding to the physical states of the MSSM.

R-parity provides as well a possible explanation for the existence of dark matter, which in the scheme of MSSM is composed of stable, neutral, supersymmetric particles (for example, the neutralino is a possible candidate). These states would not be able to decay to ordinary matter particles due to R-parity conservation and, in parallel, would not be able to decay into other supersymmetric particles since they are assumed to be the lightest ones.

It is a matter of fact that SUSY particles have not yet been observed in experiments, and their presence has been excluded – for many SUSY models – in the mass region where we observe Standard Model particles. This fact suggests that supersymmetry may not be exact, so that the states in a supermultiplet are not bound to have the same mass. This *soft breaking* of the symmetry, that is, a breaking of the symmetry which does not change the relationships between particles and superpartners but changes the mass spectrum, can be modeled by the addition of soft breaking terms [43, 42] to the MSSM Lagrangian.

1.4. The MSSM Higgs

The MSSM assumes that there are two Higgs doublets [42], which are indicated with H_u and H_d. One is needed to give mass to the up-type quarks, and one is needed to give mass to the down-type quarks and the charged leptons [39, 41, 44, 45, 46, 45, 47]. Each of these doublets and their superpartners forms a chiral supermultiplet.

$$H_u = \begin{pmatrix} H_u^0 \\ H_u^+ \end{pmatrix} \qquad \text{and} \qquad H_d = \begin{pmatrix} H_d^- \\ H_d^0 \end{pmatrix} \tag{1.6}$$

The need to have two distinct Higgs doublets for giving mass to up- or down-type quarks and leptons stems from the presence of chiral anomalies in the theory. If only one Higgs doublet was used, renormalizability would be spoiled [48].

The Higgs mechanism in the MSSM [49, 50, 42] is similar to the one in the SM: due to the particular arrangement of the self-couplings in the superpotential, the neutral states in each Higgs doublet acquire a vacuum expectation value

$$\langle H_u \rangle = \frac{v_u}{\sqrt{2}} \qquad \text{and} \qquad \langle H_d \rangle = \frac{v_d}{\sqrt{2}}. \tag{1.7}$$

The degrees of freedom corresponding to massless bosons (Goldstone bosons) mix with the massless W bosons and B boson to give birth to the physical states: $W^\pm$ and Z^0 bosons, the photon γ, and the remaining degrees of freedom give birth to five physical Higgs boson states. Two states are neutral and CP-even: the lighter h^0, which is identified with the SM-like Higgs boson observed at 125 GeV/c^2, and the heavier H^0. Another Higgs boson, A^0, is neutral and CP-odd, while the last two are charged (H^+ and H^-).

The Higgs sector of the MSSM can be parametrized by six numbers:

- Four parameters are the masses of the Higgs bosons, m_{h^0}, m_{H^0}, m_{A^0}, $m_{H^\pm}$.

- One parameter is the mixing angle between the physical neutral CP-even Higgs bosons (h^0, H^0) and the neutral states of the Higgs doublet (H_u^0, H_d^0). This angle is by convention denominated as α.

- The last parameter is another angle and is related to the ratio of the VEVs of the two neutral states of the Higgs doublet. The definition for this angle, commonly called β is the following:

$$\tan\beta = \frac{v_u}{v_d}. \tag{1.8}$$

  The $\tan\beta$ parameter has an important influence on the coupling of the heavy neutral Higgs bosons to down-type quarks, as these couplings are proportional to $\tan\beta$ and the quark mass. In the specific case of the b quark, the value of $\tan\beta$ has a strong influence on the cross section for the production of a Higgs boson associated to b quarks.

The structure of the Higgs superpotential constrains the six parameters [42], so that the behavior of the Higgs bosons at tree level is defined by only two independent parameters. It is customary to choose as independent parameters the mass of the CP-odd neutral boson m_{A^0}, and the ratio of the Higgs doublet VEVs, $\tan\beta$, defined above in equation 1.8. If higher order corrections are taken into account, more parameters of the MSSM enter into the definition of the behavior and spectrum of the Higgs sector [49, 50]. The largest corrections arise from loops containing heavy fermions (top and to a lesser extent bottom quarks) and their superpartners (stop, sbottom). It would be impractical from the experimental point of view to test all possible combinations of parameters, therefore, there have been efforts in the theoretical community to prepare benchmark scenarios [51, 52] that constrain most of the parameters and provide a reference set of models that can be used in MSSM Higgs searches, covering adequately the parameter space of interest. In this search, the m_h^{mod+} benchmark scenario [52] was used. This scenario has the attractive property of presenting a light Higgs boson h^0 with a mass of 125.5 ± 3.0 GeV/c^2, which is consistent with the Standard Model Higgs boson observed at ATLAS and CMS. The possibility that the discovered Standard Model Higgs boson is one of the heavier neutral Higgs boson states, and that the light Higgs boson lie at low masses, has not been completely ruled out but, as of today, is strongly disfavored by experimental measurements.

If the mass of the pseudoscalar boson A^0 is significantly larger than the mass of the neutral weak boson Z^0, a condition known as "decoupling limit" [42] is reached. In the decoupling limit, the indirectly measurable effects due to MSSM Higgs bosons in the electroweak sector are driven by the light Higgs boson h^0, making predictions on the mass of the other Higgs bosons difficult. In such a situation, the mass of H^0 and A^0 are almost degenerate.

The search discussed in this thesis is optimized for a high mass Higgs boson ($300 \div 1000$ GeV/c^2) and assumes that the lightest Higgs boson h^0 coincides with the discovered SM Higgs boson. It also assumes that the decoupling limit is valid. Following this latter assumption, and the consequent degeneration in mass, the H^0 and A^0 bosons are not going to be treated separately and their combined contribution to the mass spectrum is used for the search, i.e. the Higgs templates used in the statistical analysis are the sum of the templates corresponding to the two Higgs bosons.

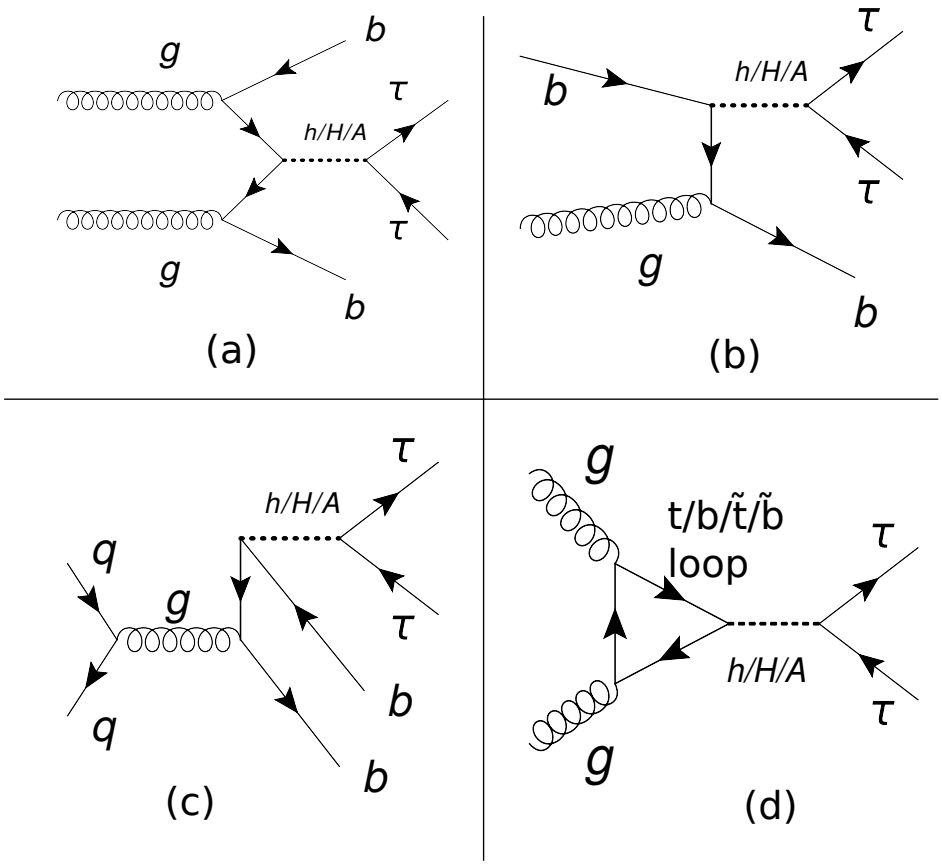


Figure 1.3.: Feynman diagrams for the production of MSSM Higgs bosons at the LHC.

1.4.1. Production of an MSSM Higgs Boson at the LHC

At the LHC, the MSSM neutral Higgs bosons are mainly produced by four processes, shown in figure 1.3. The first three processes involve the associated production of b-quarks, this makes the detection of these processes particularly sensitive to the value of $\tan\beta$. The b-associated production becomes dominant for large values of $\tan\beta$, i.e. when it is larger than 10. Figure 1.4 shows the production cross-section as a function of m_{A^0} for the A^0 boson (continuous lines) and the combination of h^0 and H^0 (dashed and dotted lines).

The last process (d in figure 1.3) is gluon fusion and it features a top/bottom quark loop. Although this is a higher order process compared to the previous three, the gluon luminosity of the LHC makes this process the most frequent one, with a cross section that is only weakly influenced by $\tan\beta$.

1.4.2. Decay into $\tau\tau \to e\mu$

Due to the relative largeness of the τ mass among the leptons, the $\tau\tau$ decay channel is the most important search channel to probe the coupling of neutral Higgs bosons to fermions. The Higgs bosons couple also to b-quarks and the branching ratio to $b\bar{b}$ is larger, as can be seen in figure 1.5, but the detection of b-quarks is comparably less efficient than in the case of τ.

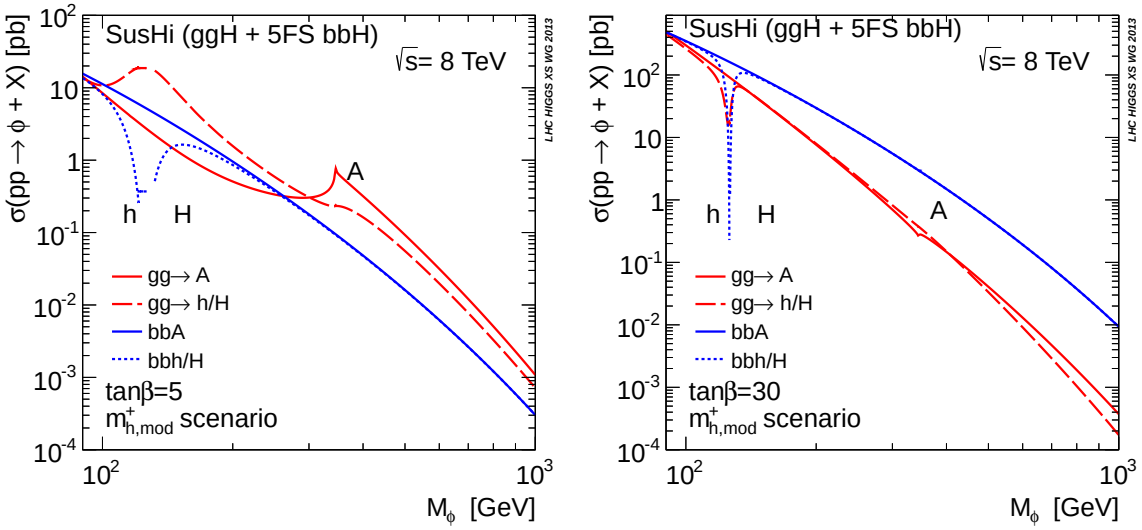


Figure 1.4.: Production cross-section for the MSSM Higgs bosons at $\sqrt{s} = 8$ TeV. From [53].

The branching ratio for the decay of a MSSM Higgs boson into pairs of particles is shown in figure 1.5. In the figure, the branching ratio is shown for the three bosons h^0, A^0 and H^0 for two values of $\tan\beta$ (10 and 50). In both cases, the $b\bar{b}$ is the dominant decay channel for all the bosons. As the mass of A^0 increases, the W^+W^- decay channel becomes important only for the h^0, while the A^0 and H^0 have a very small branching fraction to these particles. The decay into $\tau^+\tau^-$, instead, has a value always between around 10% and a few percent for all three bosons, making it suitable for performing searches in a wide mass range.

The $\tau\tau$ pair can decay to hadrons (41% of the time), to an electron (or a muon) and hadrons (23% of the time), to an $e\mu$ pair (6% of the time) or to a electron or muon pair (3% of the time). The smaller branching ratio in the case of the fully leptonic channels is in part counterbalanced by the high efficiency of reconstruction for leptons. In the case of the $e\mu$ channel, the analysis benefits from the absence of the very large $Z \to ee$ and $Z \to \mu\mu$ backgrounds present in the ee and $\mu\mu$ channels.

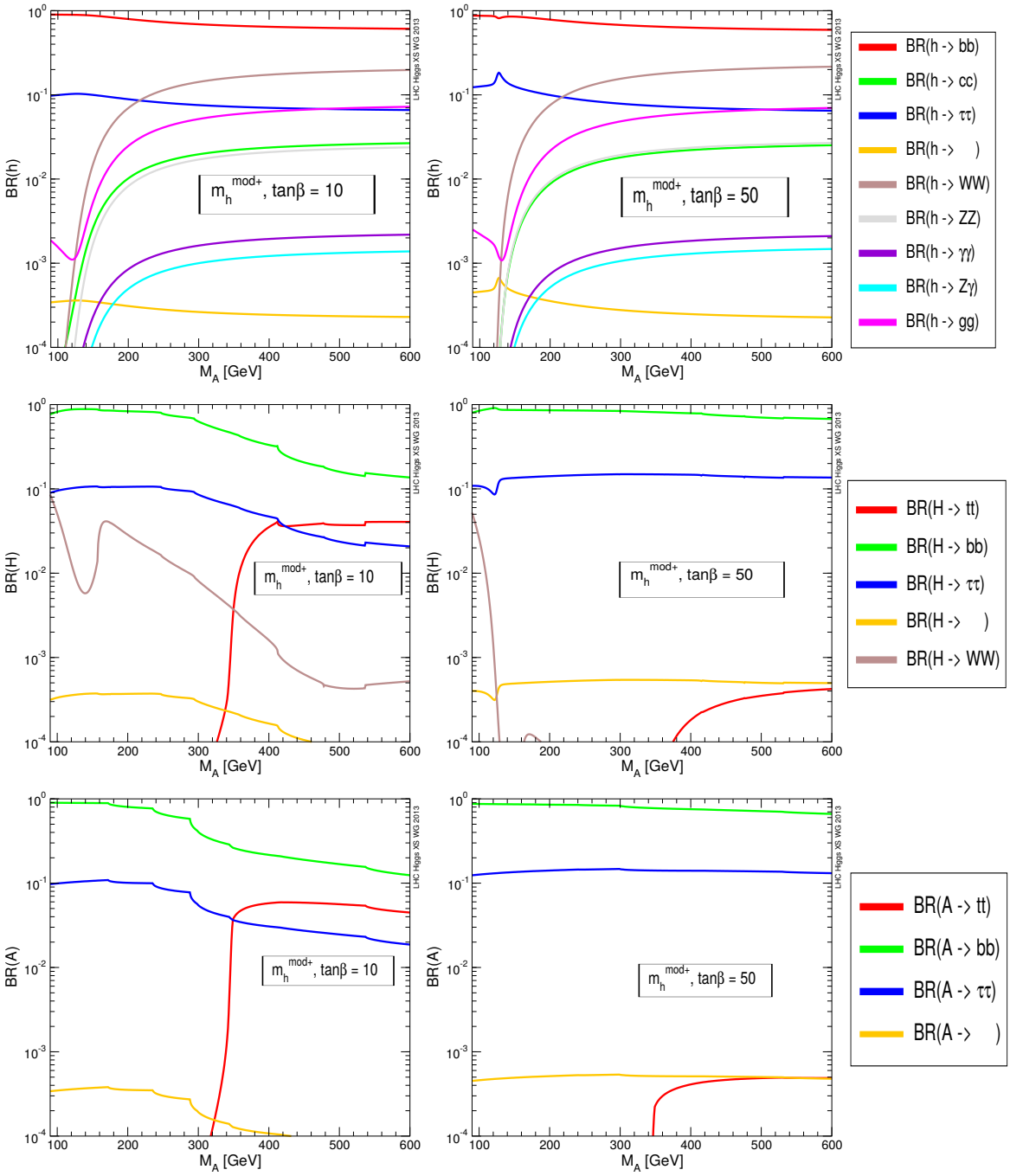


Figure 1.5.: Branching ratios of SUSY Higgs bosons in the m_h^{mod+} scenario as a function of m_{A^0}. Branching fractions for (top) h^0 , (middle) H^0, (bottom) A^0 bosons. Modified for visibility, plots from [52].

2. Description of the experiment

The European Center for Nuclear Research (CERN) has a long history of research in the field of particle physics, having started experiments in 1957 with a synchrocyclotron machine. CERN is located near Geneva, across the Swiss-French border, and the experiments that take place there have a significant impact on research, involving the collaboration of around 10000 scientists coming from every corner of the globe.

Among the operating machines, CERN hosts a proton linear pre-accelerator, a heavy ion linear pre-accelerator and corresponding accumulator ring, a small "booster" proton synchrotron, an antiproton accumulation ring, the Proton Synchrotron (PS), the Super Proton Synchrotron (SPS) and the Large Hadron Collider (LHC). The advantage for a site of this kind, sporting a great number of accelerators, is the possibility of re-using the existing ones as pre-injectors for future machines.

Following the discovery of the $W^{\pm}$ and Z^0 bosons at the $S\bar{p}pS$ in 1983-1984, CERN prepared for the start-up of the Large Electron-Positron Collider (LEP) in 1989. This collider reached an ultimate energy of 209 GeV in year 2000, after which the operations were terminated to permit the installation of the Large Hadron Collider. LEP made possible to perform high precision measurements on the parameters of the Standard Model, providing hints on the nature of the Higgs mechanism. The scalar particle of the Higgs field – the Higgs boson – was then discovered in the Summer 2012, and it is currently one of the most active fields of research for the experiments at the LHC.

2.1. The Large Hadron Collider

The Large Hadron Collider has been primarily designed to explore the range of energy between around 100 GeV and a few TeV, searching for the Higgs boson, supersymmetric particles or other particles predicted by theories beyond the Standard Model. Another important topic is the study of CP violation in the $b\bar{b}$ system. The machine also allows to perform collisions between ions. In these collisions it is possible to study a state of matter known as quark-gluon plasma.

Like other machines engineered to discover new particles, LHC collides hadrons circulating in opposite directions inside the rings of the accelerator. Currently it is possible to collide beams of protons, lead nuclei, or a combination of the two (p-Pb collisions). The option of colliding other kinds of nuclei (e.g. oxygen) is foreseen for the future. The maximum design energy in the center of mass is $\sqrt{s} = 14$ TeV for collisions between protons and 5.5 TeV/(nucleon pair) for lead ion collisions. The latter figure corresponds to a total energy of $\sqrt{s} = 1150$ TeV in the center of mass.

The search for the MSSM Higgs boson presented in this thesis is performed using a sample of data collected during the proton-proton collision mode. The following paragraphs will therefore focus on the aspects regarding this specific mode of operation of

Figure 2.1.: A view from inside one of the arcs of the LHC accelerator tunnel: the blue cylinders in the foreground are the superconducting dipoles used to steer the protons into a circular orbit. Image from [54].

the LHC machine.

The chain of machines that act as pre-accelerators for the protons injected in the LHC is shown in the scheme of the CERN accelerator complex, in figure 2.2. Protons are initially extracted from a plasma of hydrogen gas by an electrostatic field and accelerated by a small linear accelerator. They are then injected into the main linear accelerator, indicated as *Linac2*, where they reach a top energy of 50 MeV. A small circular accelerator, the *PS Booster* raises the energy of the protons to 1.4 GeV, injecting them into the *Proton Synchrotron* (PS). The top energy of the PS accelerator is 28 GeV. Protons are then injected into the *Super Proton Synchrotron* (SPS), that accelerates them to an energy of 450 GeV before injecting them into the LHC. It takes around 4 minutes and 20 seconds to fill each ring of the LHC with a nominal proton beam, this is accomplished with repeated injections by the pre-accelerators.

The LHC is a circular machine: particles travel along an orbit defined by the magnet system at a speed approaching the one of light and completing a full turn around the 26658.883 m circumference [56] of the accelerator 11245 times each second [56]. In nominal design conditions each beam is organized in 2808 "bunches", i.e. groups of protons containing around $1.2 \cdot 10^{11}$ particles each, and separated between each other by a time interval of 25 ns. The grouping of protons in bunches allows them to be accelerated by the machine radio-frequency cavity system with little losses.

The LHC re-uses the tunnel and part of the caverns that were excavated for LEP. This allowed significant savings in cost in the phase of construction. The 3.8 m-diameter tunnel rests on a slightly inclined plane, such that its depth from the surface ranges from 50 to 175 m. Vertical shafts allow the lowering of large pieces of equipment, running all the cables and pipes needed for the operations and the installation of

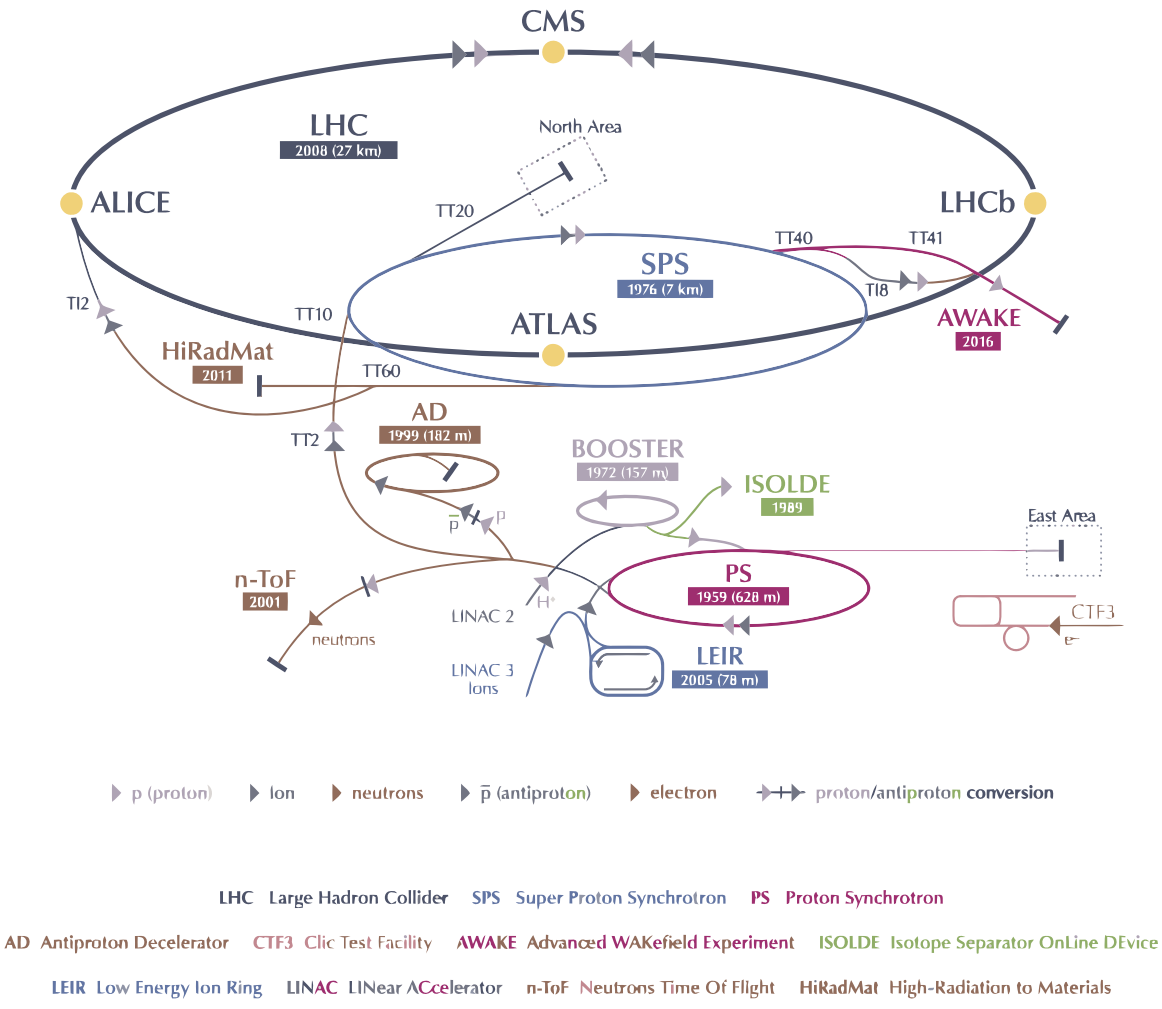


Figure 2.2.: The CERN accelerator complex. From [55].

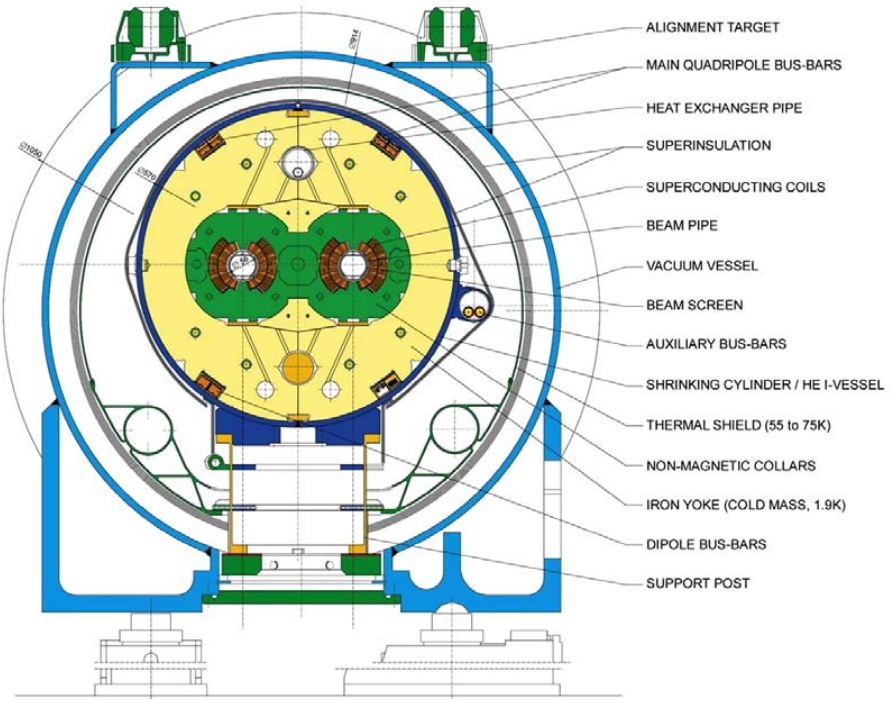


Figure 2.3.: Cross-section of an LHC dipole. From [57].

elevators for personnel access.

Protons are steered using static magnetic fields: dipole magnets are used to curve the trajectory of the beams into a circular orbit, quadrupole magnets are used to focus the protons into narrow beams that fit inside the acceptance of the machine and higher order multipole magnets correct for various aberrations of the beam optics and help keeping the beams stable. The dipoles, the quadrupoles and most of the corrector magnets of the LHC are based on niobium-titanium superconducting technology, which allows to achieve very intense fields with relatively modest power consumption. A large part of the power is used to cool the magnets at very low temperatures needed for superconductivity, of the order of 1.8 K. A cross section of the dipole is shown in figure 2.3. Almost all superconducting magnets in the machine employ the "two-in-one" design [57], where both beams circulate in different pipes of the same device, this configuration allows the accelerator to be cheaper, occupy less space and require less cooling power.

Once the LHC rings are filled by the SPS with circulating protons, the energy of the beams is “ramped” up to the maximum energy. This process needs around 25 minutes and the maximum speed at which the ramp can be accomplished is limited by the rate at which the current in the superconducting magnets can be increased.

The physics program of the LHC is focused on processes which are rare, compared to the well-known ones described by the Standard Model and studied in previous experiments. This can be expressed in terms of a small cross section σ, which has the units of an area, for the processes under investigation. The rate of events produced by the machine for a given process, R_{evt}, is proportional to the cross section for that

process and to an observable that measures the total number of collisions produced by the machine per unit of time. This observable is named *luminosity* and indicated with $\mathscr{L}$. The total number of events produced by the machine N_{evt} is the integral over time of the event rate:

$$N_{\text{evt}} = \int R_{\text{evt}}(t)\ dt = \int \sigma \cdot \mathscr{L}(t)\ dt = \sigma \cdot \int \mathscr{L}(t)\ dt = \sigma \cdot L_{int}\,. \tag{2.1}$$

The Large Hadron Collider is designed to maximize the *integrated luminosity*, L_{int}, by running the machine at high luminosity and do it for many hours, so that the fraction of time spent filling the accelerator and ramping its energy up, where no physics measurement is performed, is small. The nominal design luminosity of the LHC is $1.0 \cdot 10^{34}\ \text{cm}^{-2}s^{-1}$.

The LHC machine is divided into eight sectors, where the bending magnets are located, and eight long straight sections, in the center of which the experiments, accelerating cavities, beam dump or collimators are located. The beam is brought into collision in four *interaction points* (IP), around which the four big experiments were built:

- **ALICE** is an experiment located at the IP2 of the machine, and is dedicated to the study of ion collisions and the properties of the quark-gluon plasma [58].
- **ATLAS** is located at the IP1 close to the CERN Meyrin site, and is a general purpose detector able to study all sorts of processes. It features an innovative toroidal magnetic system for the detection of muons, which makes it the largest particle physics detector ever built at a collider [59].
- **CMS** is a general-purpose experiment at the LHC. It is located at IP5 near Cessy, France. It will be described in detail in section 2.2.
- **LHCb** is an experiment located at IP8. It is dedicated to the study of heavy flavor physics, the production of new baryons, the study of CP violation and the search for rare decays predicted in theories beyond the standard model [60].

Three smaller experiments are located close to the large ones, and pursue more specific research goals:

- **LHCf** is located near ATLAS, and measures the neutral collision products in the forward region (close to the direction of the initial protons) to improve the models describing the interaction of high energy cosmic rays with the atmosphere [61].
- **MoEDAL**, located near LHCb, searches for the production of magnetic monopoles and other highly ionizing stable massive particles escaping from the collisions at IP8 [62].
- **TOTEM**, located near the CMS experiment, observes the forward particles originating from the collisions at the center of CMS to measure the total scattering cross section for protons and to study diffractive and elastic processes [63].

In 2008 LHC suffered an accident involving tens of magnets, which stopped operations for one year. The reason for the accident was tracked back to an improper soldering between two cables, and an unsafe design of the interconnection between "dipole bus-bars" (shown in figure 2.3) of the superconducting magnets. As a consequence, during the 2009-2012 run of the LHC the machine ran at a lower energy with respect to the design one of 7 TeV per beam, as it was considered to be more safe to operate, pending improvements that have been finally installed in the Long Shutdown 1 (2013-2015).

In 2012 the energy for each proton in the beam was equal to 4 TeV, such that the total energy in the collision was $\sqrt{s} = 8$ TeV. During that data taking period, the proton bunches were separated by a time interval of 50 ns, with a consequent halving of the number of bunches in the machine, as in these conditions the operation of the machine is easier and more stable. Despite the limitations of running in this mode, LHC was able to reach a luminosity close to the nominal one, of the order of $0.7 \cdot 10^{34}$ cm$^{-2}s^{-1}$, exceeding expectations on its performance.

Luminosity is proportional to the number of bunches, so to reach a luminosity equal to 70% of the design one with only half of the bunches circulating in the accelerator, the number of proton-proton collisions in each bunch crossing needs to be higher. Indeed, this was the condition in the 2012 data taking period, when the number of *pile-up* interactions was significantly above design (around 50%). Pile-up interactions increase the number of particles traversing the tracking detectors, making it more difficult to assign each of them to their respective collision, and increase the background in the measurement of energy by the calorimeters, leading to a degradation of the resolution.

A considerable amount of effort from the experiments was dedicated to reduce the negative effects arising from the increased pile-up, and to bring the simulations in accordance with the real experimental conditions. Some of these techniques will be presented in chapters 3 and 4.

2.2. The Compact Muon Solenoid

The Compact Muon Solenoid experiment (CMS) studies the products of the collisions that take place in the interaction region. This region is determined by the LHC beam controls, that steer the two beams to collide at the center of the detector. To be able to fully reconstruct all the products of each collision, the detector is designed to almost fully envelope the interaction region, leaving only a small space for the passage of the LHC beam pipe. This hermetic configuration is called a 4π *detector*, that takes its name from the solid angle covered by a sphere.

To ease its construction and its maintenance, the cylindrical detector is divided in 11 slices that were lowered one at a time from the surface after the entire detector was built and tested on the surface. The 5 central slices follow a cylindrical geometry and compose the *barrel*, while 3 slices at each side follow a disk geometry and compose the *endcaps* of the detector.

The frame of reference of the detector is chosen such that the $\hat{x}$ axis points towards the center of LHC, the $\hat{y}$ axis points up towards the surface and the $\hat{z}$ axis points along the rotational axis of the detector in the direction of the counter-clockwise proton

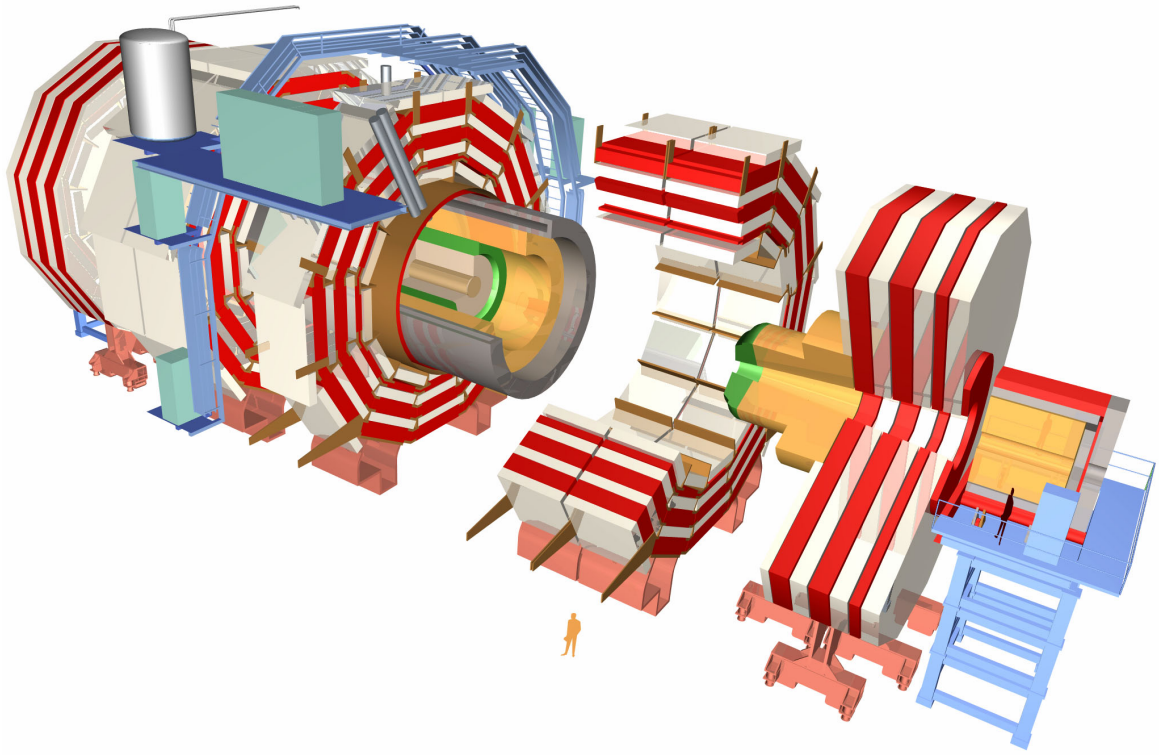


Figure 2.4.: A drawing of the CMS detector.

beam. The polar angle with respect to the positive $\hat{z}$ axis direction is called θ, the azimuthal angle in the xy plane is called ϕ. The variable called "pseudorapidity" and indicated with η is normally used to express the angle with respect to the direction of the beams, instead of θ. The pseudorapidity can be defined as a function of θ:

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right), \tag{2.2}$$

and is the massless limit for the relativistic rapidity y of a particle:

$$y = \ln\left(\tan\frac{E + P_z}{E - P_z}\right). \tag{2.3}$$

The structure of the detector is organized such that those detectors that perform a non-destructive measurement of the particles are positioned in the inner part, followed by detectors that perform a destructive measurement. The segmentation of the detector has also to accommodate the superconducting magnet, which allows for a precise measurement of the momentum of charged particles.

The different radiation conditions experienced by components in the central region and those in the forward regions, where the latter receive a significantly higher dose and flux of ionizing particles from the beam and the collisions, justify the choice of different technologies.

With reference to the exploded view in figure 2.4, the components of CMS can be ordered, from inside to the outside, in the following way:

1. Silicon pixel tracker, in darker brown.
2. Silicon strip tracker, in pale brown.
3. Electromagnetic calorimeter, in green.
4. Hadronic calorimeter, in yellow.
5. Superconducting solenoid, in dark grey.
6. Muon system, in white.
7. Magnetic flux return yoke, in red.
8. Forward calorimeter, right, in yellow.

The following sections will briefly describe each component.

2.3. Superconducting Solenoid

As the name Compact Muon Solenoid suggests, muons play a fundamental role in the physics program of the experiment. Indeed, many of the analyses require an efficient identification and high precision in the measurement of momentum of muons.

In order to achieve high precision in a wide range of muon transverse momenta, the lever arm over which the muons are bent should be large. This is achieved with a very wide magnet – it has an internal diameter of 6 m and a length of 13 m – and a very intense magnetic field in the magnetized volume. The field is generated by superconducting windings based on the same Nb-Ti technology used in the LHC magnets (Rutherford cable), extruded together with an aluminium buffer for quench protection and mechanical stability.

Figure 2.5.: The windings.

The windings are stacked radially in four layers (a section of the magnet prototype, published in [64], is shown in figure 2.5), in which a current of up to 19140 A generates a field of 4 T [64]. In normal operations the field is maintained at a slightly lower strength of 3.8 T, to lengthen the lifetime of the magnet. The CMS magnet is the most powerful single magnet ever built, storing in the magnetic field a total energy of 2.6 GJ. A specialized system takes care of removing this energy from the magnet in a controlled manner.

The magnetic flux passing through the solenoid is returned through a massive (10000 metric tons) iron yoke, in which the field saturates to a value of 1.8 T. The iron yoke reduces the effects of the intense magnetic field on machine components far from the magnet, helps increasing the field in the inner side of the solenoid and provides a volume of magnetized iron that acts as an absorber for particles different from muons and deflects the trajectory of muons, aiding the measurement of the transverse momentum of those muons that have a high p_T.

As the lever arm requirement makes the solenoid very large, a large amount of space is available inside the magnet to install various subdetectors: in CMS both

the electromagnetic calorimeter and the hadron calorimeters are installed inside the solenoid.

2.4. Electromagnetic Calorimeter

The task of the electromagnetic calorimeter (ECAL) is to measure destructively the energy of electrons and photons as they initiate an electromagnetic shower in the active material. The ECAL also identifies electrons and photons from the specific shape of their energy deposits. It is composed [64] of 61200 scintillating crystals made of lead tungstate ($PbWO_4$), a material with a very high density ($\rho = 8.28\ g/cm^3$), short radiation length ($X_0 = 0.89$ cm) and small Molière radius ($R_M = 2.2$ cm), so the ECAL can be compact and its spatial resolution can be high. $PbWO_4$ has a fast scintillation latency, so it is possible to resolve in time the light pulses and assign them to the correct bunch crossing, and also a high resistance to radiation.

The ECAL is composed of a barrel part, referred to as EB, and two endcaps, referred to as EE. The crystals in the EB (EE) have a front face measuring 22×22 mm^2 (28.62×28.62 mm^2) and they are tapered, so the back face measures 26×26 mm^2 (30×30 mm^2), these dimensions correspond to a width of 0.0174×0.0174 in the $r - \phi$ plane. Longitudinally, a crystal in the EB measures 230 mm, corresponding to 25.8 X_0 and EE crystals measure 220 mm, equal to 24.7 X_0. In both cases, the large radiation length ensures that electrons and photons start the shower inside the ECAL and release most or all of their energy in it. This way the resolution and identification of electrons and photons at high energy is improved.

The scintillation light is detected by avalanche photodiodes (APDs, shown attached to a crystal in figure 2.6) in the barrel, two per crystal, and vacuum phototriodes in the endcaps. The former type of detectors was chosen because it is more efficient, while the phototriodes were chosen as they are more radiation-resistant.

The ECAL is equipped with a preshower layer (ES) in the endcaps. This detector improves the identification performance of neutral pions, which decay to a closely spaced pair of photons. Despite the good granularity of the ECAL, the two photons are often difficult to separate by the use of this detector alone, so the ES is designed to provide a measurement of the impact points of the photons with a higher resolution. The ES also helps distinguishing electrons from minimum ionizing particles (MIPs), such as muons and hadrons. This subdetector is composed of two layers: in each layer a lead plate initiates a electromagnetic shower from the impact of photons (and electrons), then a silicon strip detector measures the extent of the shower along one coordinate, while the following layer measures the shower shape in the other coordinate.

The EB covers the pseudorapidity range $|\eta| < 1.479$ while the EE covers the range $1.479 < |\eta| < 3.0$. Finally, the ES covers the range $1.653 < |\eta| < 2.6$.

2.5. Hadronic Calorimeter

The role of the hadronic calorimeter is to measure destructively the energy of hadrons. Differently from electrons and photons, these particles tend to be more penetrating and have a longer mean free path, so they usually pass through the ECAL depositing

Figure 2.6.: A lead tungstate crystal, mated to the APDs. From [65].

little energy. This property is used to identify hadrons in the particle reconstruction algorithms.

The CMS hadron calorimeter (HCAL) is a *sampling calorimeter*, that is, a calorimeter where the absorber material and the scintillating material are interleaved: the absorber material induces a hadronic shower and the scintillating material samples the shower along its length.

The HCAL is composed of the following subdetectors: two endcaps (HE) that cover the pseudorapidity region $1.3 < |\eta| < 3.0$, one barrel subdetector inside the solenoid (HB) and one barrel subdetector outside the solenoid (HO), both covering the pseudorapidity range $|\eta| < 1.3$. In the region close to the beam, a forward calorimeter (HF) covers the pseudorapidity region $2.866 < |\eta| < 5.0$.

The HO is composed only of scintillators and uses the solenoid as absorber. It allows to measure the most energetic showers, those which propagate beyond the HB and through the superconducting coil.

Since the HCAL is inserted inside the solenoid, the absorber material needs to be non-ferromagnetic. If this was not the case, most of the magnetic flux generated by the solenoid would be "stolen" by the absorber, reducing the field in the tracker volume. Therefore, the absorber is mainly composed of C26000/cartridge brass plates [64], where the brass was obtained by decommissioned artillery shells. Plastic scintillators are inserted into slots between the brass plates. These slots are visible in figure 2.7, that shows the calorimeter endcap.

Wavelength shifting fibers are inserted inside the plastic scintillator panels to collect the generated light, the shift in color of the light makes transmission and detection of the light more efficient. The fibers are then spliced to clear fibers that carry the light to avalanche photodiodes for read-out. The HF does not use plastic scintillator panels, as they would not survive the radiation dose while maintaining a good performance. It instead uses quartz fibers as scintillating elements and – in a similar way to the other subdetectors – the light is transmitted to a read-out box, where it is measured by photo-multiplier tubes (PMTs).

2.6. Muon System

Muons can neither decay via the strong force, since they do not couple to gluons, nor cannot undergo annihilation with ordinary matter via the electromagnetic force, since ordinary matter does not contain muons. It follows that they can decay only into an electron and two neutrinos via the weak interaction, making their mean lifetime remarkably long, of the order of $\tau = 2.2$ µs. This time interval is further lengthened in the laboratory frame by relativistic time dilation. Muons lose energy only through

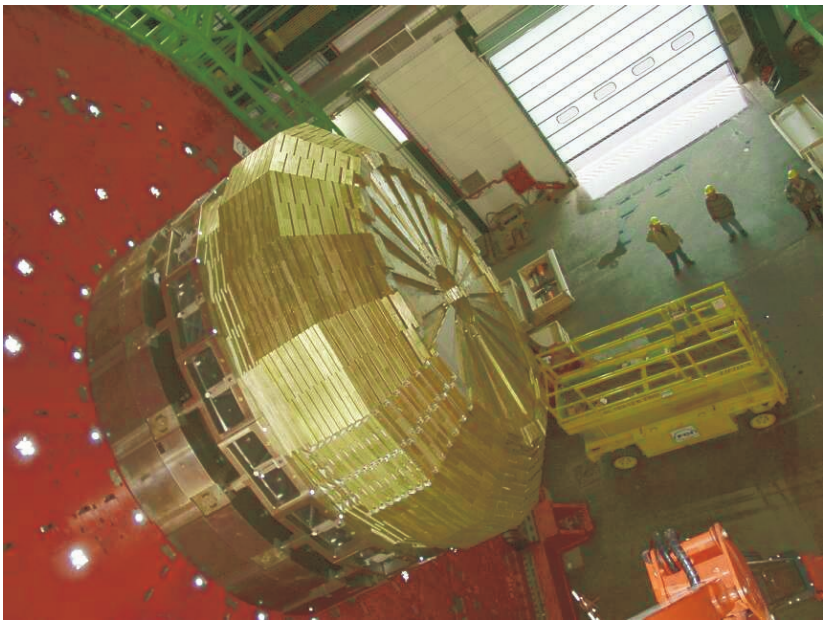


Figure 2.7.: One of the HCAL endcaps. The plastic scintillators are inserted in the darker slots carved into the brass absorbers. From [64].

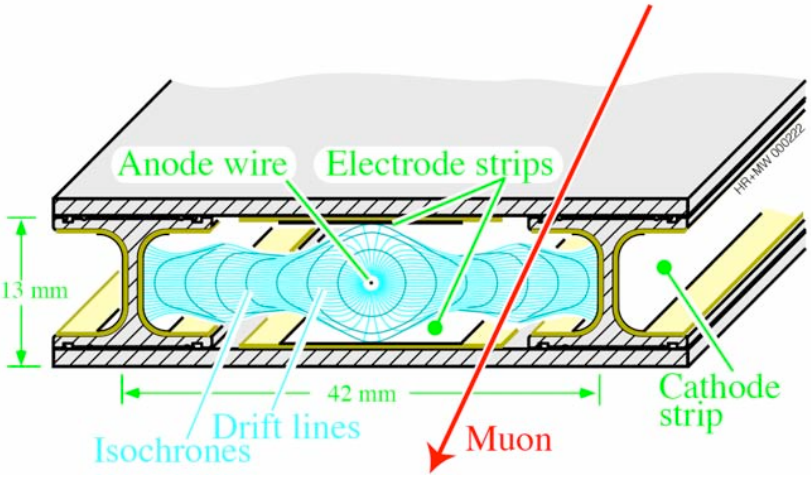


Figure 2.8.: Drawing of a drift tube cell. From [64].

ionization up to momenta of around 100 GeV/c when bremsstrahlung, e^+e^- pair production and photonuclear interactions become significant [11], and the energy loss in the ionization-dominated region is small. This fact, coupled with the long lifetime of muon, allows them to travel through the detector and ultimately escape from it. Particles of other types, instead, have a much greater probability of being stopped before reaching the external part of the detector. The identification of muons in CMS, therefore, can be accomplished by a simple system of detectors that measure charged particles as they travel through the iron yoke.

The *muon system* in CMS is composed of three types of detectors: drift tube (DT) chambers, cathode strip chambers (CSC) and resistive plate chambers (RPC) [64].

The drift tubes are installed in the barrel sections of the detector, and detect the muons by measuring the track of ionization along their path. Electrons released in the ionization process drift towards the anode wire under the effect of a strong electric field and initiate an avalanche close to the wire. The drift of the charges in the avalanche generates a current, which is amplified and read out by the detector electronics. The structure of a drift cell is shown in figure 2.8. The simple design, plates of aluminium spaced by insulated I-beams, with a filling gas mixture of 85% argon and 15% carbon dioxide, makes the detector easy and cheap to build. The drift tubes measure the position of the muon track from the time needed for electrons to drift to the wire, with a maximum drift time of 380 ns at 21 mm. This latency makes the chambers too slow to be used for event triggering at high luminosity [64]. A drift chamber determines the position of the centroid of a track passing through it with a resolution of 1.5 mm, and can measure the angle of the track with respect to the anode wire with a precision of 20 mrad.

The DT cells are arranged in "superlayers" composed of 4 stacked layers of cells. Three superlayers compose a DT chamber: two of them measure the $r\phi$ coordinate and are located at the upper and lower face of the detector, one superlayer is positioned in between and measures the z coordinate. Chambers are arranged in 4 rings (called "stations") around the beam axis: one ring sits inside the iron yoke, two rings are embedded in the iron yoke and the last one is external to it. The DT system covers the pseudorapidity region $|\eta| < 1.2$.

In the endcap region, the high stray magnetic field, its non-uniformity and the higher expected flux of charged particles discourage the use of drift detectors. Instead of drift tubes, cathode strip chambers are used to detect muons. The CSCs are multiwire proportional chambers, and use a gas under a strong electric field as active and amplification medium. They read the position of the traversing particles in two dimensions, this is accomplished through the measurement of signals on equally spaced sensing anode wires for one coordinate and on equally spaced cathode strips for the other coordinate. The measured position is an average of the wire/strip positions, weighted on the deposited charge. Each cathode strip chamber is composed of 7 panels forming the cathodes and 6 anode wire planes in between. Similarly to the DT, the CSC are fitted before, inside and past the iron yoke of the endcaps. The CSC system covers the pseudorapidity region $0.9 < |\eta| < 2.4$.

To ease assignment of muons detected by the DT and CSC systems to the correct

LHC bunch crossing, a fast detector is needed. This function is performed by the resistive plate chambers system. These detectors are composed of two parallel gaps filled with a mixture of gases (96.2% R134a ($C_2H_2F_4$), 3.5% iso-butane and 0.3% SF_6) and lined with resistive bakelite plates forming a high voltage capacitor. The passage of a ionizing particle causes a discharge between the plates, the $\vec{E}$ field disturbance is picked up by sensing wires. The resistive plates then recharge without supporting a damaging arc. The RPC system is positioned just behind DT and CSC chambers, up to pseudorapidities of $|\eta| < 1.6$.

2.7. Silicon Tracker

The CMS silicon tracker is the biggest semiconductor tracker ever deployed in a physics experiment, it totals around 200 m^2 of silicon detectors. In the inner part a pixel detector precisely measures the origin of charged tracks while the following layers, based on microstrip technology, measure the track's momentum, complement the measurement of the pixel detector and follow the track propagation up to the calorimeters and, in the case of muons, point to their impact position in the muon system detectors.

2.7.1. Pixel Tracker

Near the interaction point a very large flux of charged particles is present during normal operations. In order to avoid ambiguities in the reconstruction of particle trajectories the occupancy of the tracking detector, i.e. the fraction of channels affected by the passage of ionizing particles, needs to be kept low. As the pixel tracker is the innermost detector, close to the beam pipe, it needs to be able to achieve a high resolution both in the xy plane, to identify secondary vertices as in b-quark decays, and along the z axis, to separate particles originating from different proton-proton collisions in the same bunch crossing.

To fulfill this requirement, a pixel detector with a very fine granularity is used in the innermost part of the CMS tracker.

The pixel tracker is organized in three cylindrical layers in the barrel region and two fan shaped disks for each endcap, bringing its angular coverage up to a pseudorapidity of $|\eta| < 2.5$. The layout of the detector planes with respect to the beam axis is shown in figure 2.9. Both barrel and disk detectors use similar variations of the same hybrid detector technology.

In the barrel, the hybrid detectors are composed of two tiles, one containing the silicon sensors, the second containing the read out chips (ROC) for the sensors. A scheme of a read-out chip is shown in figure 2.11, showing the pixel read-out cells, data and time-stamp buffers and control electronics at the periphery. The hybrids are connected to the read-out back-end via a high density interconnect (HDI) board. This interconnect multiplexes the data measured by the various ROCs under its management to a single connection carrying the signals out of the detector. The three layers are joined together with glue and connected electrically with the use of bump bonds (sensor-ROC connections) and wire bonds (ROC-HDI connections), the sensor is sandwiched at the center. An exploded view of the hybrid detector of the barrel is visible in figure 2.10. A similar design is employed in the endcaps, where more sensor-ROC chip hybrids are arranged to obtain a petal-shaped blade of detectors.

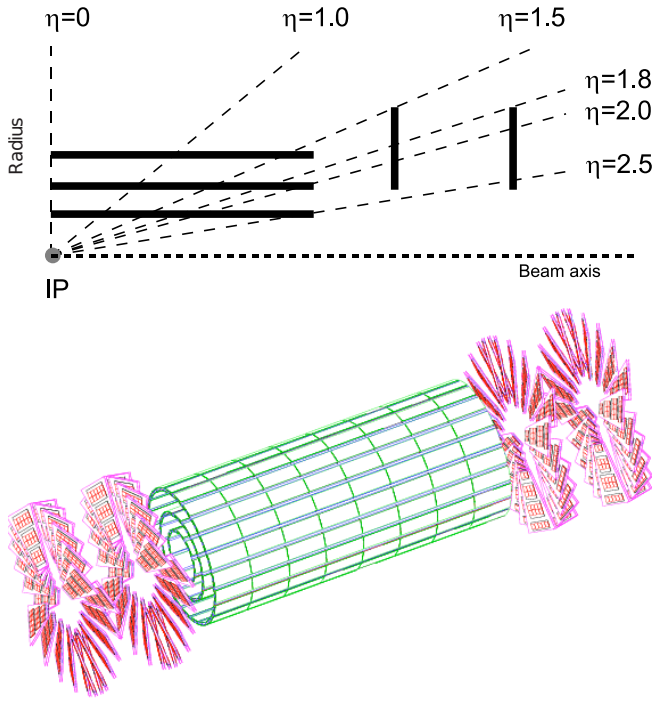


Figure 2.9.: Top: layout of the pixel detector planes in the $r-z$ plane. Bottom: a 3D view of its structure, showing the barrel layers (green) and endcap fans (pink). From [64].

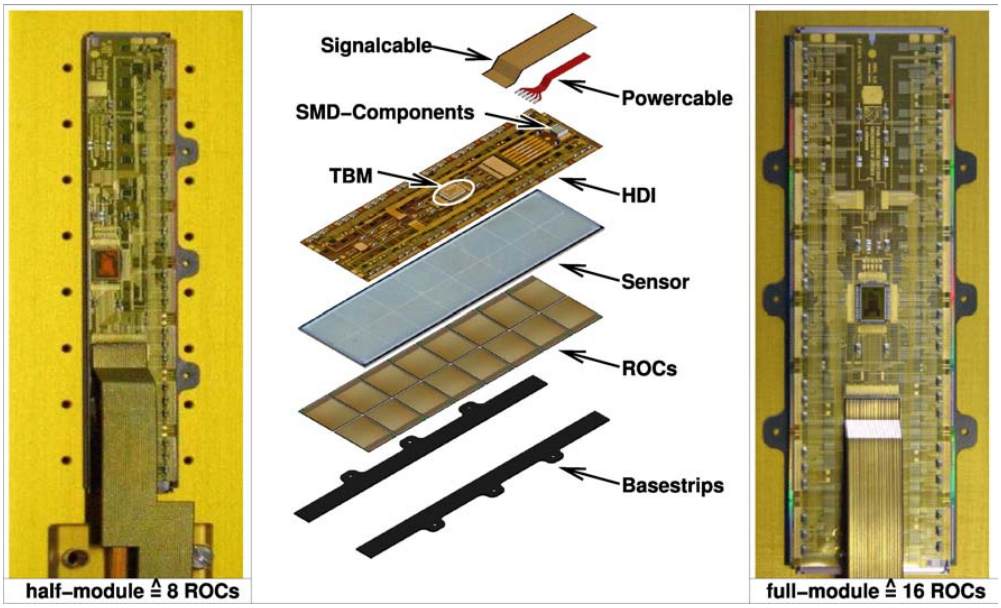


Figure 2.10.: Exploded view of a barrel pixel detector. From [64].

The silicon sensors are diodes that are reverse biased with a voltage that is increased over the life of the sensor, to counteract the effects of radiation damage. The passage of ionizing particles through the region depleted by the bias voltage frees electron-hole pairs that drift in the electric field, generating a current that is amplified by the ROCs. The sensing pixels have a size of 100 µm × 150 µm [64] and the estimation of the position of a hit is the average of the positions of the center of the pixels, where these positions are weighted on the charge collected by each pixel. This estimation can reach a precision of less than 10 µm [66].

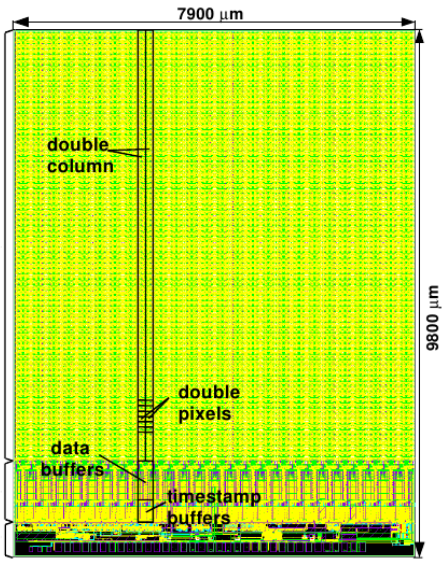


Figure 2.11.: The PSI46 read out chip. From [64].

The ROCs amplify the sensor signal received through the bump bonds. Their logic manages the amplifier gain and the threshold for the chip's own self-triggering. Given the large number of channels in the pixel detector (96 million channels), it would be impractical to read out all of them in each bunch crossing. Therefore, only pixels that show a signal above a programmable threshold are read out and their values are stored in an on-chip buffer. This buffer is needed, as the latency of the level 1 trigger of CMS (up to 3.2 µs [64]) is much longer than the bunch crossing rate of 25 ns. If a trigger signal is received about a specific bunch crossing, the ROCs transmit the relevant pixel hits to the front-end electronics outside the CMS detector for further analysis.

The logic cells in the ROCs are subject to large amounts of ionizing radiation coming from the proton beam, the collisions and the activation of the detector material. The ionization can inject charge into these circuits, changing their logic state. This kind of event is called a "single event upset", or SEU. While the ROC electronics have been designed to be resistant to these forms of noise, a small fraction of the chips in the detector (less than 1%) will randomly experience a SEU during a typical data taking run (a run is a period of time in which data taking takes place with a consistent configuration of the detector, it usually lasts between 3 and 10 hours).

Sometimes, a SEU can have a small effect on performance, for example changing the least significant bit in the triggering threshold and slightly changing the triggering rate. Other times the SEU can render a read-out chip temporarily blind to the passage of particles. Such an event can be seen in figure 2.12: in the histogram, the rate at which clusters of pixels above threshold are detected is shown (*cluster rate*). Time is expressed in LumiSections, a unit of time in use in CMS, which is equivalent to 23 seconds. In the case of a read out chip with normal behavior, shown in the top histogram of figure 2.12, the rate follows the exponentially decaying luminosity curve of the LHC machine. In the case of a chip experiencing a SEU, shown in the bottom histogram in the same figure, the rate drops abruptly for a long period of time. This is the signature of a single event upset with large effects on performance.

A software was developed to identify SEUs occurring in the pixel detector, and to provide diagnostics on their origin, evolution and eventual resolution. In normal

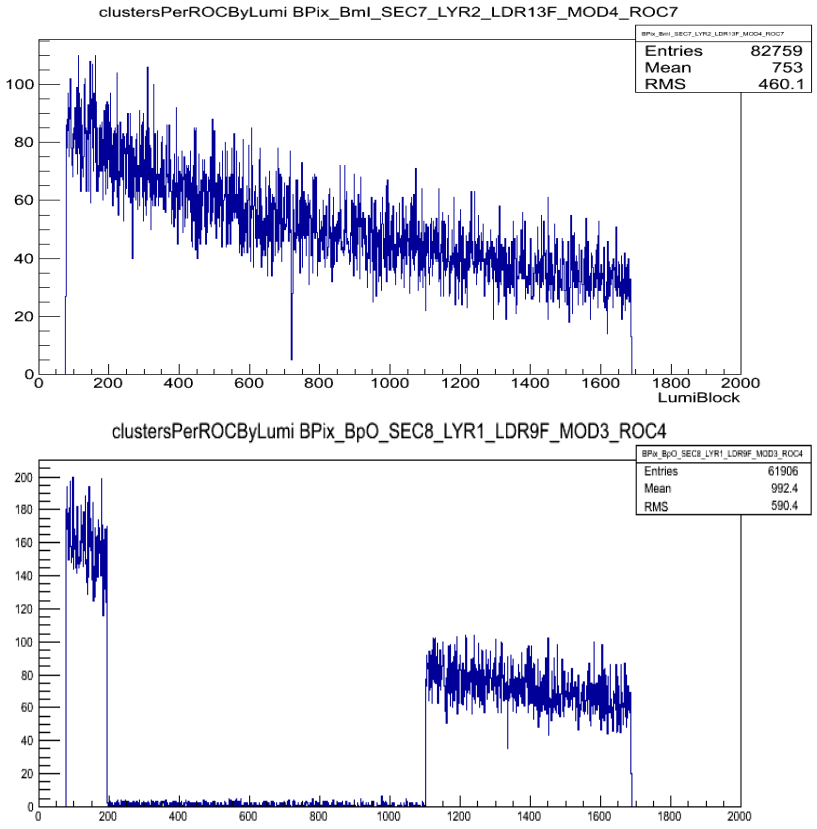


Figure 2.12.: ROC cluster rate during a 10-hour run. Top: in normal conditions. Bottom: for a chip experiencing a SEU that dramatically lowers its performance.

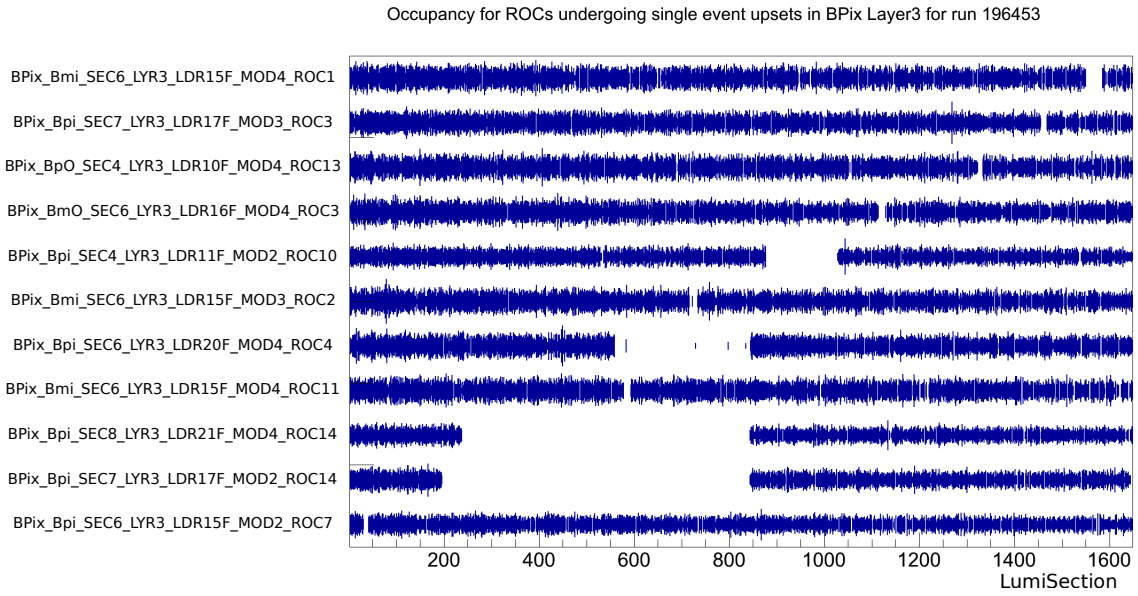


Figure 2.13.: ROC cluster rate expressed as the thickness of the blue line during the same run. Chips identified as experiencing SEUs are listed on the vertical axis.

operations, SEUs occurring in the pixel detector can be recovered by forcing a reset of the ROC electronics. With the aid of this identification software SEUs have also been observed to recover spontaneously in some instances. One of the outputs of this software is shown in figure 2.13.

In the figure, the evolution of the cluster rate is represented as the thickness of the blue line, for each ROC that has been identified as experiencing a SEU during a run of CMS. This particular image refers to the 3rd layer of the barrel pixel detector in run 196453. The graphs corresponding to each detector are ordered vertically, so each line corresponds to one ROC, and the white gaps in the time evolution of the rate are the periods of time in which the ROCs were inefficient due to a SEU. Having a view of the SEUs as they evolve in parallel is important, because it permits to identify collective effects that indicate problems in other components of the pixel detector. The SEU detection software is being used for further studies, which will help optimize the upgrades of the pixel detector in the future.

2.7.2. Strip Tracker

At farther distances from the beam line the flux of ionizing particles is lower, so microstrip detectors can be used while maintaining a low occupancy. This kind of detectors are made up of hundreds of small strips of p-type silicon, implanted on a bulk material composed of n-type silicon. A voltage depletes the sensor of free charges, so the entire sensor becomes sensitive to the passage of ionizing particles. The detectors measure the position of a hit along the coordinate transverse to the length of the sensitive strips.

In some layers of the strip tracker, the detectors are installed back-to-back to create a double layer. The strips are purposely misaligned by an angle of 100 mrad and thanks to the crossing of overlapping strips it is possible to measure the z coordinate. These layers are referred to as *stereo* layers.

The CMS strip tracker is composed of ten cylindrical layers in the barrel region (four inner ones and six outer ones), and twelve disks for each endcap (three inner and nine outer ones) [64]. The layout of the detector is shown in figure 2.14. An example of a "ladder" on which detectors have been mounted is visible in figure 2.15.

Different spacing between the sensing strips (*strip pitch*) are employed, ranging from 80 μm in the inner layers to 205 μm in the outer layers, in a compromise between lowering the number of channels to be read, maintaining a good precision in the position measurement and keeping the occupancy reasonably low, especially in the case of layers closer to the interaction point. The silicon strip tracker covers the pseudorapidity region $|\eta| < 2.5$.

The read-out of the strip sensors is performed on their periphery by custom-made ASIC chips, the APV25 [68], which is built to withstand high doses of radiation. In a similar fashion as the pixel detector case, the APV25 chips store in a buffer pulses that were measured to be above a set threshold, and transmit them outside the detector for further analysis upon receiving a level 1 trigger signal.

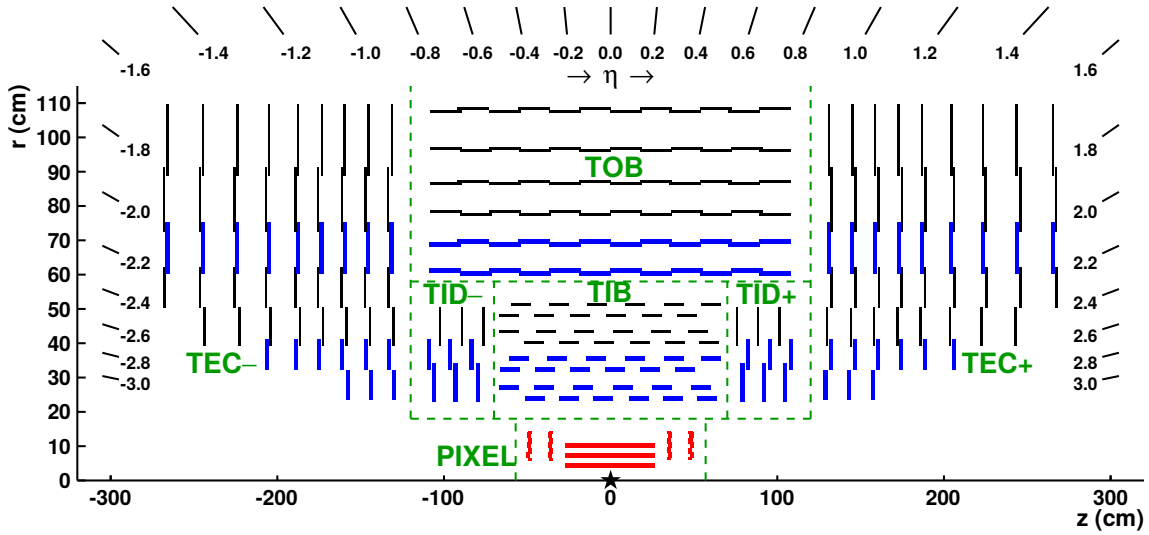


Figure 2.14.: Layout of the upper half of the silicon tracker. Single layers are shown in black and stereo layers are shown in blue. The pixel detector is in red. From [67].

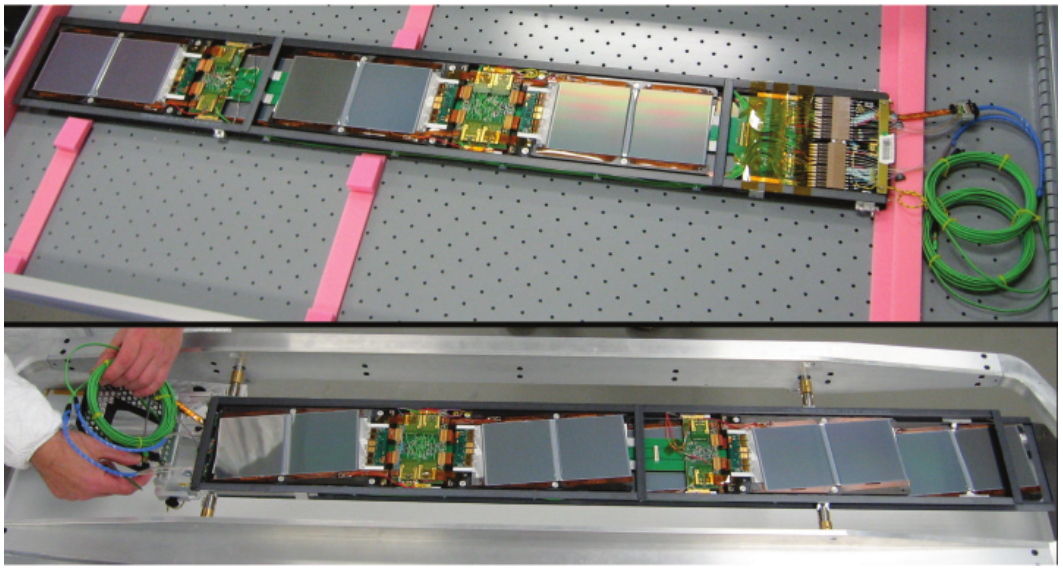


Figure 2.15.: A "rod" of silicon sensors, used in a stereo layer in the outer barrel part of the strip tracker (TOB). Optical fiber connection cables can be seen protruding from the sides of the rod. In the bottom figure, the stereo angle of the sensors is clearly visible. From [64].

2.8. Trigger System

The size in terms of data of an event in CMS is of the order of 1 MiB. Multiplying this value by the bunch crossing rate in nominal conditions (40 MHz) leads to a rate of 40 GiB/s. It is not practical, nor economically sound with current technology to attempt saving and processing this amount of data while the experiment is running.

The task of the trigger system is to select and store on disk a small number of interesting collisions out of a huge number of trivial events, arising from phenomena that were extensively studied in the past and now play the role of background in physics analyses. This tight selection is fundamental to allow the rate of events being stored and processed to reach manageable levels. The trigger system cannot afford saving events in a buffer, pending a later decision, and has to take all decisions in the time limit assigned to each of its stages. This mean also that reliability and monitorability is very important for its successful operation.

The trigger system in CMS is organized in two tiers that work serially: the level-1 trigger performs an initial selection, then the high level trigger runs more complex algorithms capable of finer selections. The two trigger stages will be described briefly below.

The level-1 trigger (L1T) is a decision logic implemented in hardware to be extremely fast. Field-programmable gate arrays (FPGA) and programmable memory look-up tables (LUT) are used where possible, to allow flexible of update of the logic. In certain cases, dedicated application-specific integrated circuits (ASIC) are used for the task. The triggering of the L1T is in connection with a specific proton bunch crossing, and provokes the read-out of all the relevant information corresponding to the triggered bunch crossing from the detector, such as the hits in the pixel and tracker detector.

Due to limitations in the latency, processing power and bandwidth of the connections between the various detectors and the trigger logic, only simple algorithms can be used. The L1T accepts the following two types of inputs as decision for issuing a trigger:

- Energy deposited in the ECAL, HCAL and HF calorimeters and variables that can be rapidly calculated from this primary information, such as number of jets, total hadronic transverse momentum, number of hadronic τ lepton decays, number of electron/photon candidates and calorimetric missing transverse energy.

- Information originating from the muon system about the presence of muons. Events containing a muon are automatically triggered. Muon track reconstruction is attempted using the hits detected by the muon system. In this respect, the precise timing function provided by the RPC is fundamental in indicating to the trigger which bunch crossings are the ones that produced muons.

The level-1 trigger lowers the rate of events from the full rate of the collisions of 40 MHz to a variable rate, which can reach a maximum of 100 kHz. This rate is low enough such that the stream of events can be piped to the high level trigger.

The high level trigger (HLT) is a software framework that shares many algorithms with the one used for offline reconstruction. It runs on a dedicated server farm, built on a cluster of commercial-off-the-shelf computers and installed in the vicinity of the

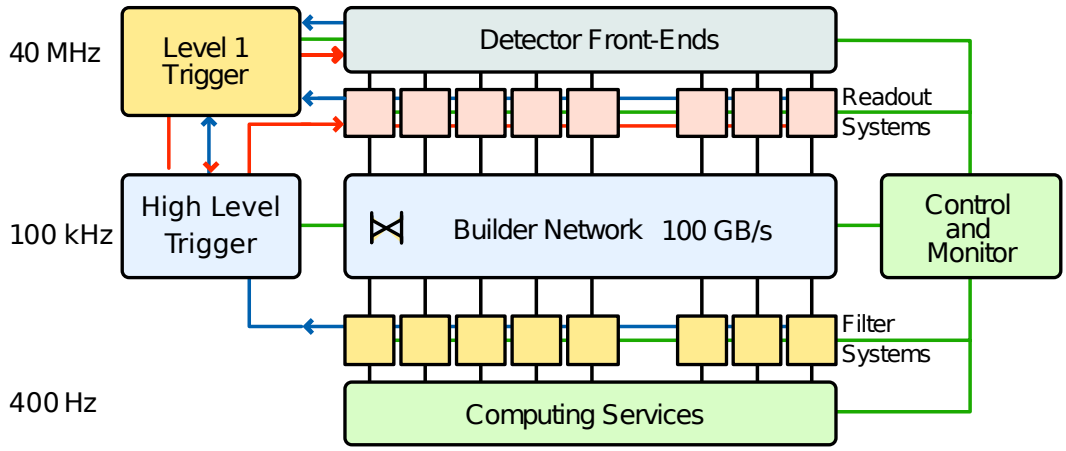


Figure 2.16.: The trigger (left), data acquisition (center), and detector control systems (right) work in parallel to select and collect events and guarantee the proper operation of CMS. Modified, originally from [64].

CMS detector. The fact that the HLT is implemented in software allows great flexibility in the choice of algorithms, some of which are even capable of performing a fast reconstruction of the tracks in the event from the hits provided by the silicon tracker, and trigger events based on the topology of the tracks, muons identified in the muon chambers and deposits of energy in the calorimeters. The HLT brings the selected event rate from the input 100 kHz of the L1T to around 400 Hz. This rate can be upgraded with the addition of more powerful computers to the server farm. The limitation of 400 Hz is mostly due to availability of storage space and processing power capacity downstream of the trigger.

The various trigger stages interact with and trigger the action of the data acquisition (DAQ) system, which takes care of reading out the data from the subdetector buffers, assembling pieces of data corresponding to a specific bunch crossing together to create the "event", the ensemble of data that describes entirely the measurements about the interesting collision. The entire process is monitored and controlled by the detector control system (DCS). This system of components is shown in figure 2.16.

3. Event reconstruction and analysis methods

The search for the MSSM Higgs boson in the $\phi \to \tau\tau \to e\mu$ decay channel presented in this thesis relies on a large set of techniques and tools which have been developed over the years in the CMS collaboration, and more generally in the high energy physics community. In this chapter the analysis methods and tools that are key to a successful the analysis are presented.

The process of enriching the sample with events of interest (which we refer to as "signal") while filtering the non interesting ones (the "background") is part of a family of problems referred to as "classification problems". In the case of this analysis, the signal from a possible Supersymmetric Higgs boson is expected to be small compared to the Standard Model backgrounds. In such a situation the efficient separation of signal and background becomes a priority if good sensitivity is desired.

Attaining an efficient separation involves choosing suitable variables characterizing the event – one example of such variables could be the D_ζ discriminant that will be introduced in section 4.3.2 – and then restricting the set of events accepted into the analysis by requiring these discriminating variables to lie in a properly chosen range in which an enrichment in terms of signal events is expected.

There are many possible variables that could be used to enrich the fraction of signal in the sample, some of which are statistically correlated. A simple method is to impose cuts on the value of each of these variables. This solution is robust and simple to understand, but does not make use of the information contained in the correlations between variables.

It is possible to account properly for the correlations between variables, even if they are nonlinear, with the use of more complex classifiers, the most common ones being artificial neural networks (ANN), support vector machines (SVM) and boosted decision trees (BDT). Due to the good compromise between performance and stability of their output, the latter have been chosen both for the task of identifying real electrons (discussed in section 3.5.1) and for the enrichment of signal events in the signal categories (described in 5.2). The next section explains how they work, why they offer an improved performance and how they are set up.

3.1. Boosted Decision Trees

Boosting [69] is a procedure developed in the field of machine learning in the early 1990s, following a paper from Robert E. Schapire [70]. It allows creating *regressors* and *classifiers*. Regressors are functions that can guess with good accuracy the value

of a variable from a collection of correlated variables. Classifiers are the special case of a regressor in which the target output value is interpreted with respect to a limited number of values, which act as labels representing the possible categories in which we can classify the objects.

The creation of a new classifier is based on a process called *training*, where the training algorithm – in our case the boosting process – operates on two or more datasets, one or more per category, representing the data we wish the classifier to separate in the course of its future application. The algorithm adjusts the classifier response by changing the parameters that define its behavior. In our specific case we are interested in the use of boosting to create classifiers to distinguish objects between two categories: signal and background. Commonly, for classification problems, events that look signal-like are assigned the label value $+1$, and events that look background-like are assigned the label value -1.

The boosting process operates on *weak learners* [70], simple decision algorithms characterized by the fact that their individual separation power is small. Separation power is a measurement on how well the classifier correctly guesses the identity of elements being classified, and can be expressed by different metrics, two of which (Gini index and cross-entropy) will be introduced later.

The weak learner is commonly chosen to be a decision tree of one or few nodes in depth. This choice – though – is not mandatory and the learner can assume very diverse forms (a Fisher linear discriminant, a k-Nearest Neighbor discriminant, a Support Vector Machine [71]...). The only requirement for a weak learner to be used successfully in boosting is that it should perform consistently better than a random choice (i.e. $p_{\text{error}} < 0.5$) [72].

During training, the boosting process generates a *forest* of weak learners, each of which is assigned a weight. This weight is used to obtain the weighted average of the outputs of each element in the forest, which will be the output value of the boosted classifier.

The weakness requirement is important: it implies that no learner is overwhelmingly important with respect to the others so it can be modified without sharply influencing the overall result, as this would complicate the convergence of the algorithm.

The weakness of the learners also protects the training process from *overtraining* the classifier. Overtraining is the condition in which the classifier response is optimized specifically on the individual points composing the training dataset, and not anymore on the general characteristics that the training dataset is meant to exemplify. If such a situation arises the question the classifier is answering is not the one we are interested in – namely, separating signal from background – and so the separation performance is reduced. Weak learners have generally few parameters, so they tend to be trained much faster and their convergence is safer than in the case of strong learners. This allows the boosting process to progress in an acceptable time.

There are various boosting algorithms available, here two of them are briefly mentioned.

3.1.1. Adaboost

Adaboost is one of the first boosting algorithms developed [73] [74] and works as follows: initially the training sample has all events assigned a unitary weight. Note that this weight is different from the weight assigned to the weak learners.

The weak learner parameters are fitted to give the best separation performance (the definition of separation can be chosen) to obtain the first weak learner of the forest. Events that are misclassified will have their weight enhanced in the next round by a factor which depends on the frequency of misclassifications of the weak learner h_m, where $m = 1 \dots N$ is the index of the iteration of the algorithm.

Letting x_i be the input parameters for an event, w_i the weight assigned to it and y_i the correct output we want to train, a training dataset with a total of N events can be written as $(x_i, y_i, w_i)_{i=1\dots N}$. We define the coefficient α_m for the weak learner m in the sequence as:

$$\alpha_m = \log\left(\frac{1-\epsilon_m}{\epsilon_m}\right) \qquad \text{where} \qquad \epsilon_m = \frac{\sum_i w_i \, \chi(h_m(x_i) \neq y_i)}{\sum_i w_i}, \tag{3.1}$$

where χ is the indicator function, so $\chi(h_m(x_i) \neq y_i) = 1$ when the learner classification is wrong, and 0 otherwise. The recursive expression for the weights of the events when training the $(m+1)$-th weak learner is (notice the normalization condition)

$$w_i^{(m+1)} = \frac{w_i^{(m)} \; e^{\alpha_m \, \chi(h_m(x_i) \neq y_i)}}{\sum_i w_i^{(m)} \; e^{\alpha_m \, \chi(h_m(x_i) \neq y_i)}}. \tag{3.2}$$

Thus, events that are misclassified will have a greater influence on the choice of the parameters for the next weak learner. This way, as the training steps proceed, various parts of the parameter space that are improperly classified by previous learners are "put under the spotlight" by the training algorithm, and a good classifier for them is optimized and added to the forest.

Each weak learner h_m will get a weight α_m in the final classifier, to obtain a combined *strong* classifier of the form:

$$H(x) = \sum_{m=1}^{N} \alpha_m h_m(x). \tag{3.3}$$

Due to the large number of outputs from weak learners being averaged together, the output value of the boosted classifier appears as a continuous variable. A value close to $+1$ will indicate that the event has a high degree of belief of being a real signal event and, equivalently, a value close to -1 indicates a high degree of belief that the event is a background one. It is then possible to choose the *working point* of the selector by imposing a cut on the BDT output variable. A working point is a choice of the cut corresponding to a specific compromise between signal efficiency and signal purity.

3.1.2. Gradient Boost

Formulating Adaboost as the minimization of a loss function led the way to the development of a training process known as *gradient boost* [75][76]. Unlike Adaboost,

this boosting algorithm does not update the weights of the training set. Instead, it interprets the iterative updates of the weak learners as converging steps to the solution of a variational problem, that is, the minimization of a suitably written loss function. The algorithm is described in more detail in [77]. Gradient boosting can attain a better performance than Adaboost due to the favorable properties of the improved loss function it uses [69], which is more robust to outliers in the training set.

3.1.3. Boosted Decision Trees

A boosted decision tree is a boosted strong classifier that uses a large number of binary trees as weak learners – typically between a few hundred and a few thousand. The binary tree is a logical structure that divides the input parameter space in regions and assigns to each of them an output score.

Figure 3.1 shows a graphical representation of a decision tree. The parameter space is initially divided in two parts at the root node by imposing a single-sided cut on the best separating variable, then the selected region is again divided in two parts by a cut on another variable and so on. This fission proceeds recursively until a node satisfies certain conditions, at which point an output score is returned. The output score at an end node corresponds to the label for the category that is most represented in that node.

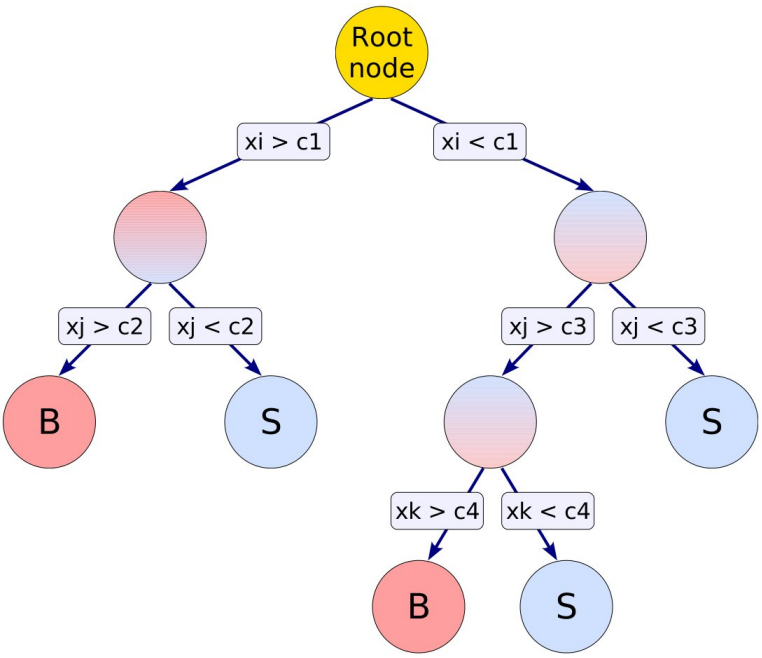


Figure 3.1.: An example of a binary tree from the TMVA guide [78].

The conditions for stopping the node splitting process are chosen as a compromise between the coverage of the rule, i.e. how large is the region of the input space on which a rule applies, and the signal/background purity achieved at the node. The first condition translates in the requirement that the node will be split only if it contains a number of events above a chosen minimum. This constraint also protects against overtraining, as the node does not get optimized against a small number of specific

events. The second condition implies that if the purity in background or signal events is below a certain threshold the node is split. Purity can be expressed for signal (P_s) and background (P_b) as a function of the weights w_i of the events in the training dataset in the following way:

$$P_s = \frac{W_s}{W_s + W_b}, \quad P_b = \frac{W_b}{W_s + W_b} \quad \text{with} \quad W_s = \sum_{i \in s} w_i, \quad W_b = \sum_{i \in b} w_i \tag{3.4}$$

The cut used in node splitting is chosen to maximize the separation between signal and background in the child nodes, and there are different ways to express this separation: a common one is considering the negative cross-entropy

$$Q_{xS} = P_s \log P_s + P_b \log P_b \tag{3.5}$$

and an alternative is using the Gini index

$$Q_{\text{Gini}} = P_s(1 - P_s) + P_b(1 - P_b) = 2P_s(1 - P_s) = 2P_b(1 - P_b) \tag{3.6}$$

as they both have a maximum at $P_s = 0.5$ (i.e. the category cannot distinguish between s and b) and a minimum value when purity (i.e. selection performance) of one of the two categories is maximal.

Boosted Decision Trees are implemented in the Toolkit for Multi Variate Analysis (TMVA) [78], which is part of the ROOT analysis framework [79]. Both the Adaboost and the Gradient Boost methods are available in the toolkit and in this thesis they are used for the identification of electrons and for the separation of genuine Higgs boson $\phi \to \tau\tau$ decays from Standard Model backgrounds.

3.2. Primary Vertex Reconstruction

In the typical experimental conditions at the LHC many protons collide in the same bunch crossing. Each of these collisions produces particles that can enter the detector, and their position in space is denominated *vertex*.

Proton collisions in a bunch crossing occur independently between each other. At the same time, interesting ones represent only a minuscule fraction of the overall number of collisions taking place in the experiment, and when a particular proton collision has produced an interesting final state, this collision will be surrounded by a number of trivial, non interesting, ones (*pile-up collisions*). The particles originating from these trivial collisions need to be excluded efficiently, to reduce their impact on measurements. The interesting collision is defined as the one for which the sum of the transverse momenta of its tracks is maximum, and is named *primary vertex*.

Vertices are reconstructed from a collection of good quality tracks selected from the event. This selection takes care that the tracks are reconstructed with high precision and that the frequency of "fake" tracks is low. Fake tracks or particles are wrongly reconstructed objects that do not correspond to a real ones, and are artifacts produced by the reconstruction and identification algorithms in presence of noisy or ambiguous data recorded by the detector.

The vertex reconstruction can be split in two stages: first, an optimal set of vertices is chosen as the origin for the provided tracks, then the position of each vertex is determined by a common fit on the tracks assigned to it. The first stage is solved by the Deterministic Annealing (DA) algorithm [80], while the second stage uses a weighted least squares method called "adaptive fitting" [81] that is more robust with respect to outliers.

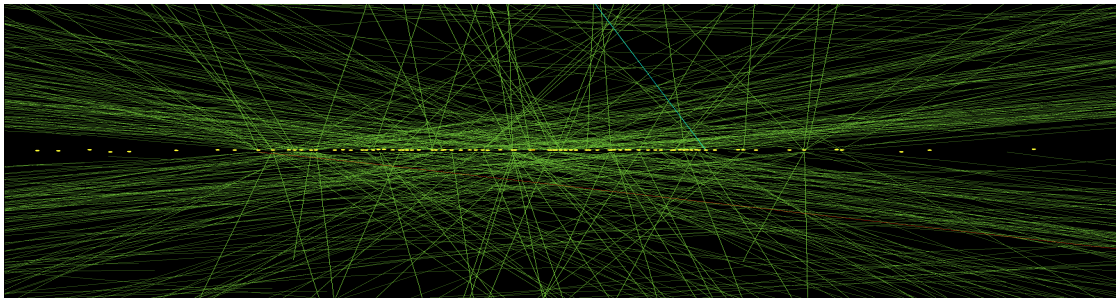


Figure 3.2.: An event in the LHC run in 2012 where a large number of primary vertices (78 of them) are identified by the reconstruction algorithm. From [82].

Cuts for vertex quality are applied: vertices are required to be at a maximum distance of 24 cm along the z axis from the nominal detector center, at a maximum radial distance of 2 cm from the beam spot in the xy plane and the number of degrees of freedom in the fit, in which their position was extracted, needs to be larger than 4.

In figure 3.2 a particularly remarkable event is shown. – which is often shown as an example of the good performance of the vertex reconstruction algorithm – in which as many as 78 primary vertices are reconstructed from a sea of thousands of tracks. In the figure, only the highest p_T tracks are shown in the picture, as it would be difficult to see anything if all the tracks were drawn.

3.3. Particle Flow Reconstruction

The Particle Flow algorithm allows combining the information provided by all the subdetectors to improve the identification of particles beyond the capabilities offered by a single subdetector, to provide a consistent global description of the collision and to get an improved measurement of their kinematic observables. This algorithm [83][84][85] was developed, drawing on the experience from *Energy Flow* algorithm, initially developed at the Aleph experiment at LEP [86][87] and refined in later experiments.

An example of the improvement arising from the use of the Particle Flow algorithm is in the determination of the p_T of jets. The dependence of resolution as a function of calorimetric energy for hadrons is shown in figure 3.3. This resolution is equal to $\Delta E/E = 104\%/\sqrt{E[\text{GeV}]} \oplus 6.7\%$ [83]. For low p_T hadrons this resolution degrades significantly. One way to improve the resolution exploits a property of jets: around two thirds of the particles in a jet are charged hadrons, and charged hadrons can be tracked with high precision by the silicon tracker. This is exactly what Particle Flow does: the algorithm combines the measurement of the p_T of the charged constituents of a jet, using the tracking detectors, with the measurement of the energy deposits in the calorimeters, thereby significantly reducing the uncertainty on the jet p_T, as the

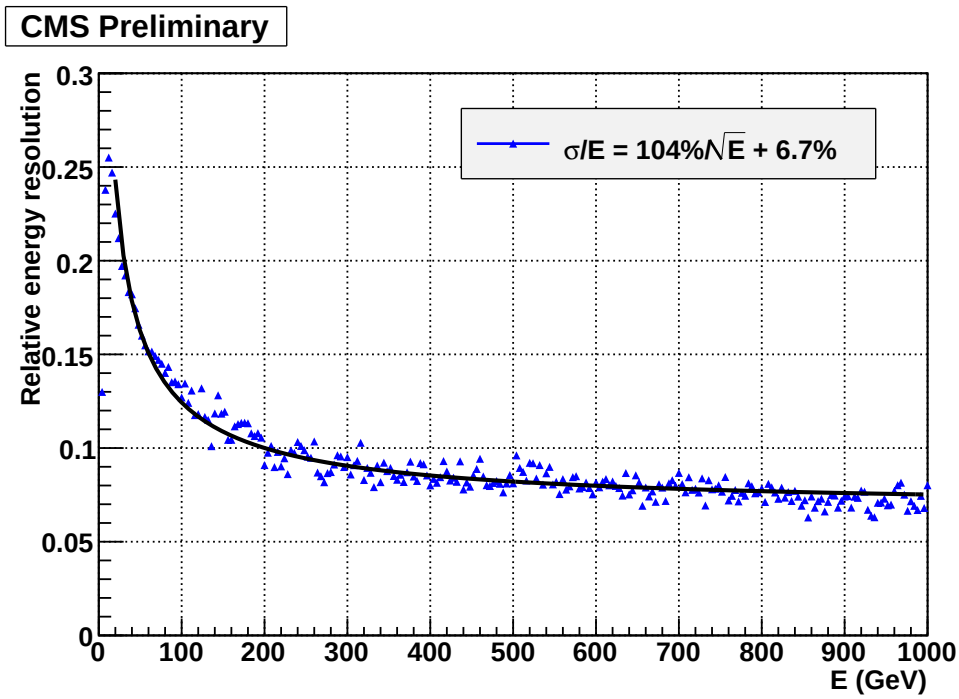


Figure 3.3.: Energy resolution of CMS calorimeters for hadrons. From [83].

calorimetric error is now relative only to the neutral fraction of the jets. In Figure 3.4 the error in the estimation of jet p_T is compared between jets measured with the calorimeters only and jets measured by Particle Flow. The PF estimation shows to have a smaller bias, and the resolution is improved.

Particle flow is divided in four main steps. First, an iterative tracking algorithm [88] based on a Kálmán filter (KF) is run on the set of hits observed in the silicon tracker. To ensure a good MET performance the efficiency for tracks has to be high and in parallel their fake rate has to be kept under control. The iterative tracking algorithm initially sets very tight cuts on the identification of tracks, then removes those hits which were associated to previously reconstructed tracks and runs again with loosened cuts on the remaining set of hits. By repeatedly loosening the cuts and – in parallel – removing already associated hits a high efficiency is obtained and the fake rate remains low (since there is less and less hits to build fake tracks from).

The second step involves clustering the energy deposits in the ECAL and HCAL [83]. Cells of the calorimeters with a measured energy above a cell-dependent threshold act as seeds for the energy clusters.

In the third step, clusters are processed by a *link algorithm*, which relates reconstructed tracks to clusters of energy in the ECAL and HCAL. A graphical view of this process is shown, for example, in figure 3.5. In the figure, tracks are shown in green and ECAL and HCAL clusters are colored red. A π^+ is identified from its reconstructed track (T2), the absence of deposits in the ECAL and the observation of a shower in the HCAL (cluster H2). There is also a dual cluster in the ECAL (E2, E3), that identifies the photons from a π^0, and a short (i.e. straight, energetic) track T1 that starts a large shower in the ECAL (cluster E1) and continues it into the HCAL (cluster H1), identified as a $\pi-$. An isolated ECAL cluster E4 is deduced to be a K_L^0 that decayed

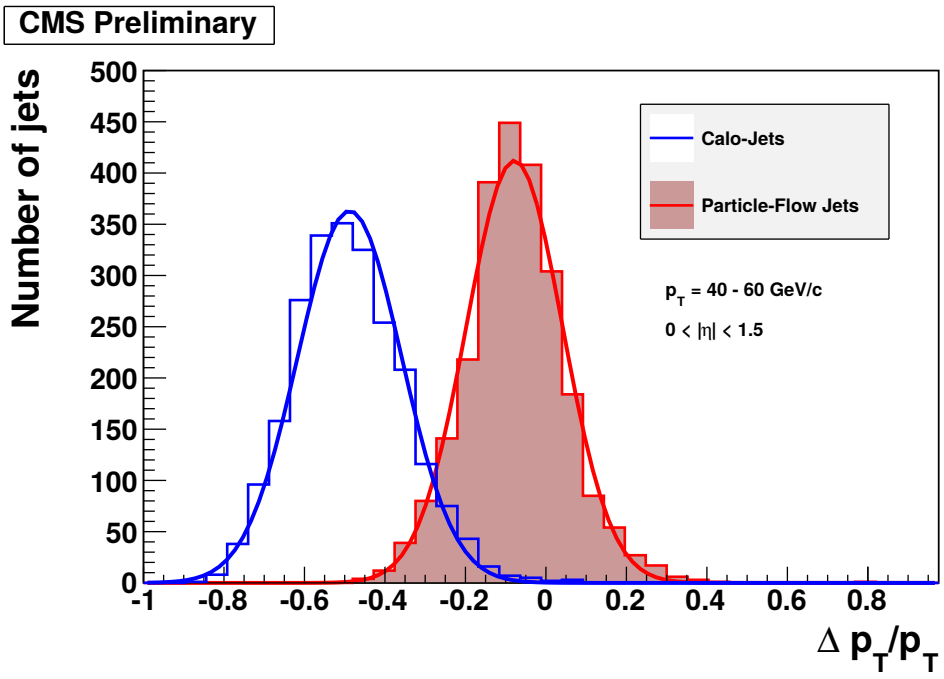


Figure 3.4.: Difference between Monte Carlo true value of jet p_T and reconstructed value for calorimeter jets (red) and Particle Flow jets (blue). From [83].

to $3\pi^0$ from the shape of its shower.

In the fourth step muon candidates are built from previously reconstructed muons and their expected energy deposits from the calorimeters are removed. What follows is the building of electron candidates, where their trajectory is followed up to the surface of the ECAL, potential radiated bremsstrahlung photons are associated to them and finally their energy deposits in the ECAL are used – together with track information – to compute their energy. The electron clusters are then removed from the set of clusters not yet associated to a particle. Unassociated ECAL clusters are then identified as photons, and remaining tracks ending in the calorimeters are identified as charged hadrons. Finally, energy deposits in the HCAL that were not yet accounted for are identified as neutral hadrons.

The availability of a consistent and complete description of all the stable particles emerging from an event allows to derive conclusions on the unstable particles that decayed in flight before being detected, including neutrinos that manifest themselves as $\not\!\!E_T$. For example, the CMS subdetectors are positioned too far from the primary vertex to directly observe the trajectory of a τ lepton before it decays into hadrons, but the observation of particles compatible with the decay products of a τ allows to reconstruct it indirectly. This process could have been potentially be performed also by each dedicated analysis, but the result would have been specific to that analysis. Particle Flow, instead, provides a reconstruction of τ leptons decaying hadronically that is consistent across different analyses.

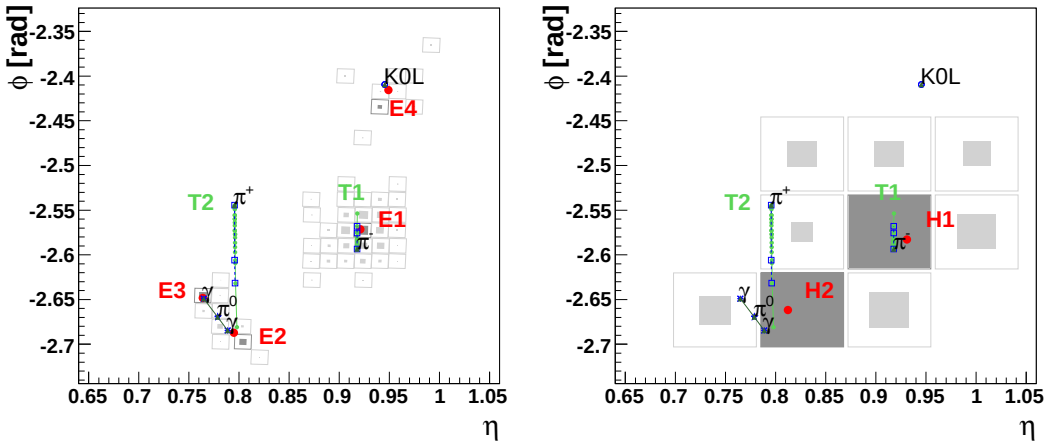


Figure 3.5.: A view in the (η, ϕ) plane of the track linking step of the Particle Flow algorithm. In the left picture the tracks (in green) are shown against clusters identified in the ECAL (red labels). In the right picture the same tracks are shown against clusters identified in the HCAL (red labels). From [83].

3.4. Jets and their Reconstruction in Particle Flow

When quarks and gluons – collectively referred to as partons – are produced in high energy collisions, they rapidly generate showers of hadrons, and those showers are called "*jets*". The development of a jet is described by quantum chromodynamics (QCD), and a more detailed treatment is available in [2]. The evolution of partons into jets happens at a length scale that is too small to allow their direct observation. For this reason, partons can only be reconstructed indirectly, through measurements of observables associated to jets.

While the jet develops, hard partons tend to radiate gluons with small momentum (*soft gluons*). These gluons in turn produce low energy hadrons that cluster around the energetic ones, i.e. those associated to the initial hard partons making up the core of the jet. Soft gluons can make the separation between jets difficult, as the ones associated to a jet often overlap with the soft gluons associated to other jets. A jet reconstruction algorithm which is stable in terms of energy and direction of the jet against the presence of soft radiation from neighboring jets is said to be *infrared-safe*.

Single partons have also a relatively high probability of splitting into two partons with energy of the same order of magnitude and with similar direction of flight. This process occurs randomly, cannot be observed directly and is only weakly related to the properties of the initial parton. The definition of a jet should take this effect into account, such that the number of jets in an event and their energy does not depend on random splittings, but only on the initial partons and their energy. Jet reconstruction algorithms that are insensitive to splitting are called *collinear-safe*. Besides a proper definition of the energy, collinear-safe algorithms are chosen because collinear-unsafe algorithms pose problems in the theoretical calculations [89].

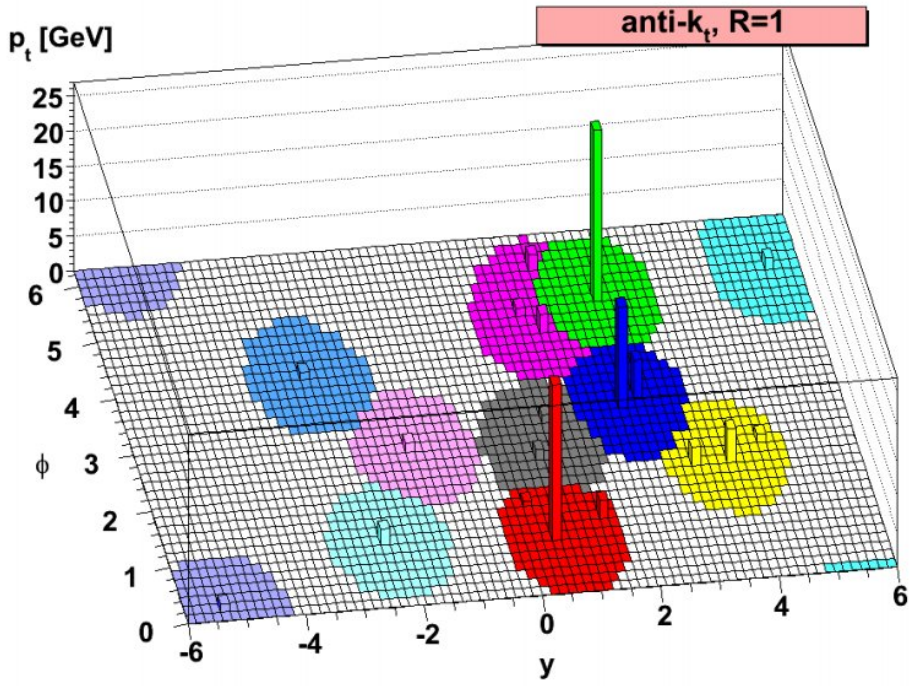


Figure 3.6.: An example of jet clustering with the anti-k_t algorithm. The reconstructed jets are shown as colored regions. From [89].

Additional challenges for jet reconstruction algorithms are the presence of the particle beams circulating in the accelerator and of additional proton-proton collisions in the same bunch crossing (pile-up). Beams are surrounded by particles produced in interactions between the components of the accelerator and the beam itself. These particles, together with jets produced in pile-up collisions, are a source of noise and good reconstruction algorithms should be able to discriminate them.

Finally, the output of a jet clustering algorithm should not depend on the order in which the input objects to be clustered are provided as input.

3.4.1. Anti-k_t Jet Algorithm

The standard algorithm used in CMS for the clustering of jets is the anti-k_t algorithm. This algorithm derives from the k_t [90, 91] and Cambridge/Aachen [92, 93] algorithms, and shares with them most of the concepts.

One fundamental concept is the definition of a metric which allows to decide which particles (also referred to as "pseudojets") are near to each other, so they can be clustered together. In the case of the anti-k_t algorithm this metric depends on the transverse momentum of the particles, which is indicated as k_t. Calling d_{ij} the distance between particle i and particle j, and d_{iB} the distance between particle i and the colliding proton beam, the following definitions apply:

$$d_{ij} = \min(k_{t_i}^{2p}, k_{t_j}^{2p}) \cdot \frac{\Delta_{ij}^2}{R^2} \tag{3.7}$$

$$d_{iB} = k_{t_i}^{2p} \tag{3.8}$$

$$\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2 \tag{3.9}$$

where the transverse momenta of particles i and j are denoted by k_{t_i} and k_{t_j} respectively and their distance in the rapidity-azimuth plane Δ_{ij}. R is a maximum jet size parameter and p is a momentum weighting parameter, these last two parameters define the overall behavior of the algorithm.

The anti-k_t algorithm [89] proceeds by considering the distances between pairs of particles i and j. If there are particles for which $d_{ij} < d_{iB}$, the closest pair is merged together (i.e. the merged particle has a 4-momentum which is the sum of i and j), while if for a given particle i there is no j such that $d_{ij} < d_{iB}$, the particle is declared a jet and removed from the particle collection. This process repeats until the particle collection is empty and all input particles have been merged into jets.

The parameter p controls the weighting of particles of different transverse momenta k_t, and thus where the clustering is started. The difference between k_t, anti-k_t and Cambridge/Aachen algorithms stands in the value of this weighting parameter, where other aspects are essentially identical. In the case of the k_t algorithm, this value is $p = 1$ and the clustering starts from the softest particles. In the case of the Cambridge/Aachen algorithm the value is $p = 0$, and only the shape of the energy deposits influences the clustering process. Finally, in the case of the anti-k_t algorithm $p = -1$ and the clustering starts with the hardest particles, to which the softest ones are then progressively "attached", mimicking in reverse the parton fragmentation and gluon emission processes. The advantages of using the anti-k_t algorithm stem from the fact that is collinear- and infrared-safe, it is simple and the definition of jet it uses can be used properly in theoretical calculations to make predictions.

The example in Figure 3.6, taken from [89], shows a result of the anti-k_t algorithm with a cone parameter $R = 1.0$. In such a picture the jets are round, have a regular boundary and are generally well separated, because soft gluon radiation has only a small effect on the contour of the jet. This property of the algorithm is called "soft-resiliency" [89]. Even in cases where there is an overlap between the clustering areas of two jets, the algorithm gives priority to the highest k_t jet, which is usually the most interesting. These are the properties for which this specific algorithm has been chosen.

In the case of the Particle Flow jets used in this analysis, the cone size parameter is set to $R = 0.5$, which is the most common and tested setting in Run I (2009-2013) for CMS.

3.4.2. Jet Calibration

Following the clustering stage, corrections [94] are applied to the jets to equalize the response for different background conditions and in different regions of the detector.

The first correction uses the jet area method [95]. This correction accounts event-by-event for sources of diffuse background and noise, such as particles produced in vertices (*pile-up vertices*) other than the primary vertex, for low energy particles (*underlying event*) associated with the collision at the primary vertex, and for noise in calorimeter electronics and nuclear interactions in the calorimeter material induced by pile-up and previous collisions.

Further corrections [94] to the jet energy account for the varying response of the detector across its geometry and the overall absolute energy scale.

3.4.3. Jet Identification

In some cases the combination of the Particle Flow and jet clustering algorithms can create a jet candidate which does not correspond to a physical jet. To suppress wrongly reconstructed jets, a series of loose identification criteria, referred to as *jet loose id* are applied to all jet candidates.

Jets are composed of charged hadrons, neutral hadrons and photons. Normally, the fraction of energy carried by neutral particles (neutral hadrons + photons) is 1/3 of the total energy of the jet. Charged hadrons carry the remainder 2/3 of the energy. This property can be used to identify other kinds of particles that were misreconstructed as jets. In CMS, less than 99% of the energy of a jet has to be carried by neutral hadrons and less than 99% by neutral particles interacting in the ECAL. Jets are also required to be composed of more than one particle.

In regions where tracking is available, i.e. where $\eta < 2.4$, it is required that the energy fraction carried by charged hadrons is larger than zero, and the energy fraction carried by charged particles interacting in the ECAL is less than 99%. Each jet must be composed of at least one charged particle. These latter cuts are designed to discriminate cases where a photon or a lepton is misidentified as a jet.

The conditions at the LHC during normal operation are harsh: around 15 to 25 protons collide in each bunch crossing (the average in 2012 was 21, and there were exceptional examples of 78 *reconstructed* primary vertices such as the one shown in figure 3.2). In these conditions it becomes complicated at times to associate jets to their correct primary vertex, when two or more vertices are close to each other.

While many collisions take place in a proton bunch crossing, only one of them per crossing may involve the production of heavy particles such as Higgs bosons, top quark pairs or W boson pairs. Indeed, the cross section for pp inelastic collisions at the LHC has been measured by TOTEM [97] to be 74.7 ± 1.7 mb and this cross-section is dominated by low energy QCD events. As a comparison, the production of a top quark pair and its decay into a pair of leptons has been measured by CMS to be $239 \pm 2(\text{stat.}) \pm 11(\text{syst.}) \pm 6(\text{lum.})$ pb [98]. This means that only around 1 in 300'000'000 pp collisions produces a top quark pair, and even smaller fractions are expected in the case of a MSSM Higgs boson, since the cross sections are expected to be smaller.

Jets originating from low energy QCD events in pile-up vertices show kinematic and particle distributions that are different from those originating from hard collisions. Therefore, a boosted decision tree has been trained in CMS to distinguish between jets originating in interesting collisions from jets originating in pile-up collisions. This boosted decision tree is given the name "Pileup Jet ID".-

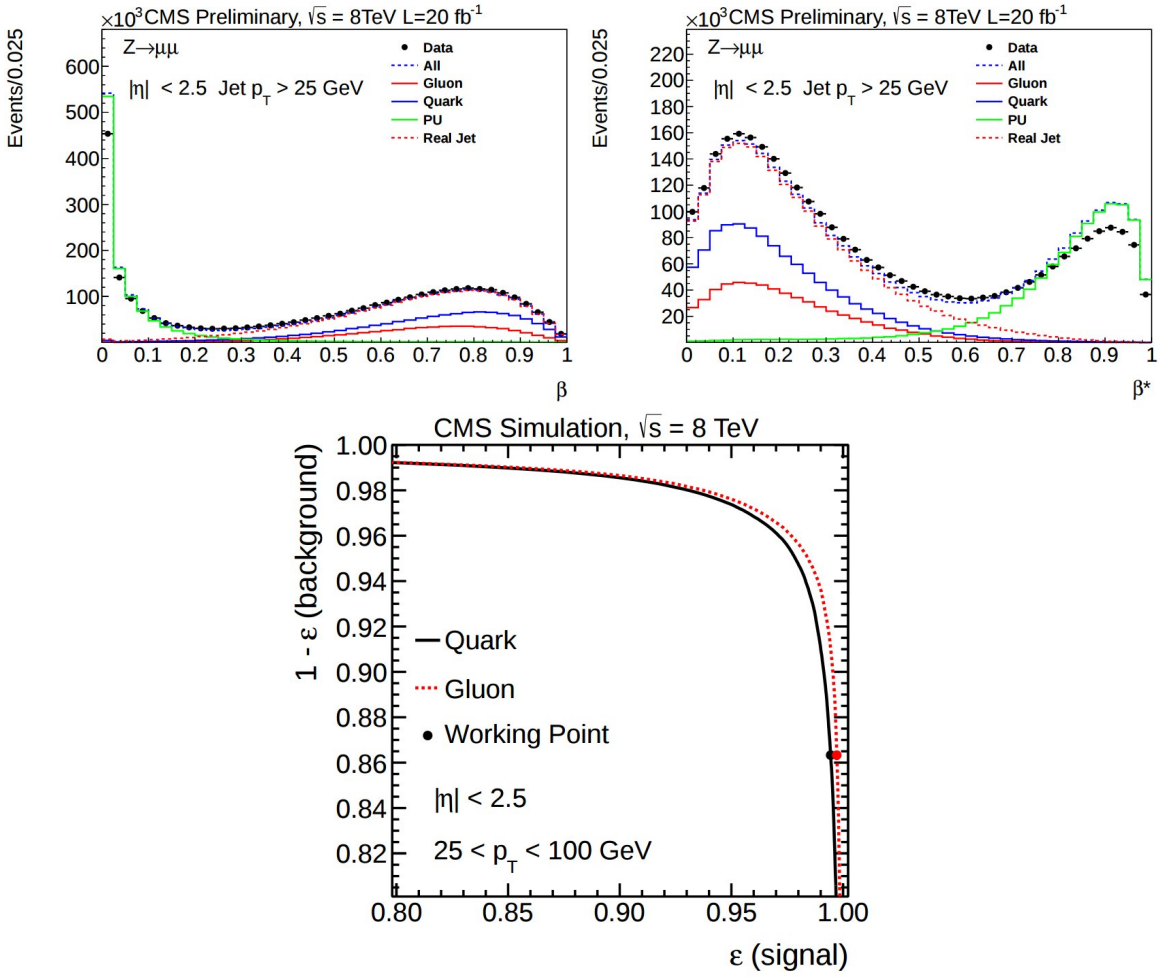


Figure 3.7.: Distribution of β (top left) and β^* (top right), for the different cases of jets originating from the primary vertex (dashed, in red and labeled "real jets") and those from pile-up jets (solid, in green). The solid blue and red lines show the two components, gluon jet and quark jet, of real jets. Also shown is the ROC curve for the Pile Up Jet ID BDT (bottom). All plots from [96].

In figure 3.7 (top left and right) the most separating variables, β and β^*, used as input for such a BDT are shown. The variable β represents the fraction of p_T in the jet carried by charged particle tracks associated to the selected primary vertex, while the variable β^* is the fraction of p_T in the jet carried by charged particle tracks associated to other primary vertices.

$$\beta = \frac{\sum_{\text{PV}} p_{T_i}}{\sum_{\text{all}} p_{T_i}} \qquad \beta^* = \frac{\sum_{\{\text{other PV}\}} p_{T_i}}{\sum_{\text{all}} p_{T_i}}. \tag{3.10}$$

In an event not necessarily all tracks are associated to a vertex, therefore in general the sum of β and β^* may be different than unity. More documentation about the PileUp Jet ID is available in [96].

At the bottom of figure 3.7 the ROC curve for the Pileup Jet ID BDT is shown, both for hard quark jets and hard gluon jets plotted against pile-up background jets. The loose working point of the Pileup Jet ID BDT, used in this analysis, is shown as a dot in the ROC curve in figure 3.7.

The ROC curve (Response Over Characteristic) curve is a 2D diagram used to represent the performance of a filter in separating signal from background. The ROC curve shows the efficiency for signal on one axis and that for background on the other axis as the parameters of the filter are varied, e.g. in the case of the ROC curve relative to the cut on a variable acting as a filter, the ROC curve is drawn as the cut threshold is progressively tightened. An ideal filter has a working point where the efficiency for signal is 100% and that for background is 0%. Real filters usually do not reach perfection, but the ROC curve can help in finding the best working point as a compromise between signal yield and purity of the sample.

3.4.4. B-Tagging

The b-associated production channel for a MSSM Higgs boson, $pp \to b\bar{b}\phi$, introduced in chapter 1, has a characteristic signature, where the two b-quarks are produced with a small p_T. A b- or $\bar{b}$-quark produced in the final state lives long enough to form hadrons. A single b-quark will hadronize into a B meson – which can be a $B^{\pm}$, B^0 (each 40.2% of the times [11]), B_s^0 (10.5% of the times) or a b-baryon (9.2% of the times).

B-hadrons have a relatively long lifetime, allowing these particles to travel a measurable distance in the inner part of the detector before decaying into other hadrons, forming a jet which is displaced with respect to the primary vertex where the b-quark was created. For example, a B^+ (mean life $\tau = 1.638 \cdot 10^{-12}$ s and mass $m_0 = 5.27926$ GeV/c^2 [11]) with a momentum $p = 15$ GeV in the rest frame of the experiment on average travels a distance ℓ before decaying

$$\ell = \tau c \beta\gamma = 1.39 \text{ mm} \qquad \text{with} \quad \gamma = \frac{\sqrt{p^2c^2 + m_0^2c^4}}{m_0c^2},\ \beta\gamma = \sqrt{\gamma^2 - 1}$$

This distance can be measured in the silicon tracker, the challenging part being to distinguish the tracks associated to the secondary vertex from other tracks. This can be difficult with a large number of pile-up vertices in close proximity. A series of algorithms have been developed in CMS to maximize the efficiency and purity of secondary vertex finding.

The algorithm used in this analysis is named Combined Secondary Vertex (CSV), which combines in a multivariate analysis discriminator information such as the impact parameter of tracks in the jet, the resolution on this impact parameter for each track, the invariant mass of the tracks constituting the jet, the number of tracks and the relative pseudorapidity of the tracks with respect to the the jet axis. The CSV algorithm provides the best possible performance among the available algorithms [99] [100].

3.5. Reconstruction and Identification of Electrons

The typical signature of an electron in CMS consists of hits in the silicon tracker and one or more clusters in the electromagnetic calorimeter (ECAL). For this reason, the reconstruction of electrons in CMS makes use of the information collected primarily by these two detectors. Other detectors participate in the task of distinguishing real electrons from background, most of which arises from the activity of jets.

Electrons are very light particles ($m_e \simeq 1/222\, m_\mu \simeq 1/273\, m_{\pi^\pm}$). They are therefore subject to relatively large changes in their directions (kinks in the trajectory) and losses of energy each time they scatter against an atom and emit a bremsstrahlung photon. The amount of material present in the CMS tracker layers is considerable, and in cases where an electron radiates a bremsstrahlung photon, as a consequence of traversing one of such layers, the energy loss can be very large: at an energy of 10 GeV around 35% of the electrons lose more than 70% of their energy via bremsstrahlung when passing through the tracker, and in 10% of the cases the energy loss reaches more than 95% [101]. This dramatic effect is visible in figure 3.8 (left) where the histogram shows the fraction of the initial energy of the electron carried by photons.

As the loss of energy due to bremsstrahlung can be very large, a correct reconstruction of the energy of electrons needs to take into account these losses. To obtain a correct estimation of the initial electron energy, the energy deposited by the electron in the ECAL crystals is summed with the energies deposited by the bremsstrahlung photons in the same detector. The matching of the photons to the electron requires the photons to lie along tangents to the electron trajectory, as shown in figure 3.8 (right). A precise reconstruction of the electron trajectory is important for this matching to be efficient.

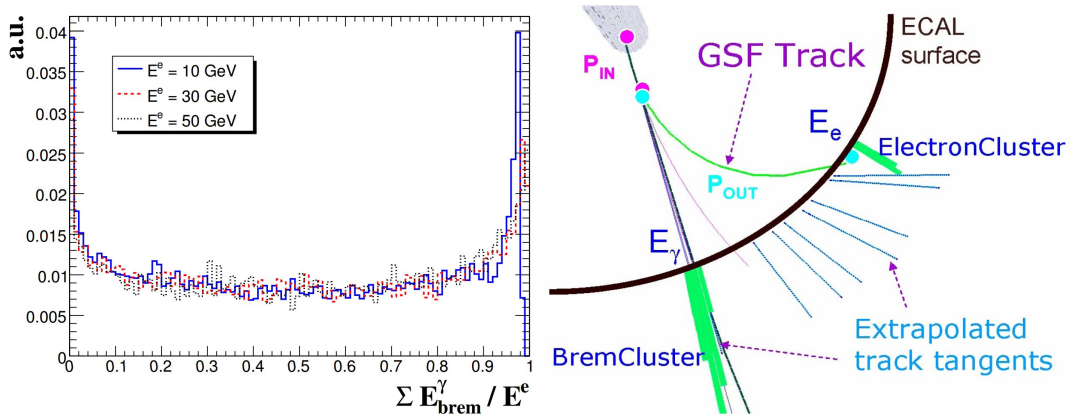


Figure 3.8.: (left) Fraction of electron energy carried away by bremsstrahlung photons after it traversed the silicon tracker, for different electron energies. (right) The bremsstrahlung photons are associated to the electron if they are tangent to the expected electron track. In that case their energy is added into to that of the electron.

The standard implementation of the Kálmán filter – used in the iterative tracking

step of the Particle Flow algorithm – models the trajectory deviations of charged particles, when passing through matter, with a single Gaussian distribution. This approximation assumes that the combined effect of many deviations, each due to the electron traversing a layer of the tracker, is much larger than the one due to each single deviation.

As discussed in the previous paragraph, in the case of electrons this assumption does not hold due to the phenomenon of bremsstrahlung, which often leads to kinks in the electron trajectory. These kinks are not modeled properly by a single Gaussian model.

An attempt to improve the description of the trajectory deviations of electrons is provided by a generalization of the Kálmán Filter: the Gaussian Sum Filter algorithm. This algorithm approximates the distribution of deviations predicted by the Bethe-Heitler model for electron bremsstrahlung with a sum of Gaussians [81].

The GSF algorithm is used by default in all electron reconstruction applications, including the Particle Flow processing chain. The electron reconstruction can be seeded by energy clusters in the ECAL (more precisely, the position of the baricenter of the energy deposits) or by pairs of hits in the pixel detector, compatible with observed ECAL clusters.

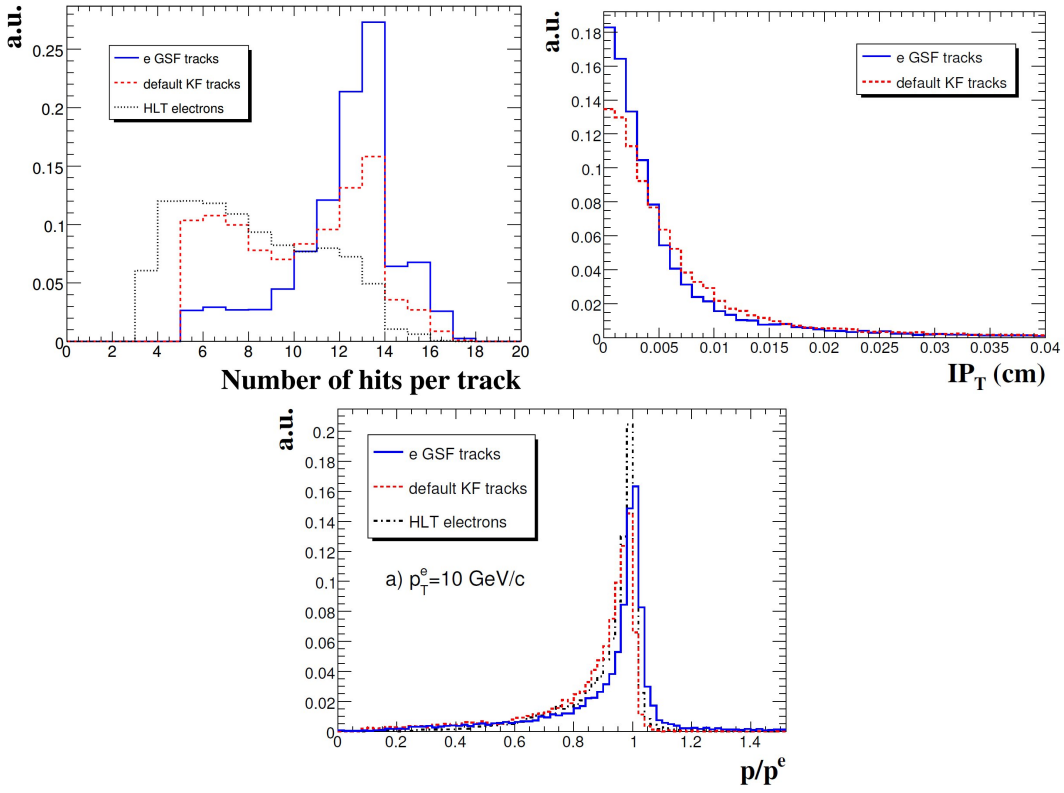


Figure 3.9.: Comparison between electrons reconstructed with KF (in red) and with GSF (in blue) filters. [a, top left] Number of hits associated to an electron track. [b, top right] Impact parameter of electrons in μm. [c, bottom] Ratio between measured and true Monte Carlo electron momentum. All from [101].

At each iteration step (i.e. tracker layer traversal), the variables describing the particle state are assumed to be drawn from a distribution which is a weighted sum of many

Gaussians. The mixture of Gaussian distributions to use for a given iteration is preset into a look up table, optimized so the mixture of Gaussians approximates optimally the Bethe-Heitler model. More details about this implementation are available in [102, 103, 101].

In figures 3.9a and 3.9b the number of hits associated to the electron track and its measured impact parameter with respect to the primary vertex are compared for the cases where the KF or the GSF algorithms are used in reconstruction. The GSF clearly is able to match more tracker hits, and achieves a more precise measurement of the track parameters than the KF.

In figure 3.9c the ratio between the reconstructed total momentum of the electron and its true Monte Carlo value is shown. It is possible to notice that, while using the KF, the measurement is biased towards lower values, i.e. some clusters that are significant to the energy computation are lost. The use of the GSF improves significantly the measurement, shifting the peak towards unity and reducing the spread of the distribution.

A feature specific to Particle Flow is that electrons whose identity is ambiguous are initially pre-filtered by running a KF reconstruction with loose cuts, then reconstructed again using the GSF algorithm (this two-stage process was designed to optimize computation time) and finally selected by a Boosted Decision Tree classifier. Further studies and improvements in the availability of computing power allowed to train an improved Boosted Decision Tree, which could be run on the entire collection of electrons, both ambiguous and directly identified by their good track-ECAL deposit matching, increasing the overall efficiency for electrons. This improved Boosted Decision Tree is discussed in the next section.

3.5.1. MVA-based Electron Identification

Certain physical processes taking place in the detector can present a signature that mimics the one of an electron. Common examples are photons converting into an electron-positron pair in the first layer of the tracker detector, narrow jets initiating showers inside the ECAL and a lone $\pi^{\pm}$ undergoing a charge exchange reaction in the ECAL crystals.

The latter case is particularly difficult to distinguish from a real electron: the pion is charged and leaves hits in the tracker as do electrons. At the same time, the charge exchange reaction, which can be written as $\pi^+ + n \rightarrow p + \pi^0$ or $\pi^- + p \rightarrow n + \pi^0$, produces a neutral pion. The latter decays into photons, which produce the same electromagnetic shower as an electron. Only subtle differences, mainly linked to the fact that electrons radiate more bremsstrahlung photons, can be used to distinguish this kind of misidentified pions from electrons. There is no single variable that can be used to effectively discriminate fake electrons from real ones with a good efficiency-purity compromise. Many variables can be combined, though, to build a multi-variate discriminant.

A BDT is trained to give an output, as is customary, of +1 on a sample of real

electrons and of -1 on a sample of fake electrons. Training is performed on a sample where the background is represented by $W + 1$jet events extracted from data – where the jet is faking an electron – while the signal events are represented by a sample of $Z \to ee$ Monte Carlo events. The BDT implementation available in TMVA [78] is used.

TMVA offers a wide range of settings for training a BDT, in this specific case the AdaBoost algorithm is used to train a forest of 800 trees with a maximum depth of 3 stages and with a limit of 1000 nodes in the whole forest. The reduction of nodes at each iteration is performed by a pruning algorithm, based on the "*CostComplexity*" cost function. The TMVA parameter for pruning strength has been chosen to have the value 5. The metrics used for node cut optimization is the Gini index. Each tree node is required to contain at least 100 events from the training dataset. In order for the cut point in each node to be well optimized, the number of sampling points for Gini index estimation has been increased from the default value of 20 [78] to 2000.

The electron identification BDT accepts as input a collection of variables, their distribution in the case of background (hatched red) and signal (uniform blue) being shown in figures 3.10 and 3.11, to give a final score for each considered electron. The variables are:

- The p_T and η of the electron, as the response of the detector changes in different geometrical regions (e.g. in the gap between the barrel and endcaps identification efficiency is degraded) and with different momenta of the electron.

- The χ^2_{KF} value of the trajectory fit, when is reconstructed with the Kálmán Filter, and the χ^2_{GSF} value of the trajectory fit, when the Gaussian Sum Filter is used. Due to the trajectory kinks resulting from bremsstrahlung the χ^2 values of the two algorithms tend to differ more between each other in the case of real electrons.

- The number of hits associated to the track when being reconstructed using the Kálmán Filter.

- The distances in η and ϕ, commonly referred to as $|\Delta\eta_{\text{in}}|$ and $|\Delta\phi_{\text{in}}|$, between the baricenter of the electron *supercluster* (i.e. the collection of the electron cluster plus all its associated bremsstrahlung clusters) and the expected impact point on the ECAL surface, extrapolated from the *innermost* state, at the vertex, of the reconstructed track.

- The distance in η, referred to as $|\Delta\eta_{\text{out}}|$, between the electron *cluster* and the expected impact point on the ECAL surface, extrapolated from the *outermost* state of the reconstructed track, i.e. the trajectory state at the last tracker layer with a valid hit.

As introduced previously in the section dedicated to the Gaussian Sum Filter used in tracking, electrons radiate frequently bremsstrahlung photons. In addition, both electrons and photons tend to deposit all of their energy in the ECAL without reaching deeper into the HCAL. Finally, the shape of the shower they produce has certain peculiar properties. These properties are condensed into a set of shower-related variables, which are fed into the BDT:

- The ratio of hadronic to electromagnetic energy, expressed as $E_{\mathrm{HCAL}}/E_{\mathrm{SC}}$, where E_{SC} is the energy of the ECAL supercluster, and E_{HCAL} the sum of the energies of the clusters lying in the HCAL behind the ECAL supercluster. In regions covered by the ECAL preshower detector (i.e. the endcaps), the ratio between energy deposited in the preshower detector and the one deposited in the ECAL is also used.

- The fraction of energy carried by bremsstrahlung photons, $f_{\mathrm{brem}} = (p_{\mathrm{in}} - p_{\mathrm{out}})/p_{\mathrm{in}}$, where p_{in} and p_{out} are the electron momenta at the vertex and at the exit from the tracker.

- A variable exploiting the loss of electron energy as bremsstrahlung, defined as $1/E_{\mathrm{SC}} - 1/\bar{p}_{\mathrm{GSF}}$, where $\bar{p}_{\mathrm{GSF}}$ is the average between the electron momenta corresponding to each trajectory state (i.e. layer with associated hits).

- Variables testing the consistency between the fitted track and the energy deposits in the ECAL. One is the ratio $E_{\mathrm{SC}}/p_{\mathrm{in}}$ and the other is the ratio $E_{\mathrm{e}}/p_{\mathrm{out}}$, where E_{e} is the energy of the cluster closest to the electron track, extrapolated from the last tracker hit onto the ECAL surface (i.e. the cluster most likely associated to the electron flying into the ECAL).

- The ratio of the energies collected in a 3×3 crystals square, centered on the most energetic crystal, and the total energy of the supercluster E_{SC}, which is given the name "R9", and tends to have a value approaching 1 for electrons since the shower tends to be compact.

- The variable defined as $(E_{5\times5} - E_{1\times5})/E_{5\times5} = 1 - \frac{E_{1\times5}}{E_{5\times5}}$, where $E_{5\times5}$ is the energy collected in a square measuring $5\eta \times 5\phi$ crystals and centered on the most energetic crystal and, similarly, $E_{1\times5}$ is the energy collected in a $1\eta \times 5\phi$ strip. This variable measures how much the energy deposit draws a "streak" along ϕ onto the ECAL surface, which indicates that the particle is charged and has a curved trajectory as a result of the magnetic field.

- The RMS widths in ϕ and η of the electron supercluster.

- The variables $\sigma_{i\eta i\eta}$ and $\sigma_{i\phi i\phi}$, which measure how the shower spreads in a 5×5 crystals square around the most energetic ECAL crystal. Letting $\langle\eta\rangle_{5\times5}$ and $\langle\phi\rangle_{5\times5}$ be the baricenters in η and ϕ of the energy deposition in the 5×5 square and i an index pointing the crystals of such square, the two variables are defined in the following equations:

$$\sigma_{i\eta i\eta} = \sqrt{\frac{\sum_{i\in5\times5} w_i \cdot (\eta_i - \langle\eta\rangle_{5\times5})^2}{\sum_{i\in5\times5} w_i}} \quad \text{and} \quad \sigma_{i\phi i\phi} = \sqrt{\frac{\sum_{i\in5\times5} w_i \cdot (\phi_i - \langle\phi\rangle_{5\times5})^2}{\sum_{i\in5\times5} w_i}},$$

$$\text{with logarithmic weights} \quad w_i = \max\left(0,\ 4.7 + \ln\frac{E_i}{E_{5\times5}}\right).$$

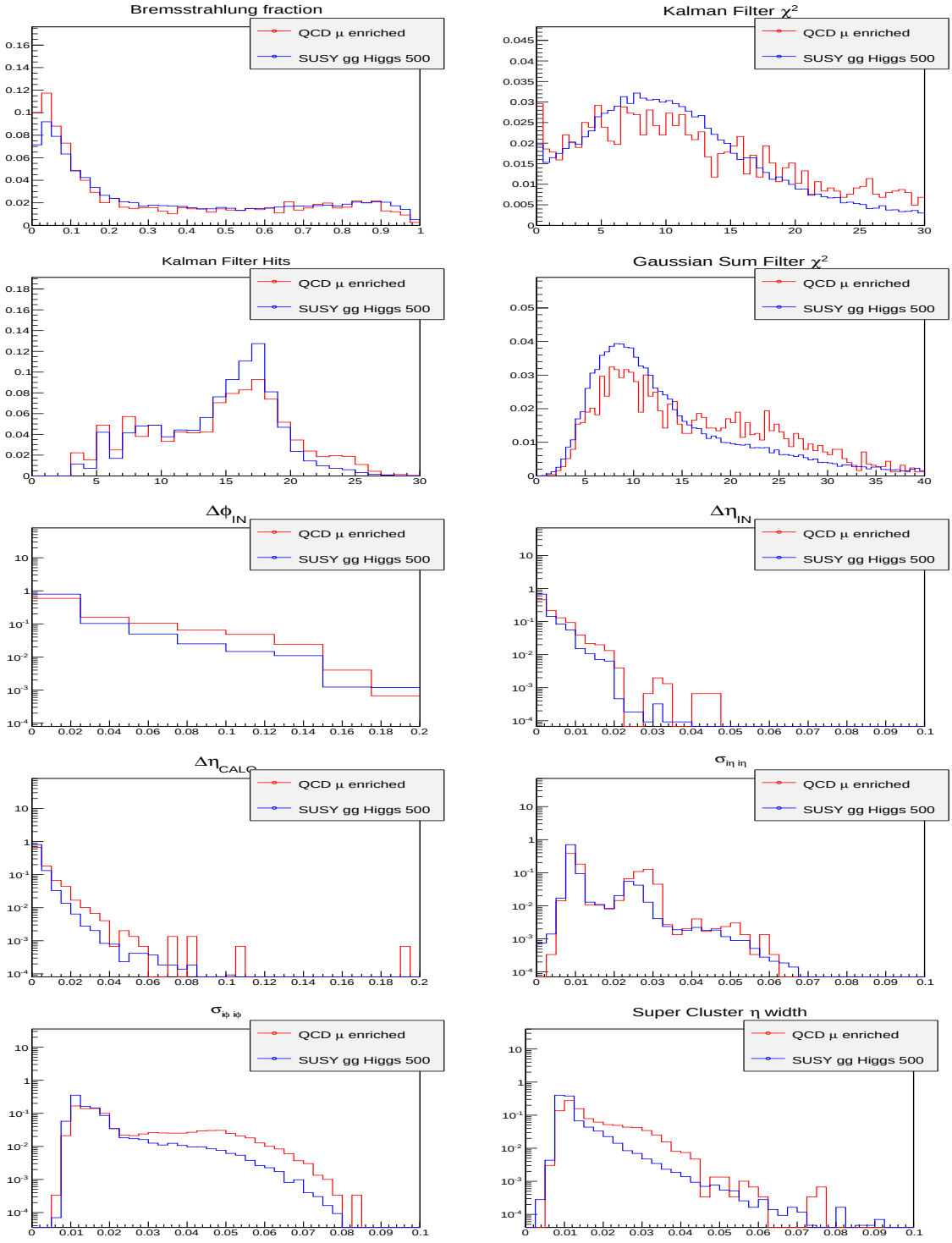


Figure 3.10.: Variables used as input for the electron identification boosted decision tree. Their description is discussed in section 3.5.1.

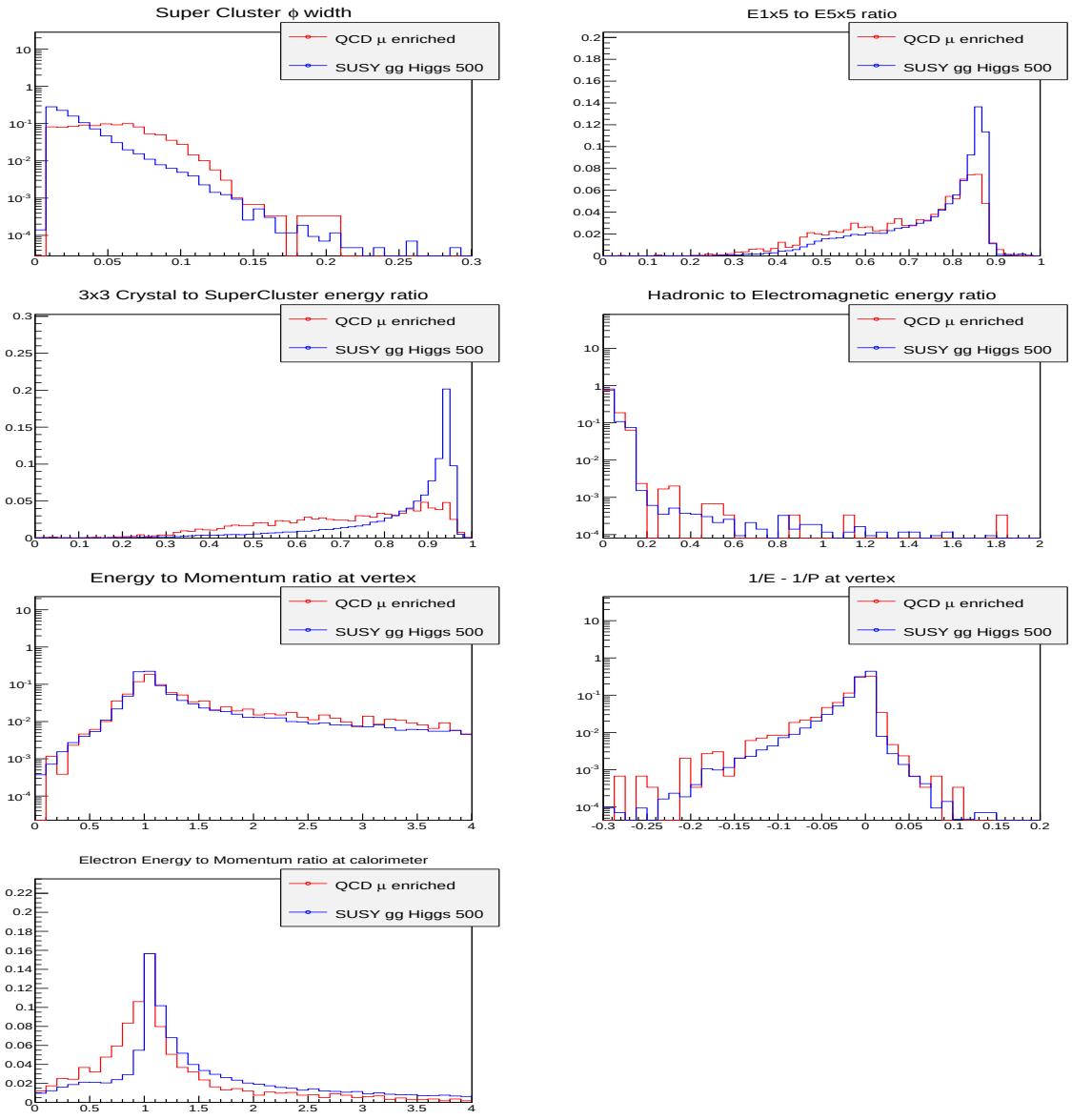


Figure 3.11.: Variables used as input for the electron identification boosted decision tree. Their description is discussed in section 3.5.1.

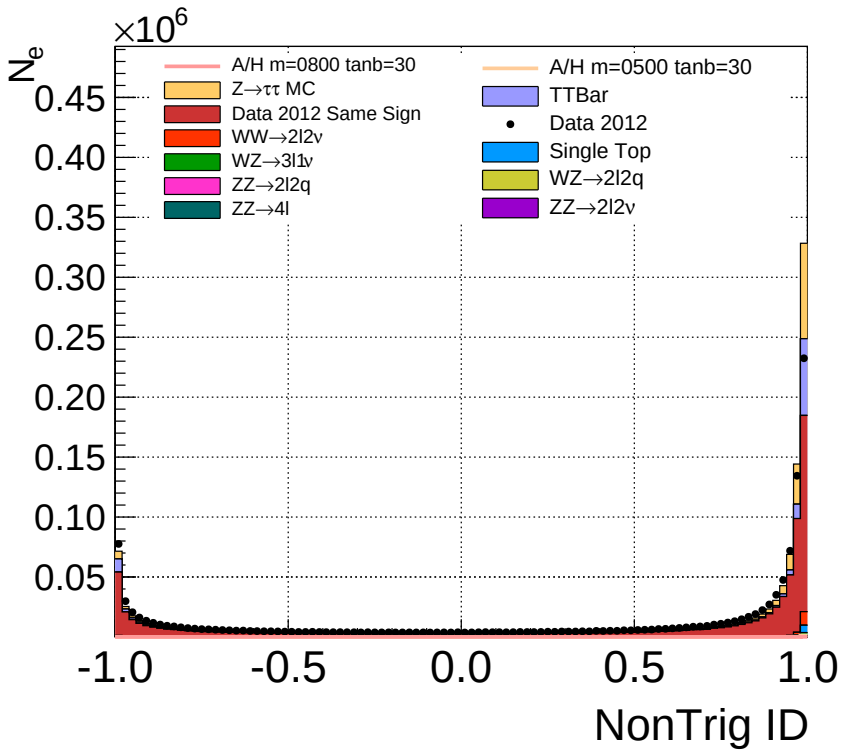


Figure 3.12.: Output of the electron identification BDT.

In figure 3.12 the output of the electron identification BDT is shown, for events that passed muon quality selection cuts. The data/simulation comparison uses the $Z \to \tau\tau$ Monte Carlo sample to represent the Z, as the exclusion cone used to simulate τ decays in embedded samples biases optimistically the electron distribution. The discrepancy between data and Monte Carlo is explained by the fact that weights (and, in particular, pile-up reweighting) is not used for this plot.

In figure 3.13 the significant improvement in terms of performance of the same BDT is shown in comparison to cut-based methods, where simple cuts have been imposed on most of the variables discussed above. The BDT allows to decrease considerably the amount of fake electrons while keeping the efficiency for real electrons constant.

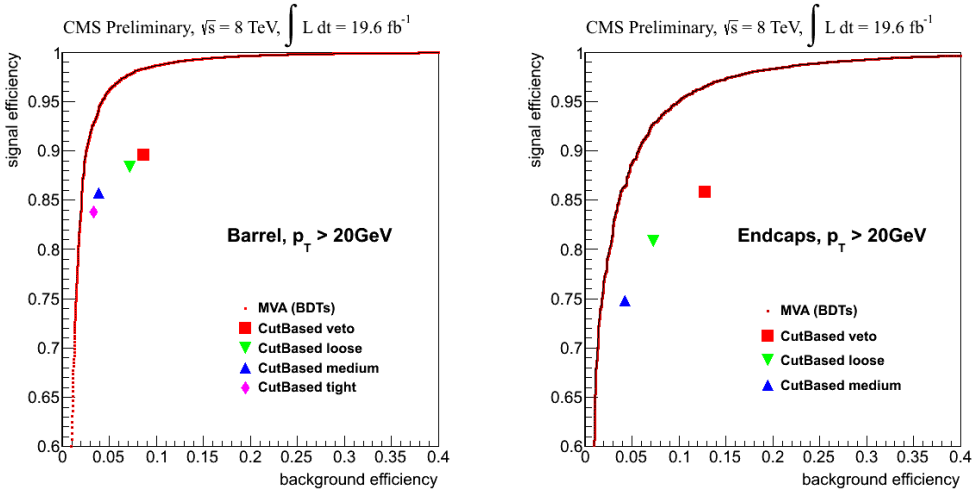


Figure 3.13.: Efficiency for real electrons vs. efficiency for fakes in the barrel (left) and endcap (right) comparing the performance of the electron ID BDT and the one of cut-based selections used before its introduction. The higher the values of electron efficiency (Y axis) for a given efficiency for fakes (X axis), the better. From [104].

3.6. Muon Reconstruction

There are two ways to identify and reconstruct a muon in CMS: the muon can be reconstructed starting from a track in the silicon detector that has a matching track when its trajectory is propagated out into more external detectors, or it can be reconstructed starting from a track identified in the muon system, and propagate its trajectory inside to find a matching track in the tracker system.

Since the muon trajectory is mainly influenced by multiple scattering, which shows a single Gaussian distribution of the trajectory deviations, it is possible to use an iterative track finder coupled with a Kálmán filter for the track reconstruction.

The two possible approaches, propagation *inside-out* and *outside-in* of the trajectory, lead to the definition of four types of muons:

- **Standalone muons** are muons that have been reconstructed from hits in the muon system, independently from tracks reconstructed in the inner tracking system. They tend to have a high purity, but the precision in the measurement of their 4-momenta is limited, due to the limitations of the muon system and the effects of multiple scattering in the yoke material. Concerning muons originating from collisions, their number is also very small (around 1%) as the inner tracking system is very efficient. It is thus very unlikely that a muon reaching the outer muon chambers would have failed to be reconstructed in the tracker. Standalone muons are indeed, most of the time, cosmic ray muons traversing the detector, as the acceptance volume of the muon system is $\sim 10^3$ times larger than the one of the tracker. Standalone muons are productively used in detector studies.

- **Calorimeter muons** are due to particles that left a track in the inner tracking system and deposited an amount of energy compatible with a muon in the

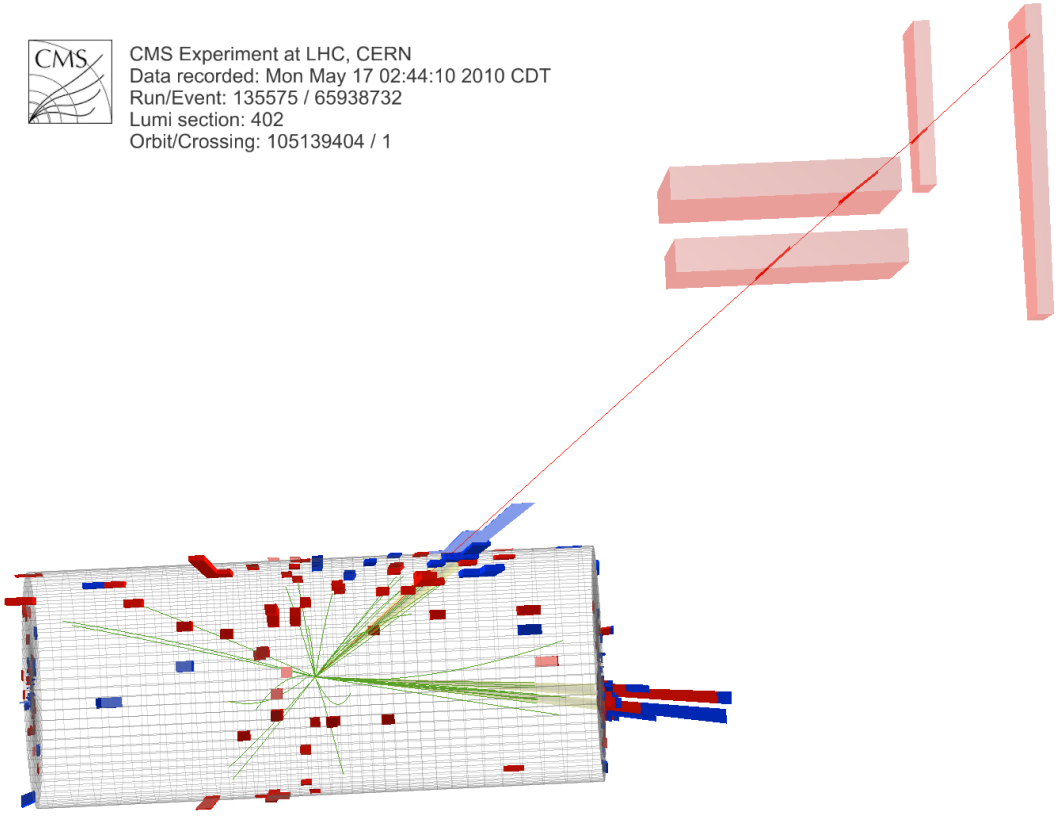


Figure 3.14.: An event display in which a Global muon traversed in its trajectory the endcap muon detectors (shown in red). From [105].

calorimeters. These muons have a very poor purity and should not be used in physics analyses, as they can be easily mimicked by other particles. They find use, though, in situations where a large number of low p_T muons are needed for detector studies.

- **Tracker muons** are seeded by a track in the inner tracking system and, by propagating their trajectory to the muon system, have an associated muon segment that has been found and matched to them. A muon segment is a line connecting matched hits in the DT and CSC system. These kind of muons have a good purity, and a very high precision in the determination of their 4-momentum, thanks to the precise measurement of their trajectory by inner tracking system.

- **Global muons** are Standalone Muons for which a matching track has been found in the inner tracking system. The Global muon track is refitted by combining the muon system and tracking system hits. This combination improves the precision on the determination of the muon 4-momenta, as at energies above $\sim$ 200 GeV the sagitta of the muon trajectory across the tracker – which is related to the p_T measurement – becomes so small that the tracker momentum resolution starts becoming dominated by the track spatial resolution. The addition of hits observed in the muon system helps counteract this loss of resolution, by providing a longer lever arm for the measurement of p_T. Global muons have a high pu-

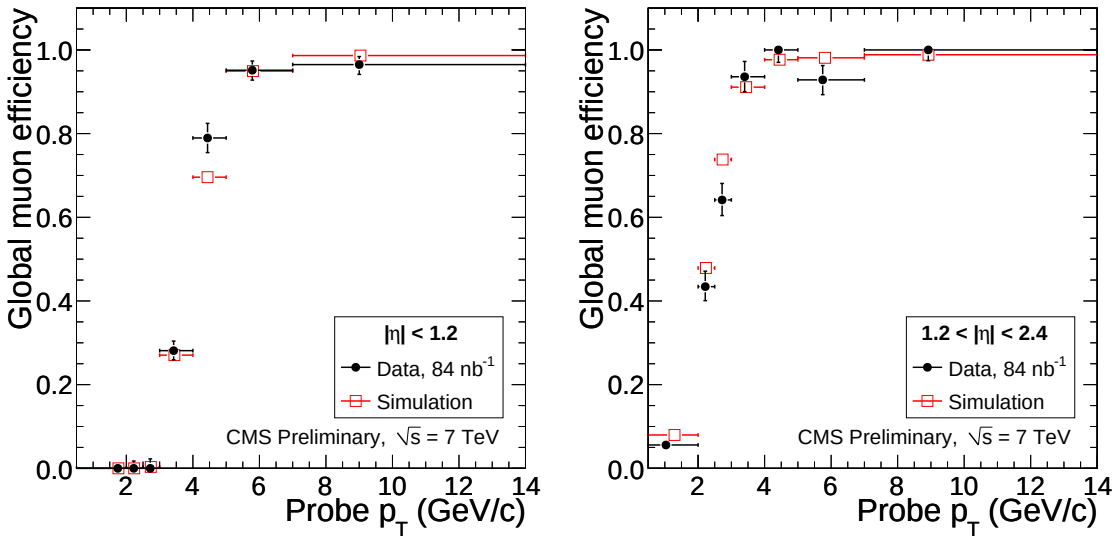


Figure 3.15.: Reconstruction efficiency for Global muons in the barrel (left) and in the endcap (right), in data (black) and simulation (red)[105].

rity, and while their efficiency is low or zero at low values of p_T, it approaches unity for muons above 8 GeV. This is visible in figure 3.15, which shows the reconstruction efficiency as a function of p_T in barrel and endcaps.

Following the availability of these reconstruction strategies, a set of three types of muon identification selection criteria have been set up, with increasing level of purity. They are

- The **Soft Muon** selection, where a muon is required to be reconstructed as a Tracker Muon and to have a matching segment in the outermost muon station that can be reached by its expected trajectory. Soft muons can be identified by this selection even if they fail to reach the last muon chamber layer because of their limited energy.

- The **Global Muon** selection, where a muon is required to be reconstructed as a Global Muon.

- The **Tight Muon** selection, where a muon is required to be reconstructed both as a Tracker Muon and a Global Muon. The p_T needs to be larger than 3 GeV/c as genuine muons below this threshold are not reconstructed properly as Global Muons. In order to limit the background from particles misidentified as muons and to filter muons whose reconstruction is of poor quality the ratio between the χ^2 and the degrees of freedom (i.e. the *normalized* χ^2) of the Global Muon track fit must be less than 10. For the same reasons, the Global Muon track fit has to use at least one hit from the muon system, there need to be at least two muon stations which are matched to the Global Muon track and the Tracker Muon track needs to have more than 10 associated hits in the silicon tracker. To reduce the background from muons originating from the decay in flight of hadrons, the impact parameter of the Tracker Muon track in the transverse plane has to be $|d_0| < 2$ mm.

Muon reconstruction is performed independently and before running the Particle Flow algorithm. Muons that are matched to the tracks found by the iterative track fitter of the Particle Flow algorithm are called Particle Flow muons (or "*PFMuons*" for brevity). The global description of the event by the Particle Flow algorithm, where all objects in the event are reconstructed in a consistent way, resolves ambiguities between tracks and reduces the frequency of charged hadrons misreconstructed as prompt muons. This improved identification leads to an improvement in the calculation of the missing transverse energy.

3.7. Lepton Isolation

Leptons can be produced directly in hard processes or in the decay of hadrons. Given the mechanism by which hadrons form in collisions at the LHC, the latter family of processes is usually associated with jets in the proximity of the leptons.

Isolation is a measurement of the amount of (hadronic) activity in the proximity of a lepton observed by the detector. **Absolute isolation** is defined as the sum of the transverse momenta carried by other particles in a cone $C(\ell, R)$ of radius R around the trajectory of the lepton ℓ and **relative isolation** I_{rel} is its ratio compared to the transverse momentum p_{T_ℓ} of the lepton itself. $I_{\text{rel}\,\ell}$ is defined as follows:

$$I_{\text{rel}} = \frac{\sum\limits_{C(\ell,R)} p_{T_i}}{p_{T_\ell}} = \frac{\sum\limits_{\text{charged}\in C(\ell,R)} p_{T_i} + \sum\limits_{\text{neutral had}\in C(\ell,R)} p_{T_i} + \sum\limits_{\gamma\in C(\ell,R)} p_{T_i}}{p_{T_\ell}} \tag{3.11}$$

where the three right-hand sums refer to charged particles, neutral hadrons and photons. The Particle Flow algorithm provides an easy way to obtain the isolation, as these sums can be calculated by summing over the transverse momenta of the various particles in the collections produced by the algorithm:

Isolation is a property that underlines the mechanism behind the production of a particle in a specific collision, and should not depend on particles produced in other ones. It is therefore desirable that the sums in equation 3.11 are corrected for the contribution of particles originating in pile-up vertices. While this pile-up removal is straightforward in the case of charged particles, as they can be traced back to individual vertices, for the case of photons and neutral hadrons the correction has to rely on an estimate. This correction is referred to as $\Delta\beta$.

It has been observed that a good estimate for the amount of transverse momentum carried by pile-up neutral hadrons and photons combined is half the transverse momentum carried by pile-up *charged* hadrons. The expression for relative isolation, taking into account the $\Delta\beta$ correction and dropping the notation for the cone becomes:

$$I_{\text{rel}\Delta\beta} = \frac{\sum\limits_{\text{charged}} p_{T_i}}{p_{T_\ell}} + \frac{\min\left(0, \sum\limits_{\text{neutral had}} p_{T_i} + \sum\limits_{\gamma} p_{T_i} - 0.5\cdot \sum\limits_{\text{PU charged had}} p_{T_i}\right)}{p_{T_\ell}} \tag{3.12}$$

In the specific case of this analysis, the recommended settings shared between many analyses in CMS are followed, where the radius of the cone in the $\eta \times \phi$ plane used for

the calculation of isolation is $R = 0.4$.

Electrons		
Particle type	Min ΔR from lepton	Min p_T [GeV]
Charged particles	0.01(barrel), 0.015(endcaps)	no cut
Photons	0.08	no cut
Neutral hadrons	no cut	no cut
Pile-up charged hadrons	no cut	no cut

Muons		
Particle type	Min ΔR from lepton	Min p_T [GeV]
Charged particles	0.0001	no cut
Photons	0.1	0.5
Neutral hadrons	0.1	0.5
Pile-up charged hadrons	0.1	0.5

Table 3.1.: Vetoes applied to the PF particles used in the computation of isolation for electrons (top) and muons (bottom).

Vetoes are applied to particles in the sums in equation 3.11, such that their contribution to the sum is not counted. Particles coming too close from the lepton of interest are vetoed, the required ΔR separation is shown in table 3.1. The veto on ΔR prevents double counting of p_T in cases where the energy deposits are misassociated, leading to the reconstruction of a fake particle instead of being assigned to the lepton.

Another veto is imposed on the minimum amount of p_T that a particle must have in order to be counted, and is shown in the same table 3.1. The veto on particle candidates below a minimum p_T avoids biases associated to the clustering of energy deposits in the vicinity of a lepton and the impact of pile-up.

3.8. Missing Transverse Energy

Missing Transverse Energy (MET) is a physical observable associated to the production of invisible particles as a result of the collisions, such as neutrinos.

In a proton-proton collision at the LHC, the momentum of the colliding particles is aligned with the $\hat{z}$ axis of the experiment. Due to conservation of linear momentum, the sum of the transverse momenta of the collision products is equal to the initial one,

$$\sum_{\text{initial}} \vec{p}_{T_i} = \sum_{\text{final}} \vec{p}_{T_i} \approx 0.$$

If the transverse momentum of the observed products is significantly different from the typical scale of the colliding partons, this condition can be interpreted as the result of one or more particles escaping undetected. The transverse momentum that would be

needed to balance to zero the sum of the collision products is given the name *Missing Transverse Energy* and indicated as $\vec{\not{E}}_T$ or MET,

$$\sum_{\text{observed}} \vec{p}_{T_i} + \vec{\not{E}}_T = 0. \tag{3.13}$$

This variable is expressed in the form of an energy for historical reasons, as at the time it was developed it was calculated by summing up the energy deposits in the calorimeters with the momentum of the muons, which escape the calorimeters.

In the case of the search for MSSM Higgs boson in the decay into a pair of τ leptons presented in this thesis, MET occurs due to the presence of neutrinos from the decay of the τ leptons, and is used to reconstruct the invariant mass of the neutrinos. MET is also used to build discriminators used in the selection of signal events. It is therefore beneficial to optimize the resolution in the determination of the MET, as this improves the sensitivity of the search in many ways.

Using only the calorimeters to calculate the MET, though, makes this calculation to depend on the resolution of the calorimeters. A more advanced approach integrates the Particle Flow algorithm in the calculation of the MET (and for this reason is given the name PF MET), leading to a significant improvements [106].

Particle Flow MET (or PF MET) uses the improved performance in terms of energy resolution of the Particle Flow algorithm to obtain the list of objects over which to compute the transverse momentum sum. The largest gains in performance arise from the ability of using the tracker to complement the energy measurement of charged hadrons (see section 3.3) and from the removal of double-countings that would arise from the use of a independent algorithms for each particle type.

$$\not{E}_{T,PF} = -\sum_{\text{PF candidates}} \vec{p}_{T_i} \tag{3.14}$$

PF MET has a limited ability to counteract the negative effects of pile-up on resolution, as the Particle Flow algorithm is able to distinguish particles coming from the primary vertex from those coming from pile-up vertices. Still, this selectivity can be improved, leading to better resolution, with a multi variate approach.

Pile-up interactions increase the number of particles over which the sum in equation 3.13 is performed. Many of these particles have low amounts of p_T, and are subject to potentially large reconstruction inefficiencies and relative p_T measurement errors. The combined effect of a large number of small losses and measurement errors add up, with the result that as the number of pile-up interactions increase, the MET resolution gets worse.

An important consideration is that pile-up vertices are low energy collisions which do not produce invisible particles. Their physical contribution to MET – aside from detector and reconstruction effects affecting the measurement – is close to zero. A good strategy to counteract the degradation on the MET resolution due to pile-up is therefore to correct for the contribution of wrongly assigned particles entering the sum in equation 3.13.

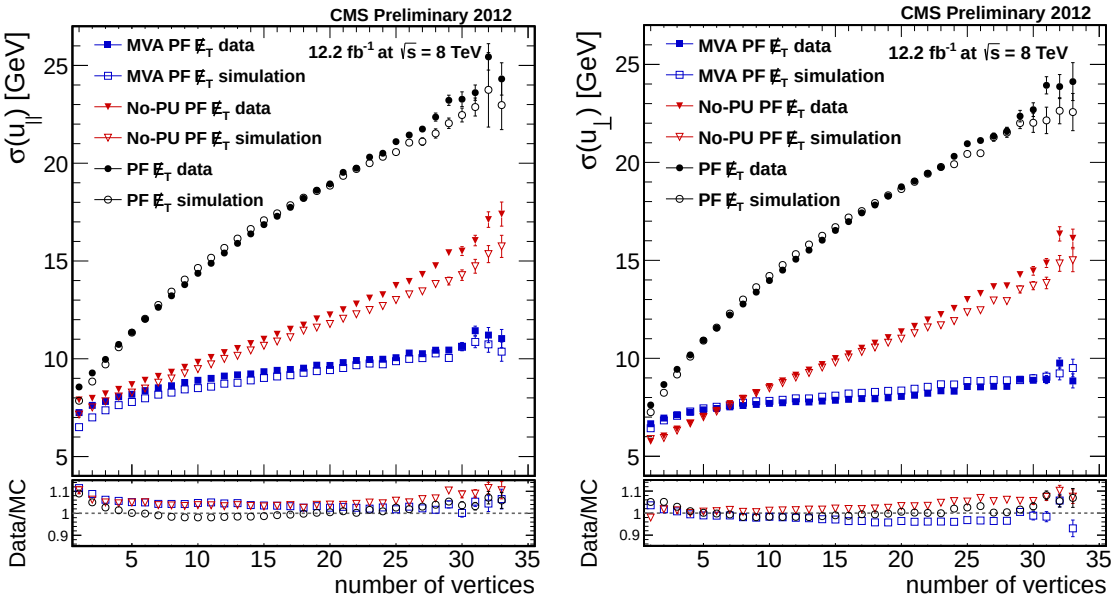


Figure 3.16.: MET resolution parallel (left) and perpendicular to the recoiling Z^0 for different algorithms as a function of pile-up: Particle Flow MET in black, No-PU PF MET (an early pile-up correcting algorithm) in red and MVA MET in blue [107].

A multi variate regression approach [107] uses the information provided by the Pileup Jet ID (described in 3.4.3) to associate jets and particles to the primary and pile-up vertices, estimating the contribution to MET from these two types of vertices, for charged particles only and for the sum of charged and neutral ones. It then combines these contributions to obtain the improved estimate of MET, which is given the name **MVA MET**.

The resolution on the MET is estimated from $pp \to Z + \text{jets} \to \mu\mu + \text{jets}$ events, in which the MET is expected to be close to zero when the muons and the jets are properly measured. The resolution of the components transverse and parallel to the momentum of the recoiling Z are shown as a function of pile up, for data and simulated samples, in figure 3.16. It is possible to see that the resolution on PF MET degrades significantly as the number of pile-up vertices increases, while the degradation in the case of MVA MET is smaller and has a smaller dependence on the number of vertices.

3.9. $\tau\tau$ Mass Reconstruction

The decay of a $\tau\tau$ system into a pair of charged leptons leads to the production of two neutrinos for each τ lepton, or four neutrinos in total. The 4-vector of each neutrino cannot be measured individually, as neutrinos escape undetected. Only the sum of the transverse component of the momenta of the four neutrinos is accessible as a measurement, manifesting itself as MET. While this loss of information on the momentum of each τ itself is not recoverable, it is possible to derive from a probabilistic argument the most likely rest mass of the $\tau\tau$ system, which becomes the Higgs candidate mass in this search.

Using such a reconstructed mass, referred to as $m_{\tau\tau}$, instead of the visible mass of the final state charged leptons, $m_{e\mu}$, has the advantage that resonances show up as narrower peaks. In the specific case of a Higgs boson with a mass around 125 GeV/c^2, the peak is significantly better separated from the $Z \to \tau\tau$ one, a process that overwhelms the Higgs one in the $m_{e\mu}$ case due to its much larger production cross section. In CMS the reconstruction of $m_{\tau\tau}$ is performed by a code named **SVFIT**.

3.9.1. The SVFIT Algorithm

The decay of a τ lepton to a visible lepton and two invisible neutrinos can be described by the mass of the τ, the 4-momentum of the visible lepton and the following parameters, which are unconstrained by the measurement since the neutrinos are not directly detected [27]:

- The fraction of τ lepton energy carried by the visible lepton, referred to as x.
- The azimuthal angle of the τ lepton trajectory in the laboratory frame, referred to as ϕ.
- The invariant mass of the neutrino system, referred to as $m_{\nu\nu}$.

A likelihood model [27] combines the knowledge of the matrix elements for each τ decay with the measurements on the visible leptons 4-momenta, the measurement of the components of the MET and their covariance matrix. The free parameters for the two τ decays are conventionally written as $\vec{a_1} = (x_1, \phi_1, m_{\nu\nu,1})$ and $\vec{a_2} = (x_2, \phi_2, m_{\nu\nu,2})$. The 4-momenta of the visible leptons are denoted by the vector $\vec{y} = (p_1^{\text{vis}}, p_2^{\text{vis}})$ and the components and covariance matrix of the MET by $\vec{z} = (\not{E}_x, \not{E}_y, \mathbf{V})$. The probability for a given invariant mass hypothesis on the $\tau\tau$ system $m_{\tau\tau}^i$ is:

$$P(m_{\tau\tau}^i|\vec{y},\vec{z}) = \int \delta(m_{\tau\tau}^i - m_{\tau\tau}(\vec{y},\vec{a_1},\vec{a_2})) \cdot f(\vec{z},\vec{y},\vec{a_1},\vec{a_2})\, d\vec{a_1}\, d\vec{a_2} \tag{3.15}$$

where $f(\vec{z},\vec{y},\vec{a_1},\vec{a_2})$ is the product of three likelihood functions. Two of them are related to the matrix element of each τ decay, and depend only on m_τ and $m_{\nu\nu}$:

$$\mathscr{L}_\tau = \frac{d\Gamma}{dx\, dm_{\nu\nu}\, d\phi} \propto \frac{m_{\nu\nu}}{4m_\tau^2}\left[(m_\tau^2 + m_{\nu\nu}^2)(m_\tau^2 - m_{\nu\nu}^2)\right] \tag{3.16}$$

The third likelihood function combines the measurement of the MET, its covariance matrix, the measurement of the momentum of the visible leptons and the free parameters $\vec{a_1}$ and $\vec{a_2}$, which enter through the expressions for the sums of neutrino momenta $\sum_\nu p_x$ and $\sum_\nu p_y$

$$\mathscr{L}_\nu = \frac{1}{2\pi\sqrt{\det(\mathbf{V})}} \cdot \exp\left[-\frac{1}{2}\begin{pmatrix}\not{E}_x - \sum_\nu p_x \\ \not{E}_y - \sum_\nu p_y\end{pmatrix}^T \mathbf{V} \begin{pmatrix}\not{E}_x - \sum_\nu p_x \\ \not{E}_y - \sum_\nu p_y\end{pmatrix}\right] \tag{3.17}$$

In figure 3.17 the normalized distributions of the mass are compared for a reconstructed $Z \to \tau\tau$ resonance and for a $H \to \tau\tau$ one, where the mass of the Higgs is assumed to be 125 GeV/c^2. The left plot shows the mass spectrum when using the mass of the visible products m_{vis} (or equivalently referred to as $m_{e\mu}$ in later chapters), while the right figure shows the distribution for the $m_{\tau\tau}$ reconstructed by SVFIT. There is a clear improvement in separation between the two peaks when using $m_{\tau\tau}$.

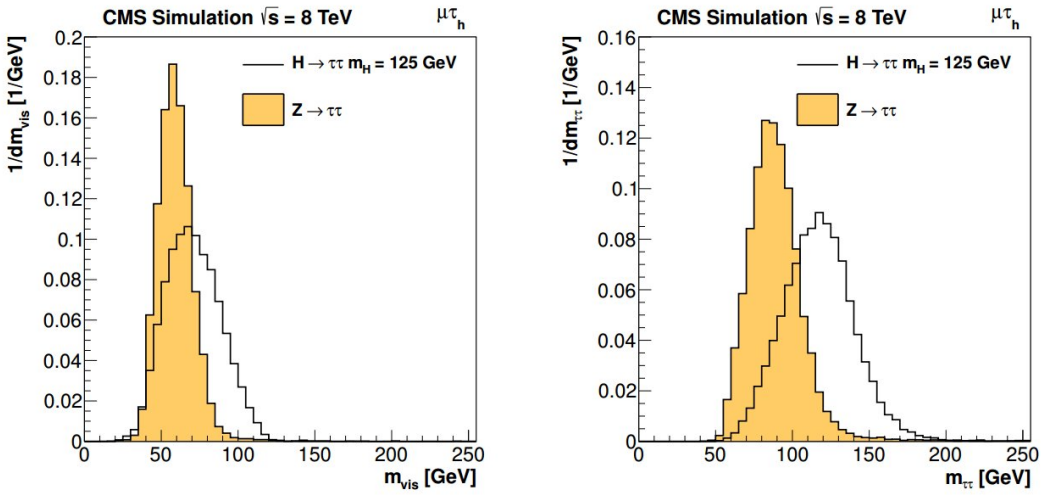


Figure 3.17.: The $Z \to \tau\tau$ and $H \to \tau\tau$ peaks, as resolved when using the m_{vis} (left plot) and $m_{\tau\tau}$ (right plot) variables. From [27].

3.10. Statistical Analysis Methods

When testing the validity of a specific physical theory predicting new phenomena ("signal") not described by already proven processes ("background"), the end result of an analysis is to either show beyond reasonable doubt that nature contradicts the predictions of that theory or show that the current accepted description of nature is not consistent with measured data and show that the measurements indicate that the new theory under test is compatible with the observations. The first case is defined as an *exclusion*, while the latter one is a *discovery*. This analysis covers the exclusion of MSSM Higgs models parametrized by the parameters m_H and $\tan\beta$, which were introduced in chapter 1.

The discriminant used to put a theory under test is a set of one or more parameters that are measured during the analysis. It is common convention to accept an exclusion if the probability that the exclusion may be attributed to known sources of fluctuation and experimental error in the measurement is less than 5%. In that case the exclusion is said to be at a $100\% - 5\% = 95\%$ Confidence Level, meaning that 95% of the times the excluded model is truly false and the exclusion identifies this condition correctly.

In the case of a discovery these boundaries are set to a much tighter level: the probability that the new observed effect be attributed to a fluctuation of should be less than the equivalent one-sided integral of a Normal distribution from a distance of 5σ from the center and beyond, which is around $2.87 \cdot 10^{-7}$. These two probability threshold define the level of reasonable doubt mentioned above.

Many different procedures have been developed over the years in the field of statistical analysis to estimate exclusion limits on the parameters characterizing a theory, such as production cross-sections and particle masses, and there is still an ongoing discussion about the advantages of using one or another. A common choice is the use of a modified frequentist method, commonly referred to as "CL_s" [108, 109].

The CMS Collaboration, together with the ATLAS Collaboration and the LHC Higgs Combination Group in 2011 agreed on a common statistical framework [110] to be used for exclusion limit estimation and Higgs signal discovery. This framework is based on the CL_s technique. The original version of CL_s requires a very large number of Monte Carlo toy pseudoexperiments to be generated, a feature which is computationally very intensive. The statistical framework agreed in [110], and used in the analysis discussed in this thesis employs a modification, referred to as the *profile likelihood asymptotic approximation* [111]. This approximation combines well-behaved results in most situations with a reduced need for computational power. It is briefly described below, following the notation in [111].

Suppose to write a model that combines the known physics (background) and additional phenomena predicted by the new theory (signal) in the modeling of the expected number of events n_i in each bin $i = 1 \dots N$ of a histogram in the measured variable that was chosen to perform the search (often this variable is the invariant mass of a system of particles). The expectation value for the number of events in a specific bin is:

$$\mathbb{E}[n_i] = \mu s_i + b_i, \tag{3.18}$$

where s_i is the expected number of signal events, b_i the expected number of background events in the bin. μ is called the *signal strength modifier* and parametrizes the problem at the core of the search exercise: if the theory under test is false and nature is fully described by the background its value is $\mu = 0$, if the theory is true it is valued $\mu = 1$. The parameters in the model that influence the predictions for s_i and b_i, such as the integrated luminosity of the experiment, the efficiency of the b-tagging algorithm, the theoretical assumptions on the constituents of the protons, are called *nuisance parameters* and are noted as the vector of values $\boldsymbol{\theta}$.

The likelihood as a function of the signal strength μ and the nuisance parameters $\boldsymbol{\theta}$, in view of the number of events counted in each bin n_i during the experiment is denoted as $L(\mu, \boldsymbol{\theta}(\mu))$ and is normally represented as a product of Poisson probabilities (see references). When keeping the value of μ fixed this function has a maximum in $\boldsymbol{\theta}$ which is denoted as $\hat{\hat{\boldsymbol{\theta}}}(\mu)$. If both the parameter μ and the parameters $\boldsymbol{\theta}$ are varied to attain the maximum of the likelihood, we call their values in this maximum $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$. The following likelihood ratio,

$$t_\mu = -2 \ln \lambda(\mu) = -2 \ln \left(\frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})} \right) \tag{3.19}$$

shows the disagreement of the model for a particular value of μ, among all the possible ones, given the measured data. This expression, though, does not take into account that some values of μ may be unphysical. For instance, in cases when the existence of an undiscovered resonance should only lead to an increase in the number of events in the histogram, values of $\mu < 0$ would be unphysical as they would imply that the new resonance inhibits background events from being produced. The following expression,

which takes into account and corrects for the possibility that the value at maximum $\hat{\mu} < 0$ and treats cases in which a large upward fluctuation pushes $\hat{\mu} > \mu$ as fully compatible with the signal hypothesis, is taken as the test statistics

$$\tilde{q}_\mu = \begin{cases} -2\ln \frac{L(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(0,\hat{\hat{\boldsymbol{\theta}}}(0))} & \hat{\mu} < 0 \\ -2\ln \frac{L(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(\hat{\mu},\hat{\boldsymbol{\theta}})} & 0 < \hat{\mu} \leq \mu. \\ 0 & \hat{\mu} > \mu \end{cases} \tag{3.20}$$

Under certain assumptions of regularity on the nuisance parameters and on the total size N of the data sample (which has to be large enough so that $1/\sqrt{N}$ is small compared to the distance in σ of a Normal distribution for which the one-sided tails have an integral equivalent to $(1 - CL_s)$, where CL_s is the requested confidence level chosen for exclusion), $\tilde{q}_\mu$ assumes a simple form from which the significance of the exclusion can be calculated as a function of $\hat{\mu}$ and of its variance σ estimated from a synthetic dataset. This dataset is built from the Monte Carlo simulated events used to model signal and background. This procedure, called the *asymptotic method* makes no use of toy Monte Carlo pseudoexperiments, and is therefore very fast compared to previously available ones. This whole procedure is implemented in the RooStats package [112], for which a front-end has been developed by the CMS Higgs to Tau Tau search group.

3.11. Tag and Probe Method

The Tag and Probe (T&P) method allows to determine from data (or simulation) efficiencies and other physical quantities representing the fraction of events passing (or failing) a given selection. It is based on the determination of the total number of events and that of those passing the cuts. The method is used frequently in detector studies due to the quality of its results, as in good conditions the extracted value can be affected only by a very small bias compared to the "real" value being estimated. It is easier to explain it using a concrete example: the selection efficiency of a series of identification cuts for real electrons in experimental data.

Measuring the efficiency of a set of identification cuts for real electrons presents a challenge deriving from conflicting needs. Indeed, it is desirable to tighten the pre-selection of the electrons so that the measurement is performed on a very pure sample.

Unfortunately, tightening the electron pre-selection biases the set of electrons over which the efficiency is being estimated, to the point that if high purity is achieved the measurement will likely be excessively optimistic. This happens because genuine electrons which were difficult to identify had been filtered out by the tight pre-selection.

One possible solution to this dilemma is using as the dataset on which the measurement is performed pairs of electrons emerging from the decay of well-known massive particles, such as J/ψ and Z^0. The line shape of the distribution in the invariant mass of the ee pair is known from theory, and the line shape of the non-resonant background can be modeled by a properly chosen function.

In the case of the Z^0 resonance, the shape of its peak is modeled by a relativistic Breit-Wigner distribution. It is common to use for representing the background the

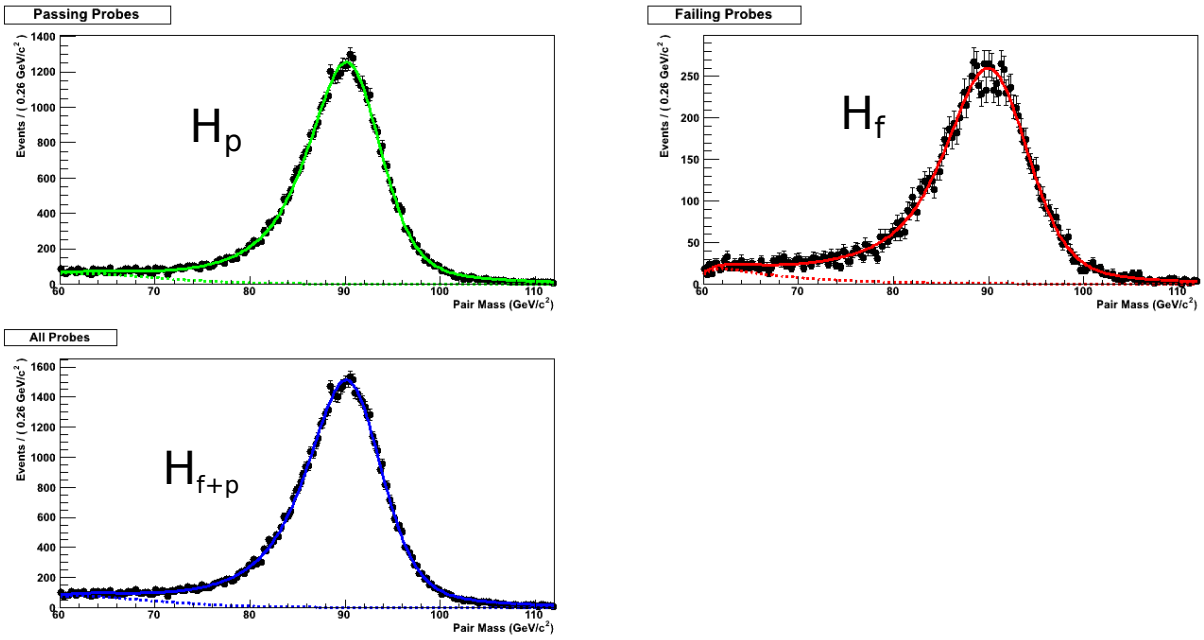


Figure 3.18.: Example from a T&P study on the efficiency of an electron ID MVA, using the spectrum in invariant mass of the electron pairs around the $Z^0 \to ee$ resonance region. Top left: pairs with a probe passing the cuts. Top right: are pairs with a failing probe. Bottom left: all the pairs.

product between the *erfc* function and an exponential. Both functions are convoluted with a function representing the response of the detector (a commonly used model is the Crystal Ball distribution [113]).

A tight selection is imposed on one (and only one) of the electrons in the pair, this electron is called the *tag*. This tight selection is chosen as to have an acceptable Z^0 signal to background ratio in the measurement dataset. The other electron, the *probe*, is subject to a set of loose selection and identification criteria that are designed to keep the bias to the measurement small.

As in figure 3.18, the invariant mass of the di-electron pair is drawn as a histogram for the entire dataset ("all pairs", where the histogram is indicated as H_{f+p}). Histograms are also drawn for the subsample in which the probe electron passes the selection criteria under study ("passing probes", histogram H_p) and for the subsample in which the probe electron fails the same criteria ("failing probes", histogram H_f).

Calling $f_s(m_{ee}|\theta_s)$, $f_b(m_{ee}|\theta_b)$ and $f_r(m_{ee}|\theta_r)$ the distributions (normalized to unity) for the Z^0 signal, the background and the detector response in the invariant mass m_{ee}, respectively, the symbols θ_s, θ_b and θ_r are chosen to represent their parameters.

A simultaneous fit is performed on the parameters of the distributions and their normalizations N, such that the fitted functions h_{s+b}, h_p and h_f shown in equation 3.21 describe the three histograms H_{s+b}, H_p and H_f.

$$
\begin{aligned}
h_{s+b}(m_{ee}) &= (N_{s,f} + N_{s,p}) \cdot f_s \otimes f_r(m_{ee}|\theta_s, \theta_r) \\
&\quad + (N_{b,f} + N_{b,p}) \cdot f_b \otimes f_r(m_{ee}|\theta_b, \theta_r) \\
h_p(m_{ee}) &= N_{s,p} \cdot f_s \otimes f_r(m_{ee}|\theta_s, \theta_r) + N_{b,p} \cdot f_b \otimes f_r(m_{ee}|\theta_b, \theta_r) \\
h_f(m_{ee}) &= N_{s,f} \cdot f_s \otimes f_r(m_{ee}|\theta_s, \theta_r), + N_{b,f} \cdot f_b \otimes f_r(m_{ee}|\theta_b, \theta_r)
\end{aligned}
$$

where the fitted normalization factors $N_{s,f}$ and $N_{s,p}$ represent the number of signal events failing and passing the selection under study, $N_{b,f}$ and $N_{b,p}$ are the equivalent numbers for the background. The estimated efficiency for real electrons of the cut under study is thus:

$$
\epsilon_{\text{cut}} = \frac{N_{s,p}}{N_{s,p} + N_{s,f}} \tag{3.21}
$$

An example from a study on the efficiency of the electron identification BDT discussed in section 3.5.1 in bins of η and p_T is shown in figure 3.18. In that figure, the blue line corresponds to the fitted distribution representing the spectrum in mass of the entire dataset, the green line is the equivalent for the passing probe datasets and the red one corresponds to the failing probe dataset. The dashed line represents the component attributed to background, and contributes in small part to the yield in all three samples, thanks to the selection imposed on the tag electron.

By selecting probe electrons in bins of η and p_T and repeating the fit for each bin, it is possible to perform a scan of the efficiency $\epsilon(\eta, p_T)$ as a function of these variables.

3.12. Data Samples

The data sample used in this analysis was collected by the CMS detector using two different dilepton triggers. The triggers require the presence of an electron and a muon, that satisfy a set of cuts. The first trigger is:

- `Mu8_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL`

  This trigger is composed of two "trigger legs". Legs are algorithms that take care of the identification of one type of lepton. Both trigger legs – the one for the muon and the one for the electron – need to have triggered for the event to be accepted. The trigger legs in use are:

  - `hltMu8Ele17CaloIdTCaloIsoVLTrkIdVLTrkIsoVLTrackIsoFilter`

    is the trigger leg associated to the electron. It requires that at least one electron candidate is present with a transverse momentum of 17 GeV/c^2. The electron satisfies cuts on calorimeter- and tracker-based isolation and on identification cuts, such as the number of hits in the pixel detector and the shape of the distribution of its energy deposits in the ECAL.

– `hltL1MuOpenEG12L3Filtered8`

is the trigger leg associated to the muon. It requires the muon to have triggered the (very fast) first level of the trigger. No selection on transverse momentum is performed at this stage. The muon is then required to have a $p_T > 8$ GeV/c^2 at the second and third levels of the trigger, which are implemented in the HLT (described in section 2.8).

The second trigger in use is named:

- `Mu17_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL`

  And requires following trigger legs of at least one lepton candidate each:

  – `hltMu17Ele8CaloIdTCaloIsoVLTrkIdVLTrkIsoVLTrackIsoFilter`

  is associated to the electron and similar to the one discussed above for the previous trigger. It requires the electron to satisfy the cut $p_T > 8$ GeV/c^2.

  – `hltL1Mu12EG7L3MuFiltered17`

  is the trigger leg associated to the muon. The muon is then required to have a $p_T > 17$ GeV/c^2 at the second and third levels of the trigger.

The trigger legs are associated to *trigger objects*, which are particle candidates built from the reconstruction algorithm in the HLT [114]. This algorithm trades high precision in the reconstruction for speed in the execution of the code. Trigger objects are, indeed, the objects on which the HLT algorithms runs.

It is important that the lepton on which the efficiency measurement is made is really the one responsible for the trigger. Selecting events that were not triggered by those leptons that are also selected in the analysis may lead to a biases. To guarantee that the leptons are responsible for the triggering of an event, the trigger objects associated to muon and electron trigger leg are matched, respectively, to the reconstructed muon or electron. This operation, which is referred to as *trigger matching* requires the trigger object to lie in a cone centered on the direction of flight of the lepton. The size of the cone in the $\eta - \phi$ plane in this analysis is equal to $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.5$.

Data recorded with the two triggers (as well as other $e\mu$ triggers) are saved into central datasets and are made available for analysis in the CMS collaboration. The datasets that are used in this analysis are the following:

`/MuEG/Run2012A-22Jan2013-v1/AOD`
`/MuEG/Run2012B-22Jan2013-v1/AOD`
`/MuEG/Run2012C-22Jan2013-v1/AOD`
`/MuEG/Run2012D-22Jan2013-v1/AOD`

These datasets correspond to the entire set of "good" events containing at least one muon and one electron collected by the CMS experiment in the year 2012 at the LHC. The events are defined as good if they were collected in experimental conditions where the LHC machine and all the detectors were in proper conditions to perform physics measurements. This certification is performed by the CMS Data Quality Monitoring and Certification group with the aid of dedicated software [115].

The total integrated luminosity for the above datasets is $\mathscr{L} = 19.7\ fb^{-1}$. The uncertainty on this value and its effect on measurement will be discussed later, in the paragraphs dedicated to the discussion of systematic uncertainties (section 5.4.1).

3.13. Simulated Samples

The simulation of a particle collision at high energy, is broken down into a series of steps, illustrated in figure 3.19.

1. Modeling of the initial state before the collision takes place. In the case of proton-proton collisions this involves the description (through *parton density functions*) of the quarks and gluons that compose the protons.

2. Modeling of the hard collision between partons constituting the protons.

3. Modeling of possible hard interactions between the remaining components of the colliding protons. This is referred to as the *underlying event* (UE).

4. Modeling of the evolution of strongly interacting elementary particles produced in the hard collision, in the process that leads to the formation of parton shower.

5. Modeling of additional collisions in the same, previous or following proton bunch crossings (*pile-up*).

6. Modeling of the interaction of the particles with the detector material, the effects of pile-up and the response of the detector to the interacting particles.

In order to generate predictions in the form of simulated events, one or more pieces of software simulate each of these steps. All the non-deterministic processes involved in the simulation, for which theory (or empirical experience) provides an expected distributions, are treated using Monte Carlo methods, from which the name "*Monte Carlo simulation*" originates.

An example of these processes is the propagation of an electron through a crystal of the ECAL, and one of the observables associated to the process is the amount of light generated by the electromagnetic shower, which the detector electronics and subsequent reconstruction convert to an estimate of the energy deposited in the crystal.

Hadronization and underlying event require a large theoretical effort to be modeled properly, as the behavior of the strong force at the characteristic energies of this processes is highly non-perturbative. The most computationally intensive task is the last one, the simulation of the detector response.

In CMS, the simulation of the interactions of particles with the detector and the response of the latter to these particles is implemented in the Geant4 software framework [117, 118]. All samples used in this analysis use Geant4 for detector simulation.

The simulation of parton showers, fragmentation and underlying event is managed by PYTHIA 6.4 (tune $Z2^*$) [119] for all simulated samples. The same package is used to generate the additional pile-up interactions. The treatment of parton density functions follows the recommendation of the PDF4LHC group [120].

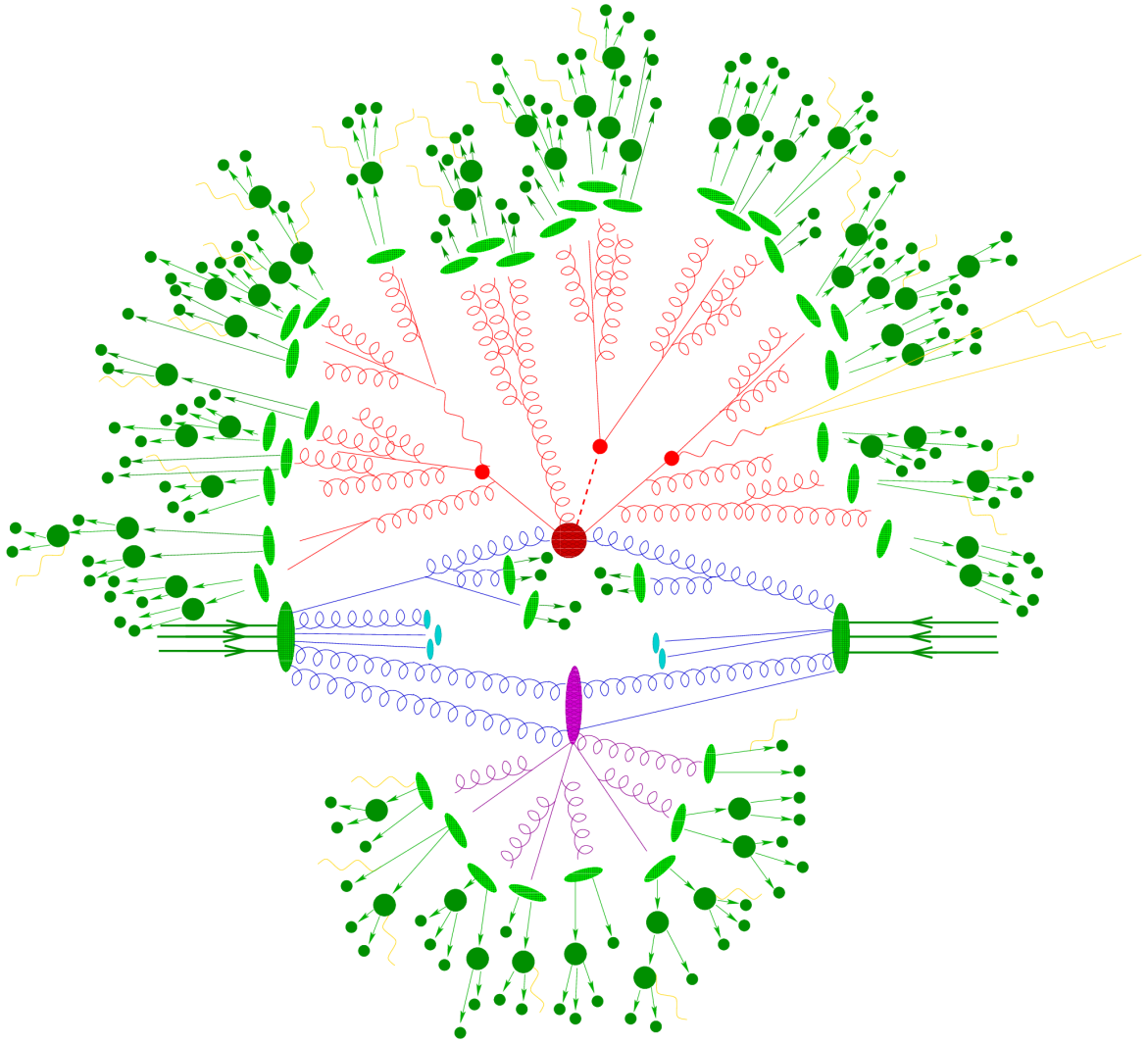


Figure 3.19.: Illustration of a generated Monte Carlo event: two protons approach from the sides (converging dark green arrows and blob), one constituent from each of them (a quark on the left, a gluon on the right) interacts in a hard collision (center of the figure and red blob). In parallel, other constituents of the protons interact with each other (UE: underlying event, violet blob), while others fly straight (beam remnants, cyan blobs). The strongly interacting particles produced in the hard collision (red lines) produce a shower of particles. The same happens to those produced in the UE, albeit in a smaller number due to the usually lower energies. Finally, all strongly interacting particles form to hadrons: some stable (small green circles), some unstable and decaying to other hadrons (large green circles). Leptons and photons (yellow) produce in the collision escape without interacting. From [116].

Process	Generator	σ_{eff} $[pb]$
$t\bar{t}$ (fully leptonic)	MADGRAPH	26.1975
$t\bar{t}$ (semi leptonic)	MADGRAPH	109.281
Single t	POWHEG	11.1
Single $\bar{t}$	POWHEG	11.1
$WW \to 2\ell\ 2\nu$	MADGRAPH	5.824
$WZ \to 3\ell\ \nu$	MADGRAPH	1.058
$WZ \to 2\ell\ 2q$	MADGRAPH	2.207
$ZZ \to 4\ell$	MADGRAPH	0.181
$ZZ \to 2\ell\ 2\nu$	MADGRAPH	0.716
$ZZ \to 2\ell\ 2q$	MADGRAPH	2.502
SM $gg \to H$ @ 125GeV/c^2	POWHEG	19.27
SM VBF $qq \to H$ @ 125GeV/c^2	POWHEG	1.643
SM $VH,\ ZH,\ t\bar{t}H$ @ 125GeV/c^2	PYTHIA	0.7046
SUSY $gg\phi$	PYTHIA	Function of $(m_A, \tan\beta)$
SUSY $bb\phi$	PYTHIA	Function of $(m_A, \tan\beta)$

Table 3.2.: List of Monte Carlo samples used in the analysis, the generator used to simulate the hard scattering and the effective cross section.

All samples involving the production τ leptons are processed by the TAUOLA generator [121] to simulate the decay. TAUOLA has been designed to account properly for the effect of spin correlations in τ decays.

Different generators are used to simulate the hard scattering process in the simulated samples: MADGRAPH 5.1 [122], POWHEG 1.0 [123] or PYTHIA 6.4 [119]. In table 3.2 the generator used to simulate the hard scattering for each sample is listed, along with the effective cross section for the process, i.e. the product $\sigma_{\text{eff}} = \sigma \cdot \epsilon_{\text{filter}}$, where σ is the calculated cross section for the process and ϵ_{filter} the overall acceptance for the filters applied to events after generation.

All the simulated samples are subject to the same selection criteria used for the selection of events in data samples. This provides consistency between data and simulation, and allows for monitoring the description of individual backgrounds in suitable control regions.

3.13.1. Correction applied to Top Pair Samples

A correction is applied to the $t\bar{t}$ sample, following measurements performed by the CMS collaboration on the differential cross section for the production of a top quark pair at $\sqrt{s} = 8$ TeV [124]. Events are reweighted as a function of the transverse momenta of the two top quarks p_{Tt} and $p_{T\bar{t}}$, where the expression for the weight $w_{t\bar{t}}$ is:

$$w_{t\bar{t}} = \exp\left(\frac{a}{2} + b \cdot \frac{p_{Tt}}{2}\right) \cdot \exp\left(\frac{a}{2} + b \cdot \frac{p_{T\bar{t}}}{2}\right) = \exp\left(a + b \cdot \frac{p_{Tt} + p_{T\bar{t}}}{2}\right) \tag{3.22}$$

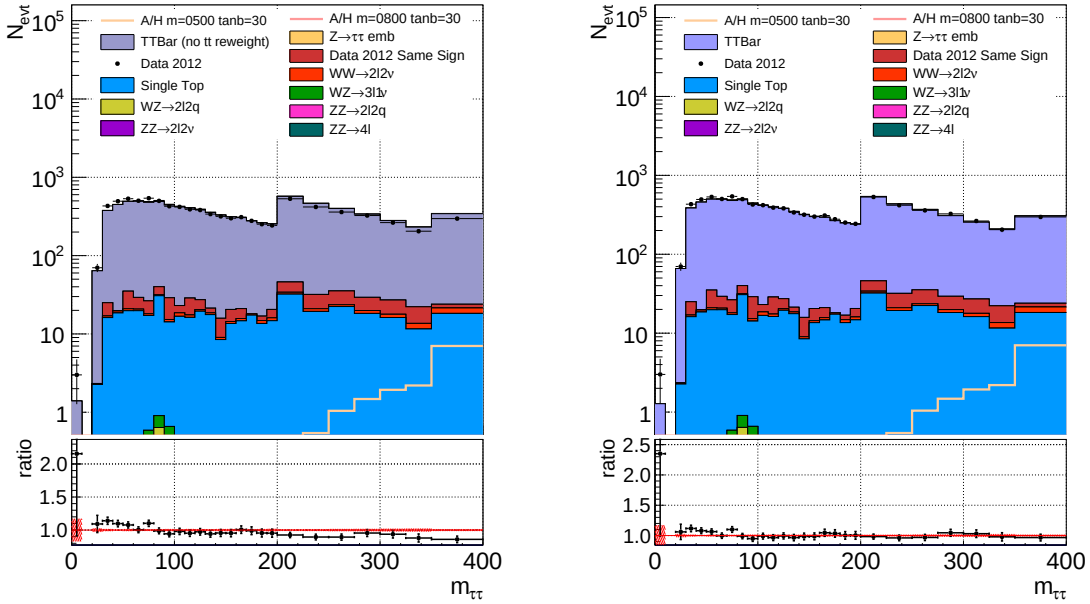


Figure 3.20.: Ditau invariant mass in the top quark pair-enriched control category. Left: before the application of top quark pair reweighting. Right: after the application off the reweighting.

where the coefficients have the values $a = 0.148$ and $b = -0.00129$ GeV^{-1}.

The correction improves the agreement between data and simulation, as it is shown in figure 3.20. The left histogram in the figure shows the top quark pair-enriched control category (defined in section 4.4) in case the weight is not applied, and a slope in the data/simulation ratio histogram is clearly visible. The right histogram shows the same distribution when the weight is applied, and in this case the slope in the ratio histogram is not noticeable.

3.14. Embedded Samples

Embedded samples are hybrid samples in which a part of the event originates from observed data and the other part is generated by a Monte Carlo simulation. They are used as an alternative to fully simulated Monte Carlo samples in situations in which there is a need to improve the predictions available in simulation.

For embedded samples to be viable, a production process with similar characteristics to the one that is being generated needs to exist. This process needs to be well known and should have a large cross section, so a large number of events can be harvested from data to be later modified by the embedding technique.

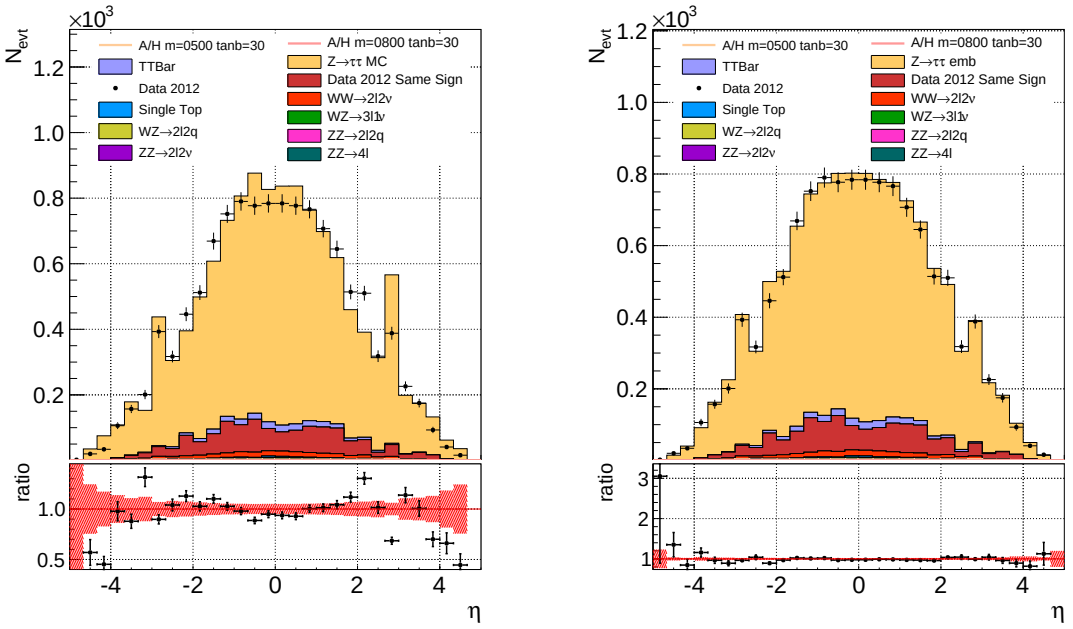


Figure 3.21.: Pseudorapidity of the leading jet in the $Z \to \tau\tau$ control category. Left: pure Monte Carlo. Right: embedded τ sample.

In the case of this analysis, embedded samples of the type "*track embedded*" are used to model $Z \to \tau\tau$ background. The procedure to generate one embedded event is illustrated in figure 3.22. The process can be broken down in the following steps:

1. Quality cuts on the leptons are imposed on experimental data to select a pure sample of events with two muons. To improve the purity sample of $Z \to \mu\mu$ events, a cut is imposed on the invariant mass of the $\mu\mu$ system: 60 GeV/c^2 < $m_{\mu\mu}$ < 120 GeV/c^2.

2. In each event of this sample the muons are removed, along with all tracks lying in a cone around them. The cone is chosen to have an angular radius $\Delta R = 0.5$.

3. τ leptons having the same 4-vector of the muons are generated in a synthetic event. This event is separate from the source one and contains only the τ lepton pair.

4. The decay of the τ leptons is simulated by the TAUOLA Monte Carlo code, and the resulting event is reconstructed normally.

5. The simulated τ sub-event is merged into the original event, with muons removed.

The embedding technique produces $Z \to \tau\tau$ events which better describe the behavior of jets in experimental data. This is visible, for example, by comparing the η distribution of the leading jet in the $Z \to \tau\tau$ enriched category (defined in section 4.4).

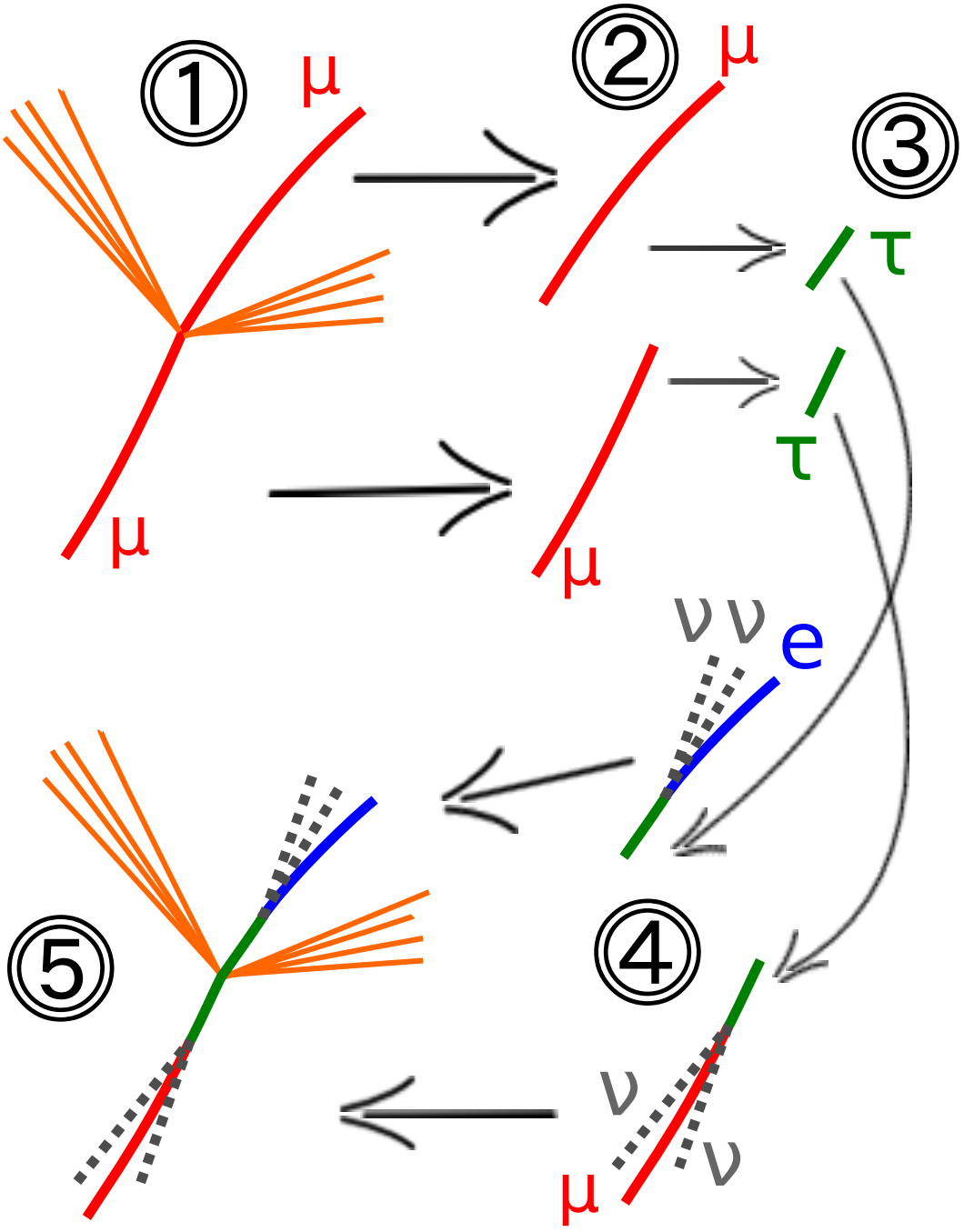


Figure 3.22.: Sketch of the tau embedding process.

Nevertheless, the purity requirements imposed on $Z \to \mu\mu$ events bias the kinematic distributions of the leptons. It is therefore important to account for this bias and correct it by reweighting the events in the embedded sample. The weight w_{emb} is expressed as a function of transverse momenta $p_T^{\mu+}, p_T^{\mu-}$ and the pseudorapidities $\eta^{\mu+}, \eta^{\mu-}$ of the original muons.

$$w_{\text{emb}} = SF_{\text{emb } p_T}(p_T^{\mu+}, p_T^{\mu-}) \cdot SF_{\text{emb } \eta}(\eta^{\mu+}, \eta^{\mu-}) \tag{3.23}$$

where the scale factors SF are extracted by a data-driven Tag and Probe fit. The

scale factors are equal to the inverse of the measured efficiency for $Z \to \mu\mu$ events when the muons lie in the specified $(p_T^{\mu+}, p_T^{\mu-}, \eta^{\mu+}, \eta^{\mu-})$ bin.

3.15. Estimation of QCD and W+Jets Background

The event selection that will be presented in the following chapters is effective at suppressing most of the W+Jets and QCD background. Nevertheless, a small contribution remains and needs to be properly modeled. This is done in a data-driven fashion, using a sample which is independent from the one used to perform the statistical analysis. The sample is referred to as "*W+Jets/QCD Fakes*", as it includes also W+Jets events where the jet is misidentified as an electron.

Events are selected from the 2012 data sample, using the same selection imposed on events in the inclusive category (these cuts will be discussed in detail in section 4.1), with one exception: the lepton pair needs to be composed of leptons with the same charge. This sample is referred to as *same sign* (SS) sample, in contrast to the one selected by imposing opposing charge of the leptons, the *opposite sign* (OS) sample. The requirement that the leptons have the same charge forces, by construction, the same sign sample to be independent from the one used for the signal categories. The spectrum for the same sign sample is shown in figure 3.23.

There is no guarantee that the SS sample has the same yield as the OS one, therefore two alternative corrections for the possible variations in the normalization of the spectrum are implemented. Both the corrections resort on the so-called *ABCD* method. This method uses a pair of weakly correlated variables to estimate the yield of events in one of four regions, defined by imposing rectangular cuts on the two variables. It is illustrated in figure 3.24. In our case, the variables chosen are the sign of the charges in the pair and the isolation of one or both leptons in the pair.

- The A region corresponds to the one we want to estimate, that is, the W+Jets/QCD spectrum in the OS sample with isolated leptons. The isolation cut (it will be described in detail in section 4.1) is $I_{\text{rel}} < 0.15$ in the barrel and $I_{\text{rel}} < 0.10$ in the endcaps.
- The B region corresponds to the sample obtained from data with a SS pair selection, described in the previous paragraph.
- The C region contains events with OS pairs and with the isolation cut on the electron, muon or both being inverted (there are three definitions for this region, as they are used to get a better estimate with the method explained below). The cut on isolation is $I_{\text{rel}} > 0.40$ in the barrel and $I_{\text{rel}} > 0.50$ in the endcaps.
- The D region contains events with SS pairs and with the isolation cut same as in C.

The first correction uses the overall normalization of the ABCD regions. N_A is the number of events in the A region and, similarly, N_B, N_C and N_D are the number of events for the B, C and D regions. The estimation for the number of the events in the A region is:

$$N_A = N_B \cdot \frac{N_C}{N_D}. \tag{3.24}$$

A diagram showing the four regions in the case of inverted isolation for both the leptons in the barrel region is shown in figure 3.23. By calling $f_B(m)$ the spectrum of events in invariant mass – of the $\tau\tau$ system or of the $e\mu$ system – obtained by selecting isolated SS pairs, the estimation for the QCD/W+Jets background in the signal (A) region is:

$$f_A(m) = f_B(m) \cdot \frac{N_C}{N_D}. \tag{3.25}$$

To get a more robust estimation of the ratio between N_C and N_D, the latter is taken as the median of the three different ratios that can be obtained by inverting the isolation of the electron only, the muon only, or both leptons.

$$\frac{N_C}{N_D} = \text{median}\left(\left(\frac{N_C}{N_D}\right)_{\text{inv } e}, \left(\frac{N_C}{N_D}\right)_{\text{inv } \mu}, \left(\frac{N_C}{N_D}\right)_{\text{inv both}}\right) \tag{3.26}$$

The second correction method accounts for the possibility that the ratio may change between the lower energy part of the invariant mass spectrum and the high energy part. The estimation of the ratio $\left(\frac{N_C}{N_D}\right)$ is changed from a global one to an estimation dependent on the invariant mass of the electron-muon pair. The ratio is estimated in coarse bins of $m_{e\mu}$, so the estimation for the W+Jets/QCD spectrum becomes:

$$f_A(m) = f_B(m) \cdot \frac{N_C}{N_D}(m_{e\mu}). \tag{3.27}$$

Again, the ratio is the median of the three values obtained by inverting the isolation of electron, muon or both leptons:

$$\frac{N_C}{N_D}(m_{e\mu}) = \text{median}\left(\left(\frac{N_C}{N_D}(m_{e\mu})\right)_{\text{inv } e}, \left(\frac{N_C}{N_D}(m_{e\mu})\right)_{\text{inv } \mu}, \left(\frac{N_C}{N_D}(m_{e\mu})\right)_{\text{inv both}}\right) \tag{3.28}$$

The shape for the W+Jets/QCD spectrum used in the analysis is the one obtained with the second method. The spectrum obtained with the first method is used in the estimation of the systematic uncertainties assigned to this sample, which will be discussed in section 5.4.1.

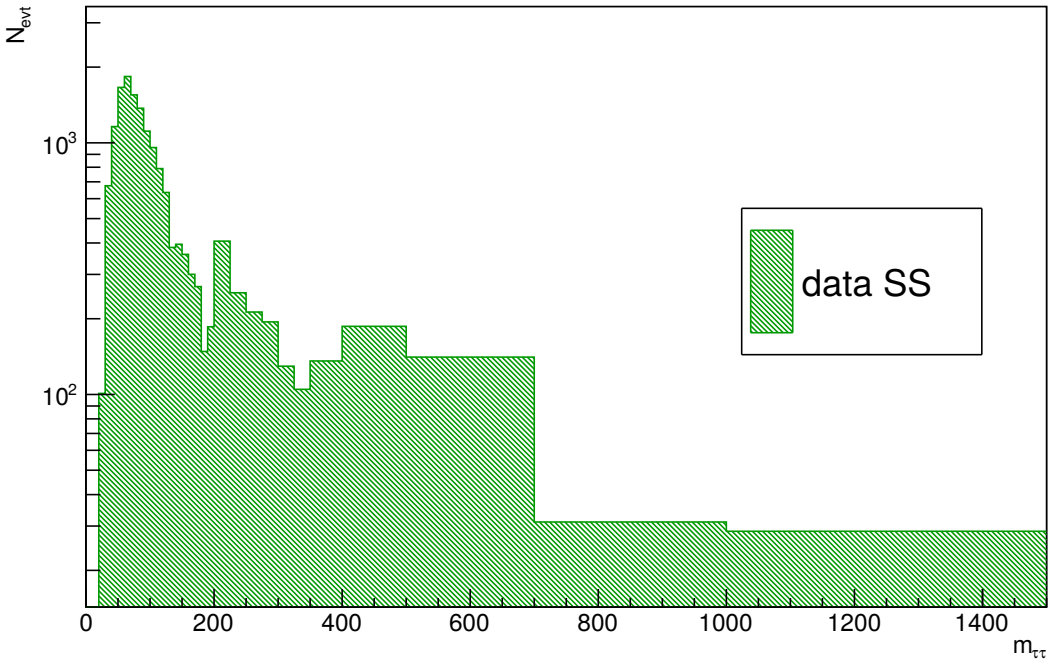


Figure 3.23.: $m_{\tau\tau}$ for the QCD/W+Jets fakes sample.

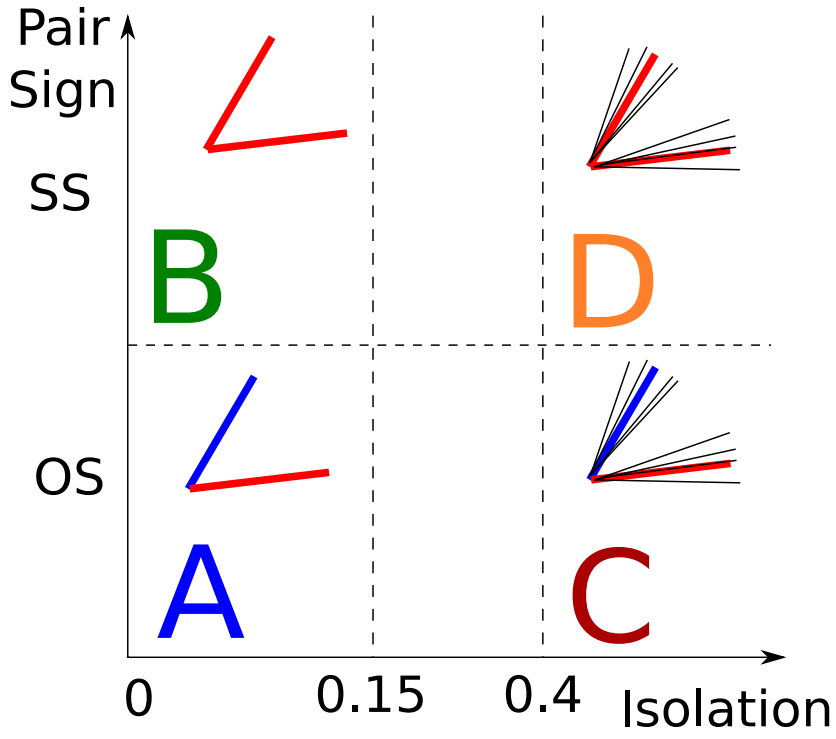


Figure 3.24.: The ABCD method, using lepton isolation and pair sign as variables.

4. Event selection, corrections and results

The signature of the MSSM Higgs search channel introduced in chapter 1 consists of an electron candidate, a muon candidate, missing transverse energy as a result of the neutrinos not being detected and generally no or few jets, which can be b-tagged in the case of Higgs produced in the b-quark associated production channel.

In this chapter the event selection is described. Its goal is to reduce the contribution from common Standard Model backgrounds events while keeping the efficiency for signal events high. Subsequently, the event categories used to control background and signal events are introduced. Finally, the results of the statistical analysis on these signal categories is presented.

Improvements with respect to the event selection presented here and to the statistical analysis procedure will be shown in the next chapter, chapter 5.

4.1. Brief Overview

The cuts applied on leptons and jets are summarized in table 4.6, and those applied on events are summarized in table 4.2. They will be described in more detail in the following sections.

Requirements for the event to be accepted in the inclusive category
Only one electron (additional electrons are counted with looser cuts)
Only one muon (additional muons are counted with looser cuts)
At least one of the two leptons with $p_T > 20$ GeV/c
The other lepton needs to have $p_T > 10$ GeV/c
Mass of the $e\mu$ pair $m_{e\mu} > 25.0$ GeV/c^2

Table 4.2.: Summary of the cuts applied to the events in the inclusive category.

Events passing these cuts are referred to as "inclusive category". This category is important as it is used, beside other monitoring functions, to fix the normalization of the $Z \to \tau\tau$ embedded sample to the one predicted by the $Z \to \tau\tau$ Monte Carlo simulation.

Cuts Applied on leptons		
Variable	Electrons	Muons
Pseudorapidity	$\|\eta\| < 2.3$	$\|\eta\| < 2.1$
Identification	No missing hits in pixel detector PF electron conversion veto Electron ID BDT (see below)	Tight Muon (section 3.6)
Trigger matched	yes	
Impact parameter w.r.t. PV	$\|d_0\| < 200$ µm and $\|d_z\| < 2$ mm	
Rel. Isolation (section 3.7)	$I_{\text{rel}} < 0.15$ ($\|\eta\| < 1.479$) and $I_{\text{rel}} < 0.10$ ($\|\eta\| \geq 1.479$)	

Electron Identification BDT cuts		
Momentum	Pseudorapidity	Cut
$p_T < 20$ GeV/c	$\|\eta\| \leq 0.8$	BDT > 0.925
	$0.8 < \|\eta\| \leq 1.479$	BDT > 0.915
	$1.479 < \|\eta\|$	BDT > 0.965
$p_T > 20$ GeV/c	$\|\eta\| \leq 0.8$	BDT > 0.905
	$0.8 < \|\eta\| \leq 1.479$	BDT > 0.955
	$1.479 < \|\eta\|$	BDT > 0.975

Cuts applied on jets	
Variable	Cuts
Pseudorapidity	$\|\eta\| < 4.7$
Transverse momentum	$p_T > 10$ GeV
Jet ID	Pass loose working point
Pile-up Jet ID	Pass loose working point
Separation from leptons	$\Delta R > 0.4$

Table 4.6.: Summary of the cuts applied to the physics objects used in the analysis.

4.2. Quality Cuts and Corrections on Physics Objects

Following the reconstruction described in chapter 3, the various physics objects such as GSF electrons, jets and muons have been selected with a set of loose cuts which are intrinsic to the reconstruction process. For the specific analysis, a set of tighter cuts that presents a good compromise between purity in the sample and signal event efficiency is chosen. The cuts selected for the present analysis are described in the following paragraphs.

In addition to the selection cuts, some correction factors are applied to simulated events. These corrections are needed to improve the description of the Monte Carlo samples when there are disagreements with data.

4.2.1. Pile-up

The distribution in the number of pile-up vertices in the Monte Carlo simulation is not consistent with the one observed in data. The number of pile-up vertices is underestimated on average. This is a result of the excellent performance of the LHC machine in the run of 2012, in which it delivered more luminosity than expected during the preparation of the data taking.

The simulation is brought to agreement with data through a reweighting: histograms representing the number of events with a given number N_{PU} of pile-up vertices are compared between data and Monte Carlo. The weight is expressed as

$$w_{\text{eff}}(N_{PU}) = \frac{n_{\text{events data}}(N_{PU})}{n_{\text{events MC}}(N_{PU})} \tag{4.1}$$

where $n_{\text{events data}}(N_{PU})$ and $n_{\text{events MC}}(N_{PU})$ are the populations of the bin N_{PU}, respectively, in the data and Monte Carlo histograms. This weight is applied to all MC samples.

4.2.2. Electrons

Electrons should lie in the geometrical acceptance region of the silicon tracker ($|\eta| < 2.5$) and in the acceptance of the electron trigger ($|\eta| < 2.3$), since we require them to be matched to electron trigger objects in order to be accepted. Therefore, all selected electrons lie in the pseudorapidity region $|\eta_e| < 2.3$.

Real electrons, not directly produced in the physical event of interest, can emerge from the decay-in-flight of hadrons. One example is the decay of kaons, where the process $K^+ \to \pi^0 e^+ \nu_e$ occurs $5.07 \pm 0.04\%$ of the times [11]. Another example is the conversion of photons into electron-positron pairs in the detector material. These electrons tend to originate far from the primary vertex and, thus, can be rejected by imposing cuts on the minimum distance of their track from this vertex. As an example, it is possible to see in figure 4.2 that electrons originating from $\phi \to \tau\tau \to e\mu$ decay have on average smaller transverse impact parameters, compared to those reconstructed from a QCD sample, which includes electrons from decays in flight of hadrons. In this analysis, electrons are required to have an impact parameter in the transverse

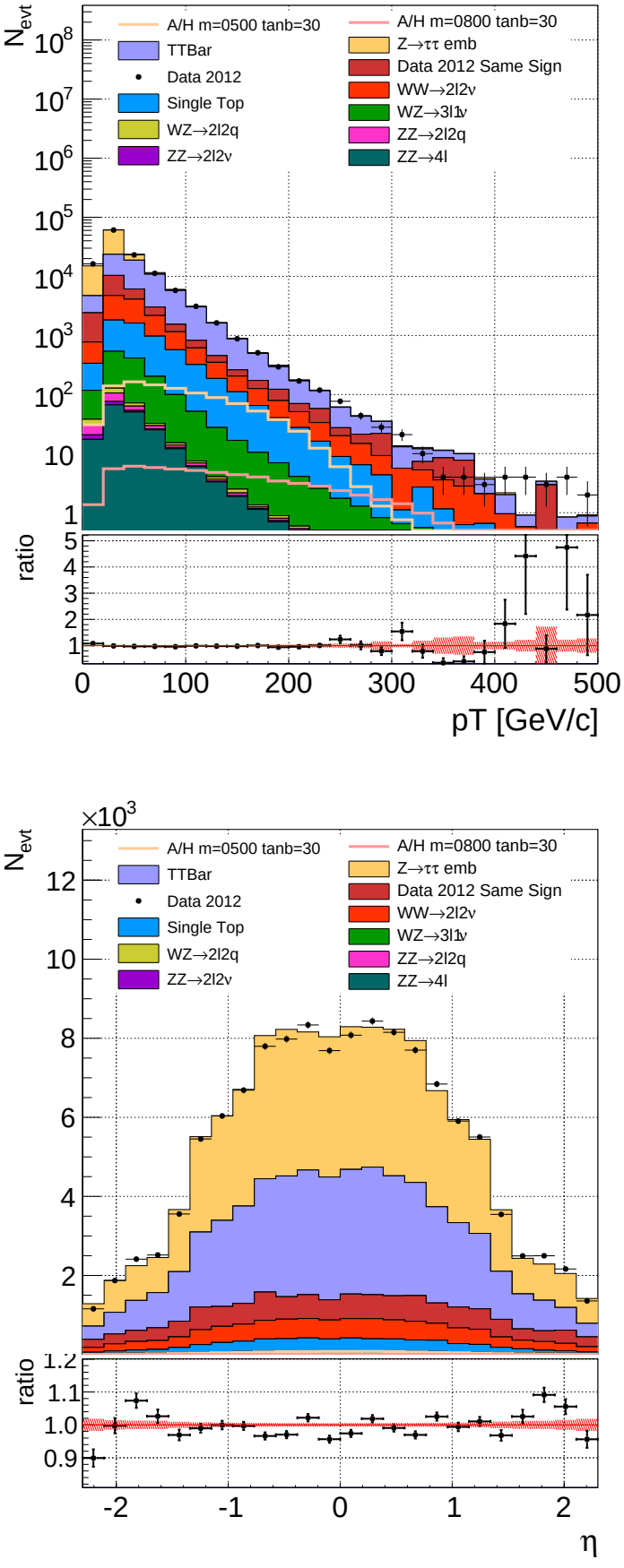


Figure 4.1.: Spectrum in p_T and η distribution of electrons in the inclusive category.

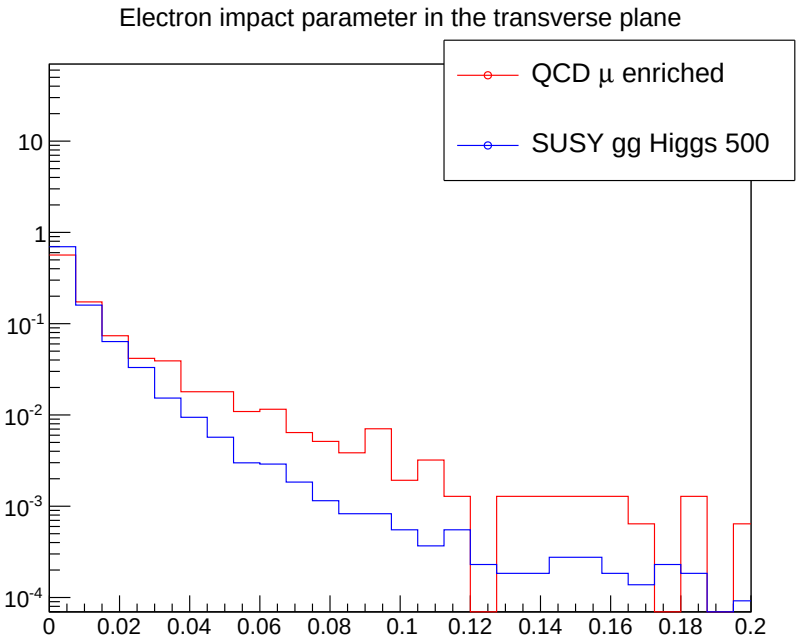


Figure 4.2.: Electron impact parameters in cm along the beam axis, comparing electrons from a sample of Higgs decays to misidentified electrons and electrons from decays in flight from a QCD sample.

plane $|d_0| < 200$ μm and along the beam axis $|d_z| < 2$ mm.

Noise in the tracker electronics and the occurrence of nuclear interactions in the detector material can produce randomly distributed hits that are not associated to charged particles originating in collisions. In turn, random hits can lead to the reconstruction of trajectories that do not correspond to any real particle. These artifacts are referred to as *fake tracks*. The CMS tracker has been designed to reduce the rate at which random hits occur, so it is rare that fake tracks, when erroneously reconstructed, are associated to a large number of hits. To reduce their fraction and that of photon conversions, in this analysis every electron track is required to have an associated hit in each of the traversed layers of the pixel detector.

As discussed in section 3.5.1, a common source of fake electron candidates are those particles which are produced in association with QCD activity. The electron identification tool described in section 3.5.1 is employed for this analysis at its medium working point, which provides an efficiency of around 85% (barrel) and 75% (endcaps) for real electrons and an efficiency of around 3−4% for particles being misidentified as electrons.

The efficiency for electrons to fire the trigger and pass the identification cuts may not be modeled properly by the Monte Carlo simulation. For this reason a weight $w_{\text{eff } e}$ is applied to simulated events as a function of electron η and p_T, in order to bring the simulation to agreement with data. $w_{\text{eff } e}$ is expressed as

$$w_{\text{eff } e} = \frac{\epsilon_{\text{e trig data}}(p_T, \eta)}{\epsilon_{\text{e trig MC}}(p_T, \eta)} \cdot \frac{\epsilon_{\text{e id data}}(p_T, \eta)}{\epsilon_{\text{e id MC}}(p_T, \eta)} \tag{4.2}$$

where the efficiencies ϵ as a function of p_T and η are estimated with the *Tag and Probe* method, described in section 3.11.

The spectrum in p_T and the distribution in η of the electrons in the inclusive category – which is defined in section 4.1 and includes events passing all the quality cuts – are shown in figure 4.1. There is a good agreement between data and simulation up to $p_T = 350$ GeV/c. The step-like peak at low values of p_T arises from the trigger thresholds and lepton p_T cuts being asymmetric.

4.2.3. Muons

Muons are selected with a high level of purity and a good efficiency following the Tight Muon identification criteria, described in section 3.6. Selected muons should lie in the geometrical acceptance of the muon system ($|\eta| < 2.4$) and that of the muon trigger ($|\eta| < 2.1$), so all muons are required to satisfy $|\eta| < 2.1$. Muons are required to be matched to a muon trigger object.

Similarly to electrons, muons can originate from the decay in flight of other particles. To improve the rejection of these kinds of muons the cut on impact parameter is tightened with respect to the one in the Tight Muon selection and is the same as applied to electrons: muons are required to have an impact parameter in the transverse plane $|d_0| < 200$ µm and along the longitudinal axis $|d_z| < 2$ mm. Similarly to the case of electrons, a weight is applied to simulated events to account for the different trigger and identification efficiency – estimated via Tag and Probe – between data and simulation. The weight $w_{\text{eff } \mu}$ is a function of muon p_T and η:

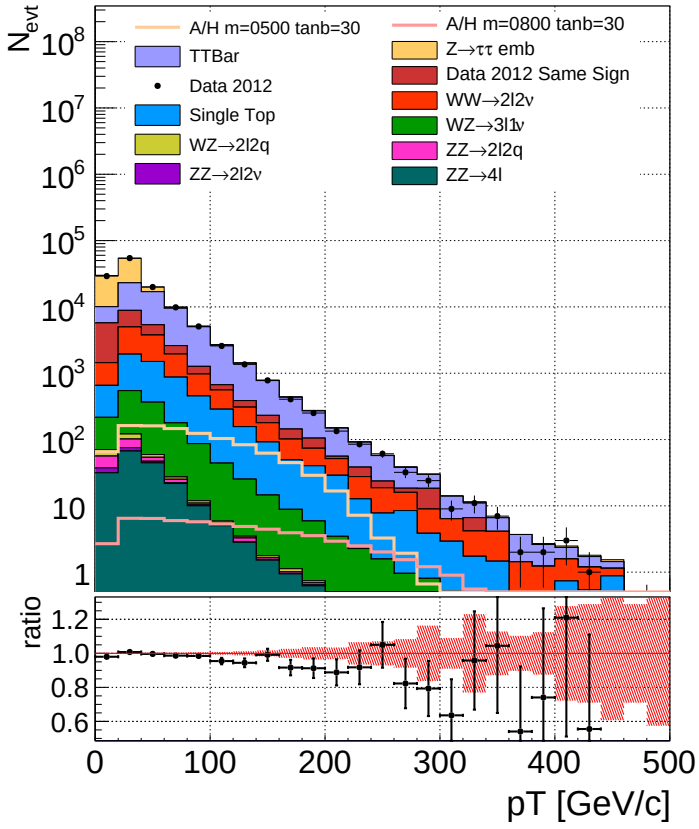
$$w_{\text{eff } \mu} = \frac{\epsilon_{\mu \text{ trig data}}(p_T, \eta)}{\epsilon_{\mu \text{ trig MC}}(p_T, \eta)} \cdot \frac{\epsilon_{\mu \text{ id data}}(p_T, \eta)}{\epsilon_{\mu \text{ id MC}}(p_T, \eta)} \tag{4.3}$$

The p_T spectrum and the η distribution of muons in the inclusive category are shown in figure 4.3. A good agreement between data and simulation is observed. The slight slope in the ratio plot can be attributed to the choice of the top quark p_T reweighting factor, which will be described later as a systematic uncertainty to the cross-section limit. The double peak at low values of p_T arises from the trigger thresholds being asymmetric.

4.2.4. Lepton Pairs

In addition to the cuts imposed to individual leptons, others are applied to the lepton pair.

Two electron+muon triggers are used to acquire the data used in this analysis, where one sets a higher p_T threshold for electrons, and the other sets a higher p_T threshold for muons. These triggers require an electron (or, respectively, muon) with at least a transverse momentum of 17 GeV and a muon (or, respectively, electron) with at least



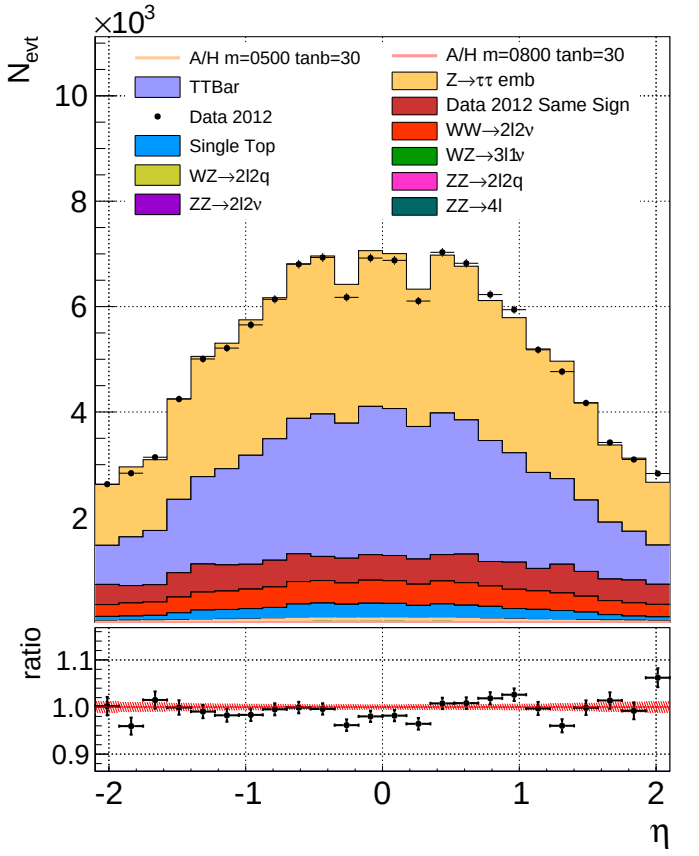


Figure 4.3.: Spectrum in p_T and η distribution of muons in the inclusive category.

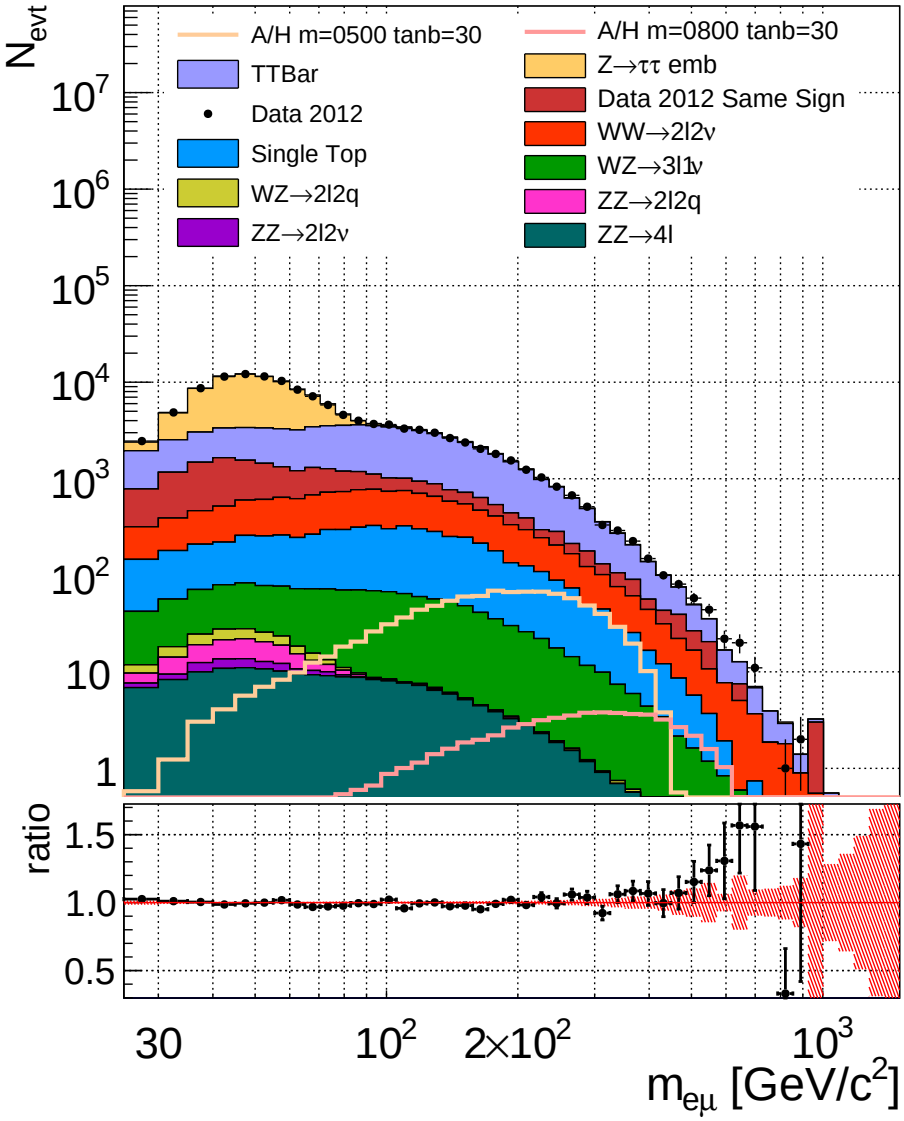


Figure 4.4.: Invariant mass of the $e\mu$ pairs in the inclusive category, in log-log scale.

a transverse momentum of 8 GeV. The trigger efficiency for each type of lepton in the region close to the p_T threshold varies significantly, from zero to an asymptotic value which depends on η, but is generally around 95% or more.

The trigger efficiency may not be modeled in Monte Carlo simulation exactly as it occurs in data. For this reason the already mentioned correction scale factors are applied to simulation to correct for this difference. In the p_T region close to the trigger threshold, though, the required correction factors may be very large and subject to significant systematic uncertainty. For this reason, a cut on the minimum p_T of the leptons is applied a few GeV above the threshold of the trigger, such that the region where the efficiency changes significantly is excluded. This cut does not have a critical impact on the analysis, as leptons originating from heavy Higgs boson decays have usually larger transverse momenta. The cut on p_T is chosen to be 20 GeV for the harder lepton and 10 GeV for the softer lepton in the pair.

The dilepton phase space region with invariant masses of the $e\mu$ system below $m_{e\mu} = 25.0$ GeV/c^2 has been observed to be affected by poor agreement between data and simulation. This may be explained by the fact that a variety of processes occurring at low energy have not been included in the simulation samples. For these reasons a cut is introduced on the $e\mu$ mass of the pair, which is required to satisfy $m_{e\mu} > 25.0$ GeV/c^2 for all events. The invariant mass between $m_{e\mu} = 25.0$ GeV/c^2 and 1500.0 GeV/c^2 is shown in figure 4.4. In that region a good agreement between simulation and data is observed.

4.2.5. Jets

All the accepted jets are required to lie in an acceptance region defined by $|\eta| < 4.7$.

The large amount of pile-up vertices in collisions during the 2012 run of the LHC can degrade the performance of the jet reconstruction algorithms (described in section 3.4). Indeed, not keeping under control the effects of pile-up has many negative effects, such as an incorrect estimation of the jet energy, the transverse momentum or the b-tagging score, and a wrong counting (usually an excess) in the number of jets produced in the collision, which increases with the number of pile-up interactions.

The number of jets is an important variable as it is used to discriminate between Higgs signal and $t\bar{t}$ background, and therefore it deserves a special mention. To better identify and reject those jets that are associated to pile-up interactions the Jet ID MVA and the PileUp Jet ID MVA (see section 3.4.3) are used, with a cut imposed at their loose working points.

The embedded τ technique used to model $Z \to \tau\tau$ samples (see section 3.14) may introduce biases if jets are too close to the original muons. To reduce the impact of the mentioned biases, the acceptable phase space for jets needs to be restricted: jets are required to have their momentum vector separated by at least $\Delta R = 0.4$ from the momentum of the leptons, and they are ignored if they don't satisfy the condition. Electrons and muons emerging from $\phi \to \tau\tau$ decay are expected to be isolated, so the impact on genuine signal events is expected to be small.

Following the cuts mentioned above, the p_T spectrum of all the jets passing the identification cuts in the inclusive category is shown in figure 4.5, where it is possible to see that there is a good agreement between data and simulation. In the same figure

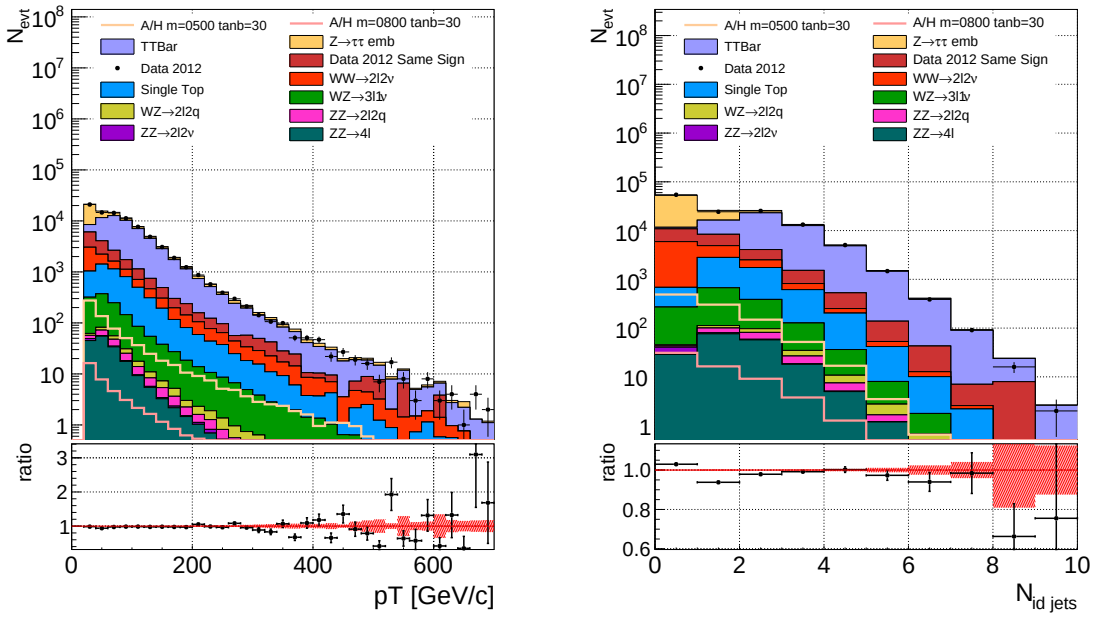


Figure 4.5.: Left: spectrum in p_T of the leading jet in the inclusive category. Right: number of identified jets with a $p_T > 30$ GeV/c in the inclusive category.

the number of identified jets with a $p_T > 30$ GeV/c is shown in the same category, and a slight disagreement is visible. This disagreement is taken into account in the limit calculation, where the systematic uncertainty on the jet scale is implemented as a nuisance parameter.

4.2.6. B-tagging

Jets which pass the selection criteria described in section 4.2.5 are analyzed by the b-tagging algorithms described in section 3.4.4. In the case of data events, a jet is defined to be b-tagged if it passes the medium working point of the Combined Secondary Vertex discriminator, i.e. if the score assigned to it by the algorithm is above a value of 0.679.

It has been noticed that the Monte Carlo simulation of the b-tagging algorithms does not fully reflect the one observed in data, therefore a statistical correction is applied to bring the Monte Carlo b-tagging values closer to the ones in data. The correction is implemented with the help of a random selection algorithm. It consists of re-declaring a small fraction of jets below threshold as b-tagged jets and - in a similar fashion - randomly declaring jets above threshold as non b-tagged ones, the ratio between the frequency of promotion into the b-tag class and demotion out from the b-tag class is chosen to obtain a match between Monte Carlo and data.

4.2.7. Missing Transverse Energy

The missing transverse energy is computed using the MVA MET method, described in section 3.8. No cuts are applied for the selection of events based on its value as it

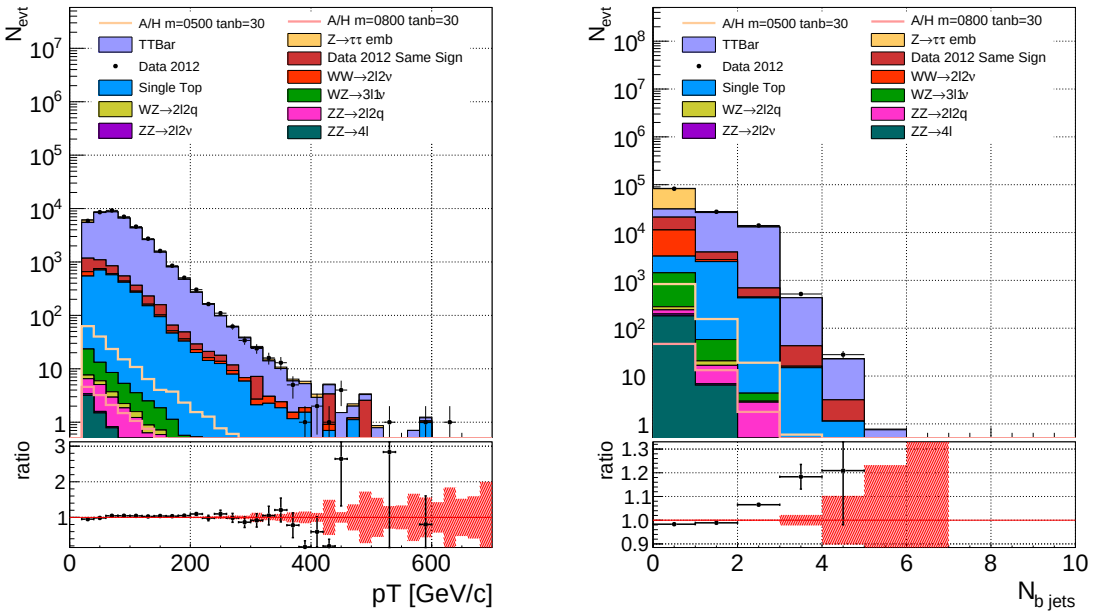


Figure 4.6.: Spectrum in p_T of leading b jet and number of b-jets with a $p_T > 20\text{GeV}/c^2$ in the inclusive category.

is preferred to use discriminators such as D_ζ, which will be later described in section 4.3.2.

The MET distribution in the inclusive category is shown in figure 4.7: while simulation describes well the MET distribution in data for low values of MET, a slight discrepancy is visible in the figure at MET values above 120 GeV. This will have an impact on the analysis as discussed in chapter 5.

4.3. Suppression of Standard Model Backgrounds

There are various processes described by the Standard Model that can produce real $e\mu$ pairs. One is the Drell-Yan process $pp \to Z/\gamma^* \to \tau\tau \to e\mu$, which is the most important at $\tau\tau$ invariant masses of the order of the Z rest mass (91.1876 ± 0.0021 GeV/c^2 [11]) and having a very large cross section at the LHC.

Another process with $e\mu$ pairs in the final state is the production of a pair of top quarks in cases where the top quarks decay leptonically to a e and a μ. This is the dominant process in the inclusive category at high mass (beyond $m_{\tau\tau} \simeq 200$ GeV/c^2).

An event can also present an $e\mu$ signature as a result of the production and decay of a pair of weak gauge bosons such as in the $pp \to WW \to e\nu\mu\nu$ process, which is an important source of background in the high mass region along with $t\bar{t}$.

In some cases more than one $e\mu$ pair in the event can be identified, one example being the decay chains $ZZ \to ee\mu\mu$, $WZ \to e\nu\mu\mu$ and $WZ \to \mu\nu ee$. The pair can also be emulated by the decay $Z \to \mu\mu +$ jets where one of the jets is misidentified

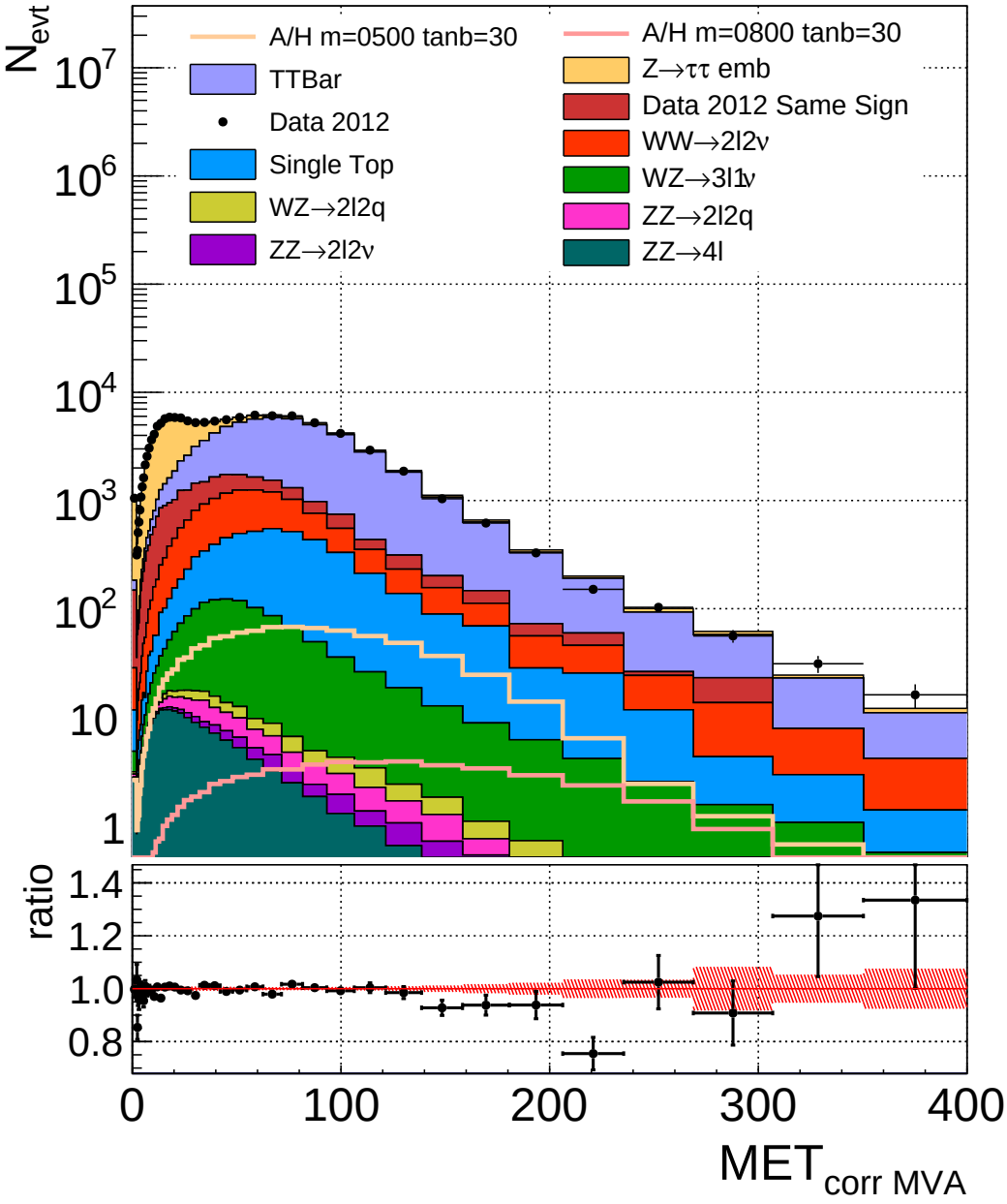


Figure 4.7.: Missing transverse energy distribution in the inclusive category, as computed by the MVA MET method.

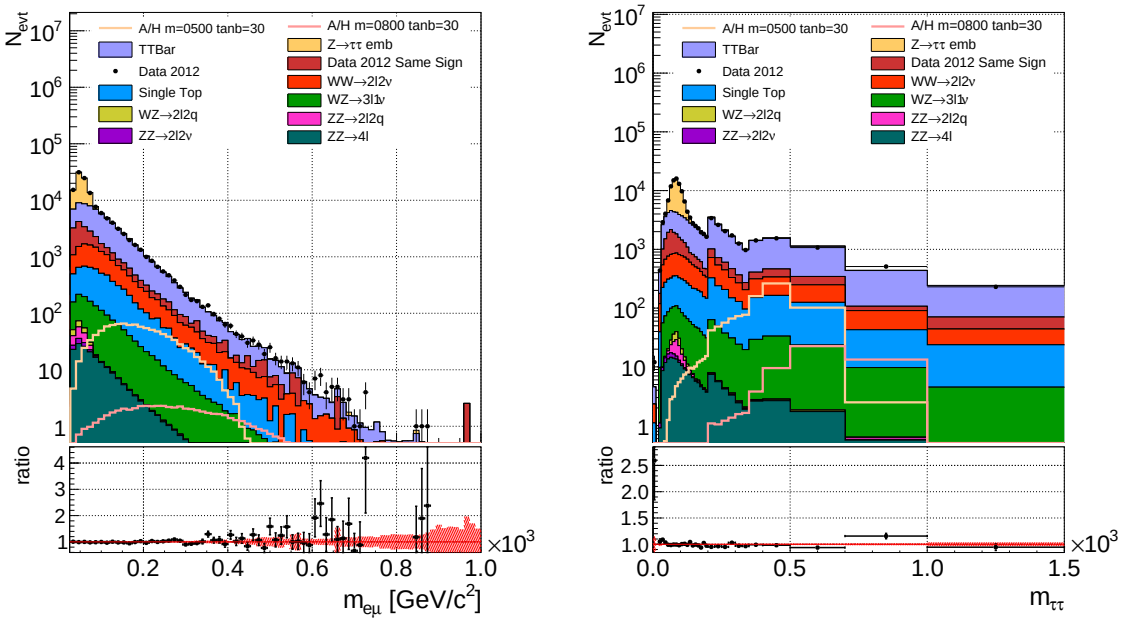


Figure 4.8.: Invariant mass spectrum of the $e\mu$ pairs and $\tau\tau$ pairs as reconstructed by the SVFit algorithm in the inclusive category.

as an electron. In the case of these decays it is very likely that more than two lepton candidates lie in the acceptance of the detector and are identified, so a good reduction of these backgrounds is achieved by rejecting events where one or more additional leptons pass a looser version of the identification cuts used to select the main pair.

Other processes that produce potentially misidentified $e\mu$ pairs are the $W+\text{jets} \rightarrow \mu\nu + e_{\text{fake}}$ process and various types of QCD processes, in which a jet is misidentified as an electron. The cross section in these cases is very large, but the quality cuts described in the previous sections suppress this background to the point that the contribution in the signal categories is small. The $W+\text{jets}/\text{QCD}$ background is modeled by the fakes sample, described in section 3.15.

4.3.1. Irreducible Backgrounds

Following the cuts mentioned earlier, the main sources of background are processes that produce single $e\mu$ pairs, where both leptons lie in the detector acceptance and are properly identified. These background events cannot be filtered by tightening further the quality cuts on the leptons and other physics objects and are therefore referred to as irreducible backgrounds.

The biggest contributions to the latter backgrounds come from Drell-Yan events in the low dilepton mass region and from $t\bar{t}$, single top quark and $WW \rightarrow \nu\bar{\nu}e\mu$ events in the high dilepton mass region, as is evident in the spectrum of the invariant mass of the $e\mu$ pairs and that of the $\tau\tau$ pairs in the inclusive category (Figure 4.8).

A number of event properties can be used to further enrich the sample with signal

events against backgrounds. For example, the number of hard jets, and especially the number of hard b-tagged jets, is different on average for the Higgs boson signal and the backgrounds. Therefore, the jet and b-jet multiplicities are used for event categorization, and will be further described in section 4.4.

Furthermore, the weak decay of a τ into lighter leptons occurs at a much smaller energy than the weak decay of a top quark or an electroweak boson into the same leptons. This property can be used to separate genuine $\tau\tau$ events. The difference is exploited by performing a cut on a suitably constructed variable, which is described in section 4.3.2.

Finally, the τ lepton has a mean lifetime which is long enough for it to have traveled a macroscopic amount of space at the time of decay. This distance between the primary vertex and the tau decays reveals itself as a displacement between the tracks of the electron and muon and the primary vertex. This variable is used in the selection described in chapter 5, which has been developed to optimize the search for a high mass Higgs boson.

4.3.2. Enrichment of $\tau\tau$ Events based on the Kinematics of their Weak Decay

A discriminant to distinguish between genuine τ decays and other weak decays can be built from the momenta of the leptons and the missing transverse energy. Such discriminant is referred to as D_ζ. The ζ refers to the bisector between the directions of flight of the leptons.

Let $\hat{p_e} = \overrightarrow{p_{T_e}}/|\overrightarrow{p_{T_e}}|$ and $\hat{p_\mu} = \overrightarrow{p_{T_\mu}}/|\overrightarrow{p_{T_\mu}}|$ be the directions of the electron and muon momenta in the transverse plane of the laboratory frame. The bisector between the direction of the electron and that of the muon can be written as

$$\hat{\zeta} = \frac{\hat{p_e} + \hat{p_\mu}}{|\hat{p_e} + \hat{p_\mu}|} \tag{4.4}$$

The modulus of the projection of the dilepton momentum in the transverse plane onto the $\hat{\zeta}$ axis is named $p_{\zeta\mathbf{vis}}$ and can be written as

$$p_{\zeta\mathbf{vis}} = \left(\overrightarrow{p_{T_e}} + \overrightarrow{p_{T_\mu}}\right) \cdot \hat{\zeta} \tag{4.5}$$

In a similar fashion, the projection of the missing transverse momentum on the $\hat{\zeta}$ axis is referred to as p_ζ and written as

$$p_\zeta = \vec{\not{E}}_T \cdot \hat{\zeta} \tag{4.6}$$

The discriminant D_ζ is then defined as follows

$$D_\zeta = p_\zeta - \alpha\ p_{\zeta\mathbf{vis}} = \vec{\not{E}}_T \cdot \hat{\zeta} - \alpha\ \left(\overrightarrow{p_{T_e}} + \overrightarrow{p_{T_\mu}}\right) \cdot \hat{\zeta} \tag{4.7}$$

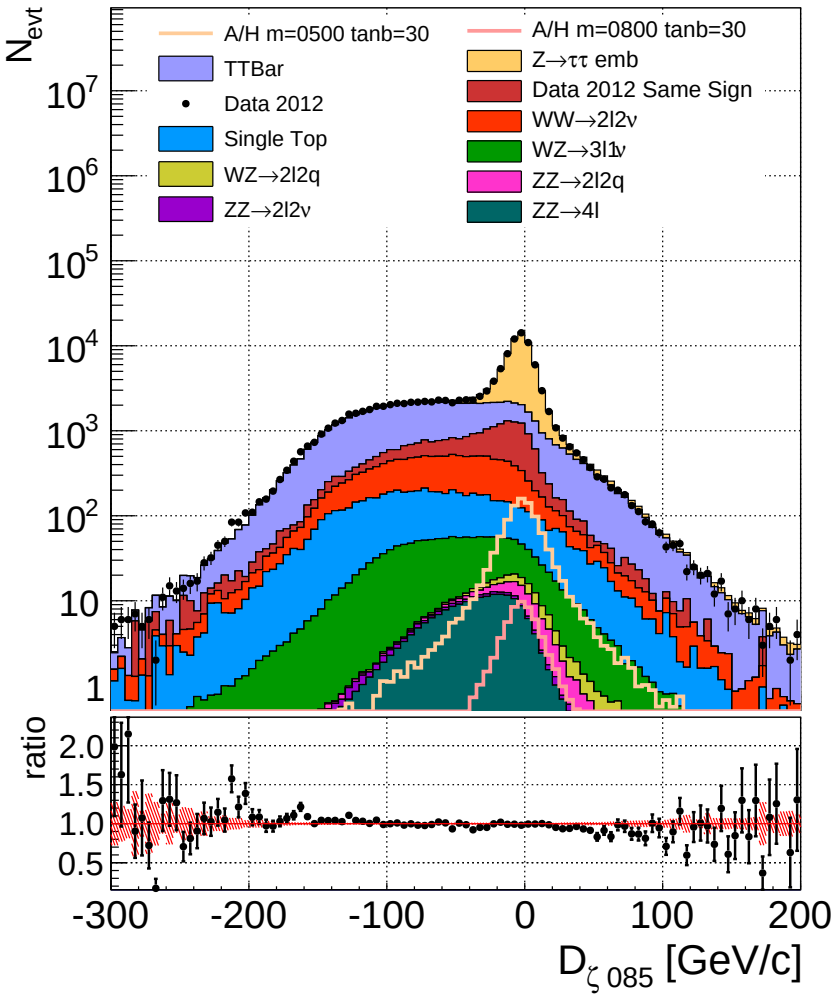


Figure 4.9.: Comparison between data and Monte Carlo prediction for the D_ζ variable in the inclusive category.

where the value of $\alpha = 0.85$ has been optimized in the past by dedicated studies [125]. Figure 4.10 gives a graphical representation of the aforementioned definitions in the case of a $\tau\tau$ decay and in the case of a WW or $t\bar{t}$ decay.

As one can see in figure 4.9, $\phi \to \tau\tau$ and $Z \to \tau\tau$ are distributed around zero while top quark pair production and electroweak processes (i.e. di-boson and single top) spread to larger negative values. Requiring the events to have a $D_\zeta > -20$ GeV/c reduces considerably the $t\bar{t}$ and electroweak background.

Explanation of D_ζ effectiveness from τ decay kinematics

The energy available to the decay products from the leptonic (i.e. weak) decay of a τ lepton is relatively small

$$Q(\tau \to \nu_\tau e \bar{\nu_e}) = (m_\tau - m_{\nu_\tau} - m_e - m_{\bar{\nu_e}})\, c^2 \simeq 1.78 \text{ GeV}$$

$$Q(\tau \to \nu_\tau \mu \bar{\nu_\mu}) = \left(m_\tau - m_{\nu_\tau} - m_\mu - m_{\bar{\nu_\mu}}\right) c^2 \simeq 1.67 \text{ GeV}$$

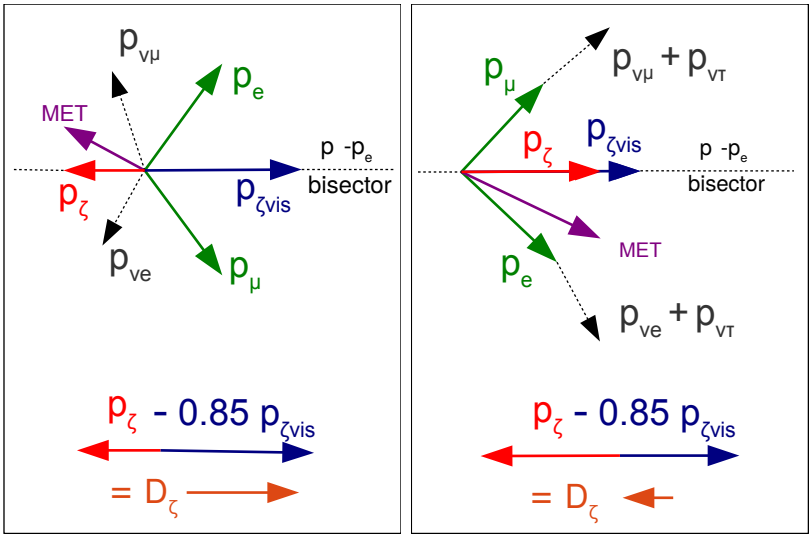


Figure 4.10.: Graphical representation of the D_ζ variable (orange arrow), in the case of $t\bar{t}$ decay (left) and in the case of $\phi/Z \to \tau\tau$ decay (right, where the kinematics of the $\tau\tau$ decay give it a small value.

As a consequence of this small amount of available energy, the decay products of highly boosted τ leptons — such as in the case of Higgs decays — tend to be contained inside a very narrow cone around the trajectory of the original τ and tend to share in approximately equal parts momentum and energy of the original τ. This fact implies that in Higgs events the missing transverse energy, being composed to a large part by the sum of the 4-momenta of the neutrinos, is highly correlated with the sum of the 4-momenta of the two leptons. This is the basis on which the D_ζ variable is built (see Figure 4.10).

Top quark decay kinematics

The top quark decays into a W boson and a quark, where the quark is almost always a b quark [126]. In cases where the W boson decays into a neutrino and a charged lepton the latter can pass the cuts of the lepton selection.

If both the top quarks in a pair decay leptonically (event that occurs $(9.4\%)^2 = 0.88\%$ of the times [126]), the event can mimic a $\tau\tau$ event and be accepted into the analysis selection.

Contrary to the case of τ, the decay chain of a top quark shows markedly different kinematics: the difference in the scale of energy of the primary top quark decay and the secondary W boson one is not so pronounced.

$$Q\left(t \to Wb\right) = \left(m_t - m_W - m_b\right) c^2 \simeq 88.5 \text{ GeV}$$

$$Q\left(W \to \nu_\mu \mu\right) \simeq Q\left(W \to \nu_e e\right) = \left(m_W - m_{\nu_e} - m_e\right) c^2 \simeq 80.3 \text{ GeV}$$

As a consequence, the products of the W boson decay have a kinetic energy in the frame of reference of the W boson which is of the order of the boost of the latter in the rest frame of the laboratory. It follows that the charged leptons emerging from a top quark decay chain can have trajectories far apart from the associated neutrinos, and therefore the missing transverse energy is much less correlated with the common 4-momentum of the two charged leptons (see Figure 4.10).

WW and other electroweak decay kinematics

Events involving the production and leptonic decay of two W bosons are kinematically similar to leptonic $t\bar{t}$ decays, since the W bosons have a decay energy which is of comparable order to their boost in the laboratory frame. As a result, strategies effective in reducing $t\bar{t}$ background using MET and lepton kinematics work suitably well in reducing WW backgrounds.

Other electroweak processes with smaller cross section and acceptance — such as WZ, ZZ and single top quark production — are similarly characterized by a smaller correlation between the momenta of the leptons and the MET than in the case of $\phi \to \tau\tau \to 2\ell 4\nu$, and therefore their contribution is also reduced by imposing a cut on the D_ζ variable.

4.4. Background Control Event Categories

Each Standard Model background is monitored by using one or more dedicated control categories. The purpose is to verify the quality of the background description in the Monte Carlo samples and the appropriateness of the corrections further applied to those samples. The control categories cover Drell-Yan processes, top quark pair and WW production.

4.4.1. Drell-Yan

Drell-Yan processes are monitored using events from a category that is defined as follows.

To filter out $t\bar{t}$ events no b-jet is allowed and MET cannot have a value above 40 GeV, the latter cut also filters part of the WW background. Furthermore, the minimum distance between the electron and muon track has to be less than 200 μm, as this reduces the amount of QCD/W+Jets events entering the selection.

The fact that the only significant Drell-Yan process among the $e\mu$ channel backgrounds is the one that produces a $\tau\tau$ intermediate state implies that the D_ζ variable can be used to select those decays. In this case a cut $D_\zeta > -20$ GeV/c is applied.

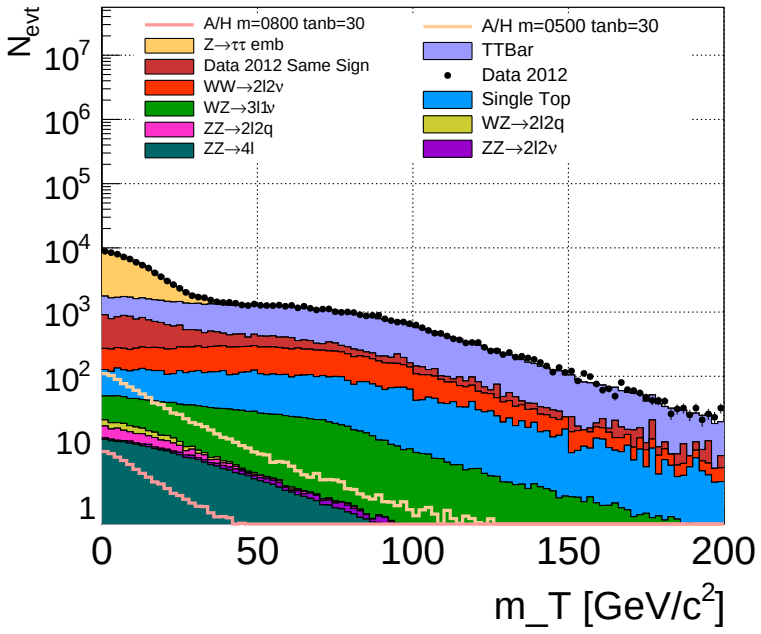


Figure 4.11.: Minimum between the transverse mass of the electron and that of the muon in the inclusive category.

A particularly sensitive variable that helps separating Drell-Yan events from the other processes is the minimum between the transverse mass of the electron and that of the muon, shown in figure 4.11, where the definition of the transverse mass for a lepton is show in equation 4.8. Differently from events from other processes, the Drell-Yan ones (yellow color in the figure) cluster around values of zero, providing good separation power against other kinds of events.

$$m_{T_\ell} = \sqrt{2\, p_{T_\ell} \not{E}_T (1 - \cos|\phi_\ell - \phi_{\not{E}_T}|)} \tag{4.8}$$

In the Drell-Yan category a cut is imposed on the minimum transverse mass, such that events have to satisfy $\min(m_{T_e}, m_{T_\mu}) < 15$ GeV/c^2.

The distribution of the invariant mass of the $e\mu$ pair and that of the $\tau\tau$ pair are shown in figure 4.12.

4.4.2. Top Quark Pairs

Top quark events tend to contain hard jets - many of which are b-tagged - and a large missing transverse energy. For these reasons the monitoring category dedicated to these processes is chosen to have two or more hard jets ($p_T > 30$ GeV/c), at least one b-tagged jet and a missing transverse energy larger than 80 GeV.

In figure 4.13 the invariant mass of the $e\mu$ pair and that of the $\tau\tau$ pair are shown for the top quark enriched category. The small disagreement at high dilepton mass is properly managed by the systematic uncertainty from $t\bar{t}$ p_T reweighting, described in the next chapter.

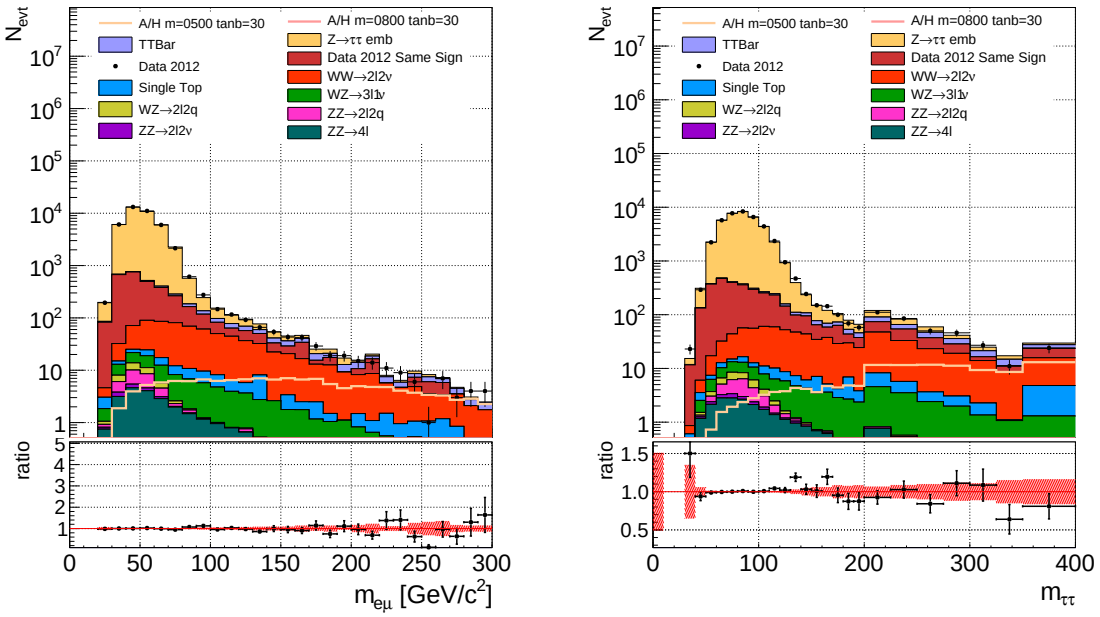


Figure 4.12.: Invariant mass of the $e\mu$ pairs (left, in log scale) and $\tau\tau$ pairs (right, in log scale) in the Drell-Yan control category.

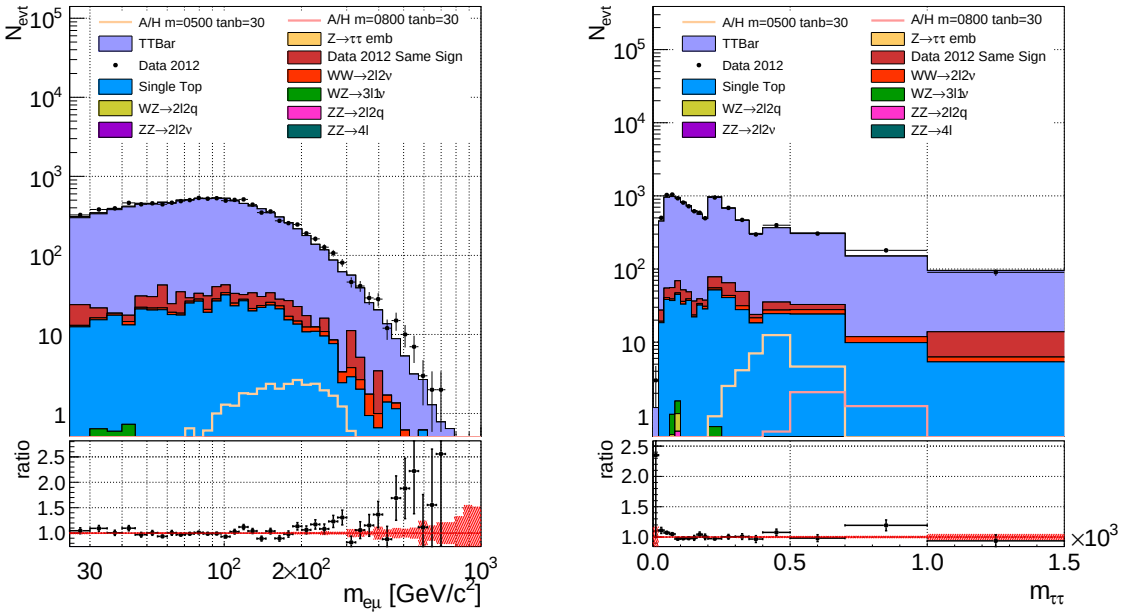


Figure 4.13.: Invariant mass of the $e\mu$ pairs (left, in log scale) and $\tau\tau$ pairs (right, also in log scale) in the top quark pair control category. The category reaches a good purity of $t\bar{t}$ events, with some contribution from single top.

4.4.3. Di-bosons

The production process of W pairs has a very low probability of producing b quarks (see figure 4.6), therefore no b-jets are allowed for the events in this category. The WW component is enriched by requiring MET to be in the range 40 GeV $< \not{E}_T <$ 80 GeV. Additionally, since the $e\mu$ pairs from a rapidly decaying boson as the W are prompt, a cut on the minimum distance between the electron and the muon track is imposed: $\mathrm{DCA}_{e\mu} < 80$ μm.

4.5. Signal Categories

In section 4.3.2 the construction of the D_ζ discriminant, its physical justification and the benefits of using it were described. As discussed in section 1.4, the significant production processes for a neutral MSSM Higgs boson at the LHC are two: the gluon fusion process (referred to as "ggϕ") and the b-quark associated production process (referred to as "bbϕ"). These processes have different signatures, therefore it makes sense to use two different signal categories, each optimized for the specific process. The selection for the signal categories is summarized in table 4.8.

As the signal categories are designed to enrich the fraction of genuine $\phi \to \tau\tau$ events, a cut on the value of the D_ζ discriminant is applied in both of them. The cut is the same for both categories: D_ζ is required to have a value above -20 GeV/c.

In order to simplify the statistical treatment of the backgrounds and to reduce the number of nuisance parameters that the limit setting package has to manage, some backgrounds are grouped together. The diboson and single top quark production processes are grouped into a single "Electroweak" background template.

		$N_{\mathbf{jets}}$ with $p_T > 30$ GeV		
		0	1	2+
$N_{\mathbf{b\text{-}jets}}$	0	no-btag category		
	1+	btag category		reject event

Table 4.8.: Graphical representation of the signal categories.

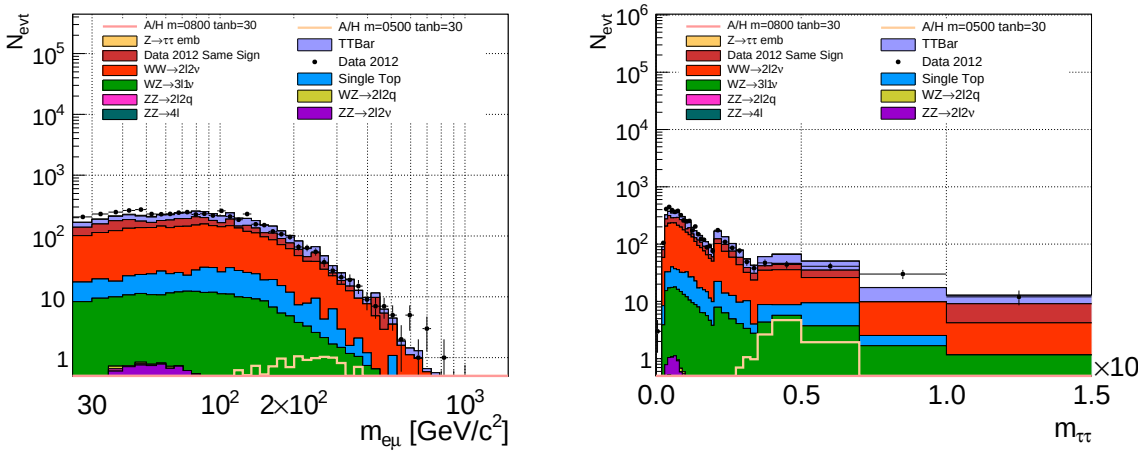


Figure 4.14.: Invariant mass of $e\mu$ pairs (left) and $\tau\tau$ pairs (right) in the top quark pair control category. The category reaches a good purity of $t\bar{t}$ events, with some contribution from single top.

4.5.1. Btag Category

The cuts used to define the btag category are tuned to select bbϕ events. At least one b-jet with a $p_T > 20$ GeV/c is required, but no more than one "hard" jet, that is, one with a $p_T > 30$ GeV/c. The effectiveness of these two cuts lies in the fact that b quarks associated to Higgs boson production are soft, while b quarks emerging from top quark decay are significantly harder.

Relative to this category, figure 4.15 shows the spectrum in $m_{e\mu}$ and $m_{\tau\tau}$ of the central Monte Carlo templates as provided as input to the statistical analysis package, where the latter has been described in section 3.10, while figure 4.16 shows the $m_{\tau\tau}$ histogram after the maximum likelihood fit performed during the limit extraction process.

4.5.2. Nobtag Category

The cut used to define the nobtag category is tuned to select ggϕ events, although a fraction of bbϕ events also enters the selection (i.e. those events in which the harder b-jet is not identified, or in which none of them reach the 20 GeV/c threshold). The requirement is simple: no b-jets should be present in the event.

As most $t\bar{t}$ events will have either one or two b-jets properly identified, there is a significant suppression of this kind of background, to the point that a considerable fraction of the background in the high mass tail is composed of WW events (see the $e\mu$ and $\tau\tau$ mass spectra before the maximum likelihood fit in figure 4.17). At low mass, instead, the dominating background are Drell-Yan events.

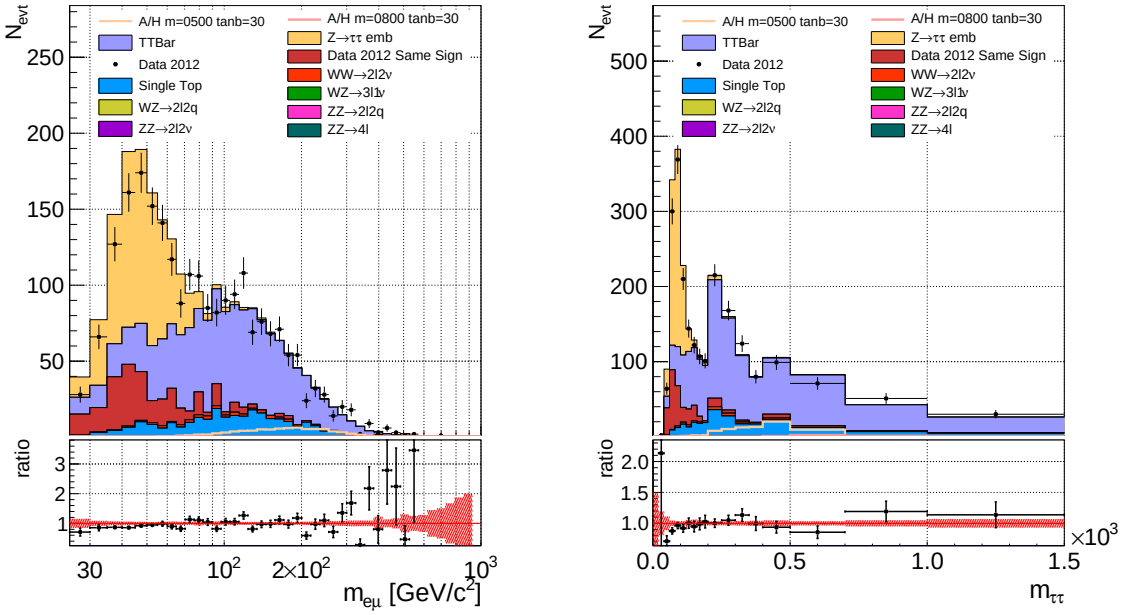


Figure 4.15.: Invariant mass of the $e\mu$ pairs (top) and $\tau\tau$ pairs (bottom) in the btag category.

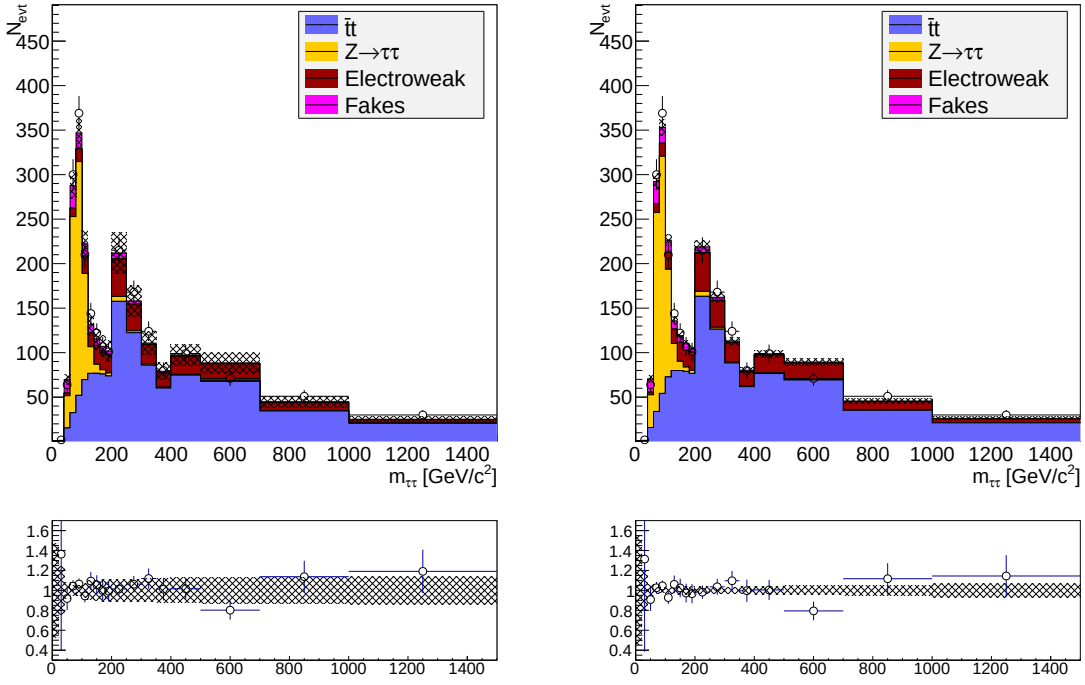


Figure 4.16.: Invariant mass of the $\tau\tau$ pairs in the btag category as seen by the statistical analysis package, showing the background templates before (left) and after (right) the maximum likelihood fit has converged.

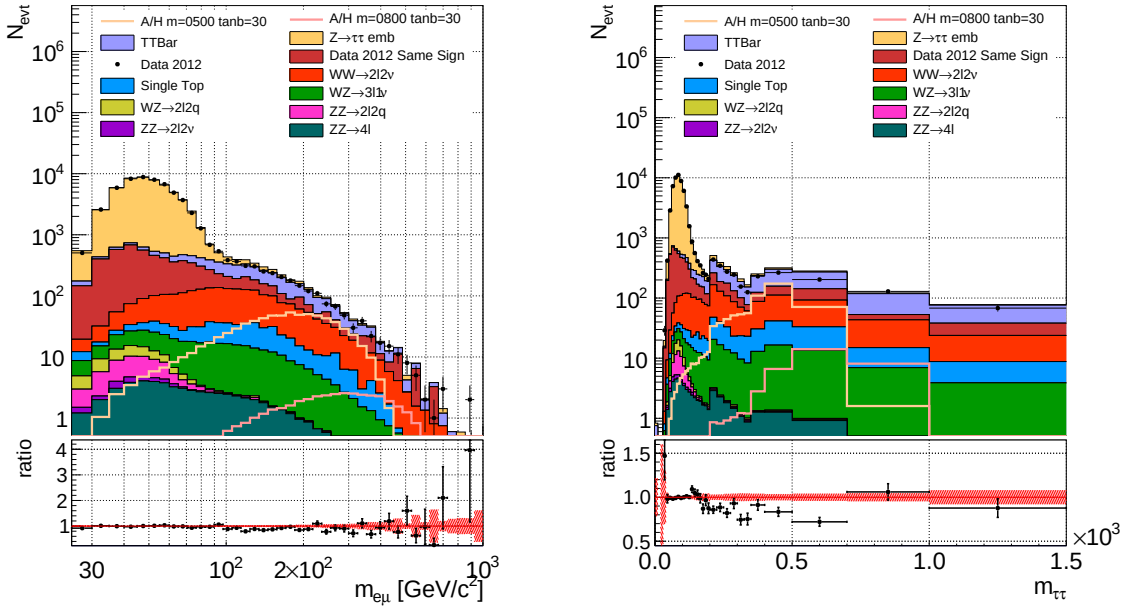


Figure 4.17.: Invariant mass of the $e\mu$ pairs (left) and $\tau\tau$ pairs (right) in the nobtag category.

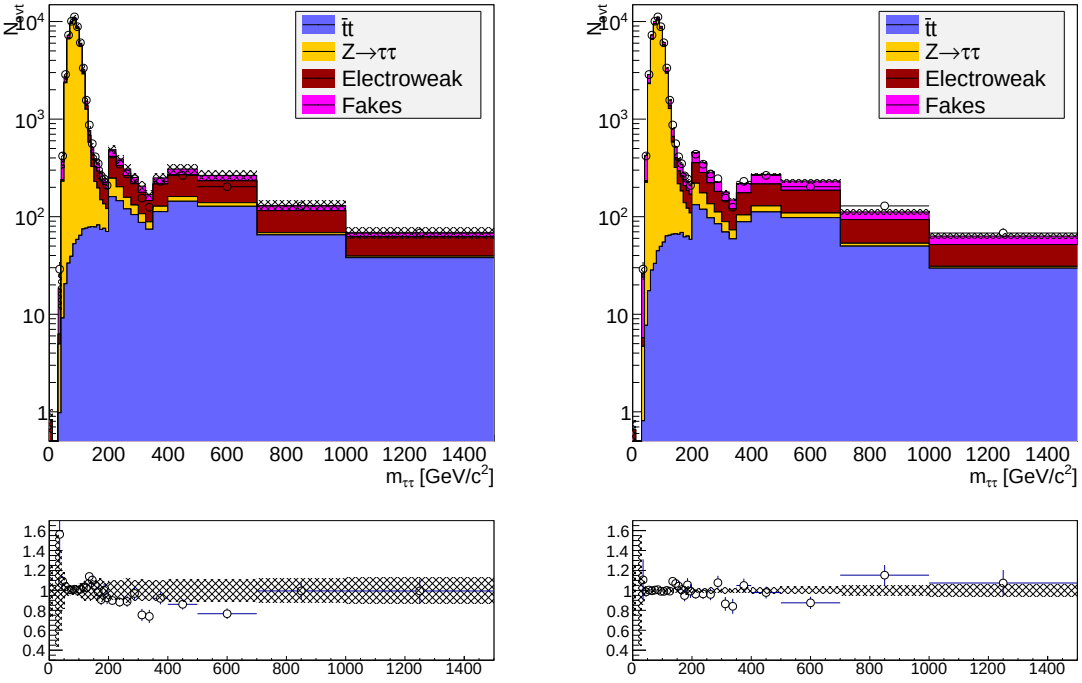


Figure 4.18.: Invariant mass of the $\tau\tau$ pairs in the nobtag category as seen by the statistical analysis package, showing the background templates before (left) and after (right) the maximum likelihood fit has converged.

4.6. Limits extracted from the Signal Categories

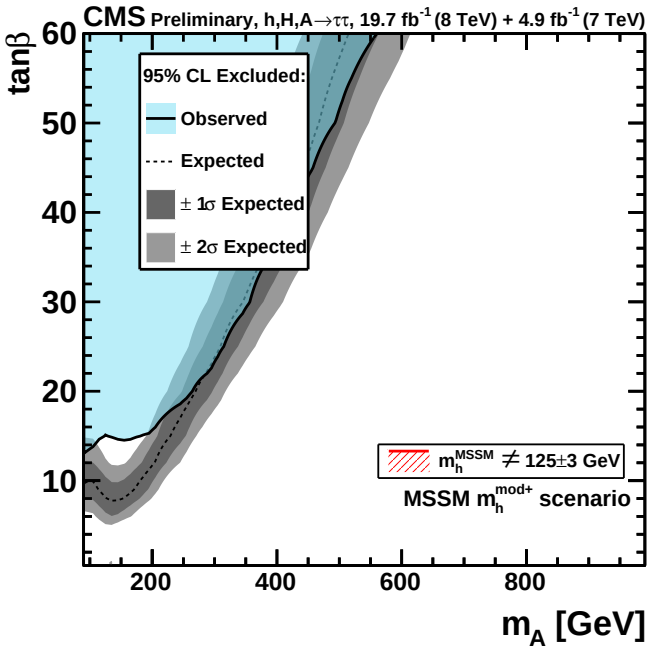


Figure 4.19.: Asymptotic limit in the $m_A - \tan\beta$ plane for the selection based on the D_ζ variable, using $\tau\tau$ mass as the probe statistics.

The $m_{\tau\tau}$ variable is reconstructed by the SVFIT code, which has been introduced in section 3.9. SVFIT is used to calculate the most likely $\tau\tau$ mass from the kinematic variables of the leptons and from the missing transverse energy. This mass is used by the statistical analysis framework – described in section 3.10 – to derive limits on the cross section for the production of a Higgs boson, as well as limits in the $m_A - \tan\beta$ plane in the context of the MSSM model. The treatment of the systematic uncertainties on the measurement is described in detail in the following chapter, in section 5.4.

The use of $m_{\tau\tau}$ instead of the simpler $m_{e\mu}$ originated in the search of a light Higgs boson at small masses, where it is necessary to attain a good sensitivity in the search. The improvement this choice brings on the separation between the Z peak and that of the Higgs boson one has been discussed in section 3.9. The possibility of using $m_{e\mu}$ for the search at high mass will be investigated in chapter 5.

The asymptotic limits on the cross section for the bbϕ and ggϕ processes, multiplied by the branching ratio to $\tau\tau$ are shown in figures 4.20 and 4.21. In the case of the ggϕ process the limit exhibits an excess of the order of 3σ in the low mass region and a smaller deficit in the high mass region. Such a large excess has not been observed in the other $\tau\tau$ decay channels. It is reasonable to assume that the mismodeling of the MVA MET in Monte Carlo simulation has an impact on $\tau\tau$ mass reconstruction, as the SVFIT code needs such variable as an input and is critically dependent on its value.

The limit in the $m_A - \tan\beta$ plane for a MSSM Higgs boson is shown in figure 4.19. Similar to the cross section limits discussed above, the observed limit shows a disagreement with the expected one at low Higgs masses.

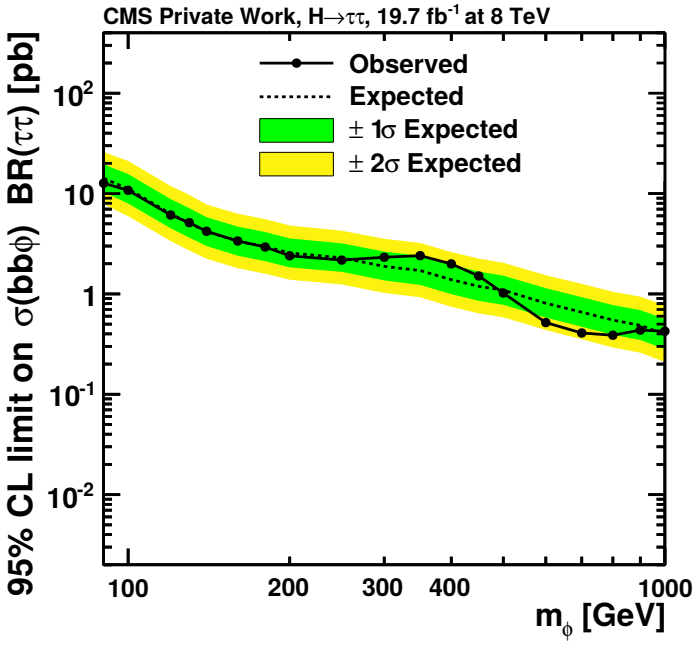


Figure 4.20.: Asymptotic limits on $\sigma(\mathrm{pp} \to \mathrm{bb}\phi) \cdot \mathrm{BR}(\phi \to \tau\tau)$ for the selection based on the D_ζ variable, using $\tau\tau$ mass as the probe statistics.

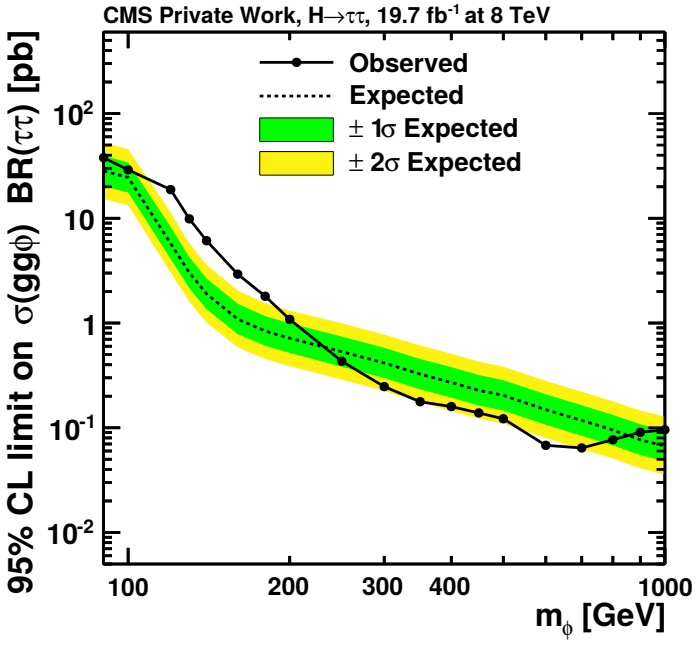


Figure 4.21.: Asymptotic limits on $\sigma(\mathrm{pp} \to \mathrm{gg}\phi) \cdot \mathrm{BR}(\phi \to \tau\tau)$ for the selection based on the D_ζ variable, using $\tau\tau$ mass as the probe statistics.

5. Optimization for the search at high mass

In this chapter possible optimizations of the event selection are investigated, with a the aim of increasing the sensitivity in the search of a high mass MSSM Higgs boson. The optimization involves the study of new discriminants and their combination in multivariate analysis discriminators implemented as boosted decision trees: one for each of the two signal categories.

Following the optimization of the event selection, this chapter covers the results of the statistical analysis when performed on the $m_{e\mu}$ distribution, where the result is compared to the case in which $m_{\tau\tau}$ is used. This arrangement is interesting since in the search region for a high mass MSSM Higgs boson (300 GeV$< m_{\phi} <$ 1000 GeV) the properties of the Standard Model background are substantially different from those in the low mass search region. This change in behavior is due to the fact that the composition of the background changes radically: the previously dominant Drell-Yan process gives way to top quark pair production and – to a lesser extent – W^+W^- production. These backgrounds are non-resonant, therefore the requirement of attaining a very high mass resolution is relaxed. It is thus possible to simplify the approach and use $m_{e\mu}$ for the statistical inference.

5.1. Improvement of Signal Discriminants

The argument discussed in the previous chapter for the use of D_ζ as a discriminant remains valid at high mass, as the backgrounds to which D_ζ is sensitive do not change and the decay mode of the Higgs boson does not change either. Nevertheless, since the momentum scale of the participating processes is different, one can expect that the optimal construction of the kinematic discriminant may change.

In the following sections possible optimizations of the α parameter in the expression for D_ζ are explored. New kinematic discriminants are built and the sensitive ones are combined in a Boosted Decision Tree. The improvement in terms of limits is then evaluated.

5.1.1. Optimization of D_ζ

The D_ζ discriminant has been defined in equation 4.7 as

$$D_\zeta = p_\zeta - \alpha \cdot p_{\zeta\mathbf{vis}} = \vec{\not{E}}_T \cdot \hat{\zeta} - \alpha \left(\overrightarrow{p_{T_e}} + \overrightarrow{p_{T_\mu}} \right) \cdot \hat{\zeta} \tag{5.1}$$

where the parameter α was given the value 0.85, optimized for the light SM Higgs boson search. In the following, the optimization of α for the different categories of the

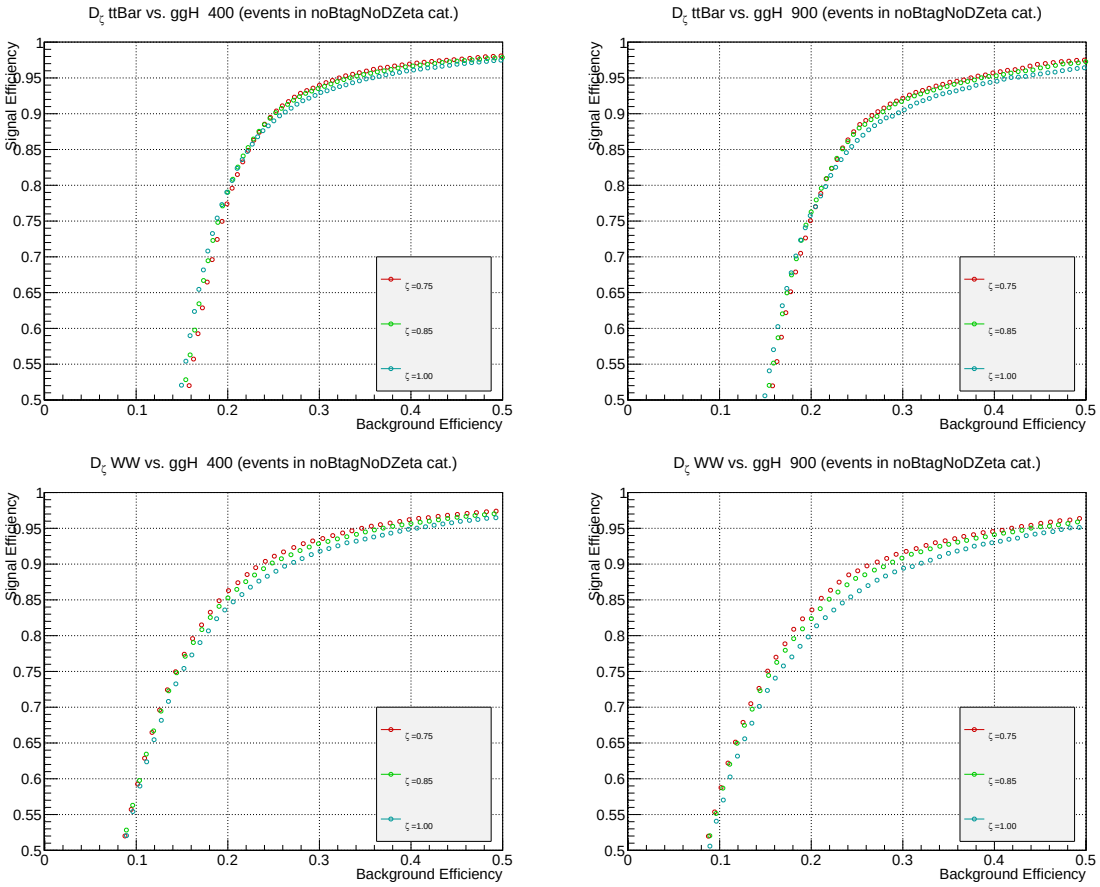


Figure 5.1.: ROC curves of D_ζ for events in the nobtag category for $\alpha = 0.75,\ 0.85,\ 1.0$. Signal is $gg\phi$ with a mass of 400 GeV (left) and 900 GeV (right) and is compared against $t\bar{t}$ (top) and W^+W^- (bottom) backgrounds.

high-mass Higgs search is discussed. The separation power of D_ζ is represented as a ROC curve (previously introduced in section 3.1) for many different possible values of the α parameter. The set of test values

$$\alpha = \{0.20, 0.25, 0.33, 0.50, 0.66, 0.75, 0.85, 1.00, 1.20, 1.33, 1.50, 1.66, 1.85, 2.00, 2.20, 2.50\}$$

has been chosen to uniformly cover the range in which the optimal value is expected to reside.

In the ROC curve representation, the value of α for which the best separation with respect to $t\bar{t}$ and WW backgrounds is achieved corresponds to the one for which the curve comes closest to the upper left corner in the diagram. This value of α will be taken as the one to use in the updated discriminant.

For the **nobtag category**, the comparison between the three values of α close to the optimal one is shown in figure 5.1. To study the behavior at both edges of the mass interval under search, two sets of figures are produced: in the left figures the mass of the Higgs boson is equal to 400 GeV/c^2, while in the right ones the mass is equal to 900 GeV/c^2. The Higgs boson is assumed to be produced via the $gg\phi$ process, which

is the dominant MSSM Higgs production process in that category. The ROC is drawn against the two main backgrounds, $t\bar{t}$ (upper plots) and WW (lower plots).

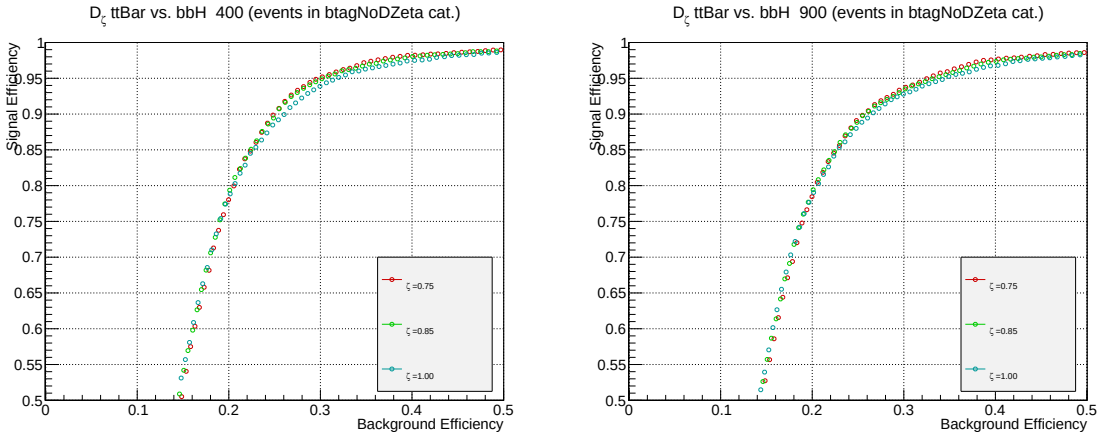


Figure 5.2.: ROC curves of D_ζ for events in the btag category for values of $\alpha = 0.75,\ 0.85,\ 1.0$. The signal is $gg\phi$ with a mass of 400 GeV (left) and 900 GeV (right) and is compared against $t\bar{t}$ background.

The best separation against $t\bar{t}$ background is obtained with values of α between 0.75 and 1.0, with the intermediate value being 0.85. The observed behavior of D_ζ as a function of α is consistent throughout the range of masses considered for the Higgs boson. In the case of WW background, best separation is obtained for lower values of α, but the difference is small. Since $t\bar{t}$ is the main background, the value $\alpha = 0.85$ is chosen.

For the **btag category**, only the $t\bar{t}$ background has been considered, as the contribution from WW is negligible. ROC curves are shown for $m_\phi = 400$ and 900 GeV/c^2, and the optimal value for α is found to be $\alpha = 0.85$, as in the case of the nobtag category. This result is again consistent between different Higgs mass hypotheses.

These results justify the choice of a value of $\alpha = 0.85$ in the expressions of D_ζ in both categories. The variable D_ζ will be further used as an input to the signal enrichment BDT.

5.1.2. The E_ζ Discriminant

The D_ζ discriminant depends on two sources of experimental uncertainty: one is the uncertainty on the kinematics of the leptons, the other is the uncertainty on the missing transverse energy direction and amplitude. Thanks to the very good tracking precision of the CMS detector for single charged particles, the former uncertainty is negligible when compared to the latter one. Missing transverse energy is indeed a variable that depends on a variety of factors, and its uncertainty can vary significantly between events. The observable D_ζ, being the difference between a term proportional to MET and a term depending only on the lepton kinematics, suffers from the uncertainty of MET.

A possible improvement involves dividing the entire expression for D_ζ by the magnitude of MET

$$\frac{D_\zeta}{|\vec{\not{E}}_T|} = \widehat{\not{E}} \cdot \hat{\zeta} - \alpha \, \frac{|\overrightarrow{p_{T_{e+\mu}}}|}{|\vec{\not{E}}_T|} \, \widehat{p_{T_{e+\mu}}} \cdot \hat{\zeta} \tag{5.2}$$

This expression has interesting properties. The first one is that the absolute value of MET appears only in the ratio term comparing it to the dilepton p_T: the contribution associated to the direction of MET and that associated to its magnitude appear in separate terms.

In the decay of τ leptons, the momentum of the τ is on average shared equally between the decay products, one charged lepton and two neutrinos. Therefore, for this kind of events the ratio $|\overrightarrow{p_{T_{e+\mu}}}|/|\vec{\not{E}}_T|$ should be independent of p_T and be distributed around a constant value. Bounds can be derived for the expression in equation 5.2. Indeed, the bisector between the p_T of each lepton is always contained in the hemisphere centered around the sum of the p_T of the leptons. It follows that:

$$0 \leq \widehat{p_{T_{e+\mu}}} \cdot \hat{\zeta} \leq 1 \tag{5.3}$$

And considering that α and $\frac{|\overrightarrow{p_{T_{e+\mu}}}|}{|\vec{\not{E}}_T|}$ are positive we have

$$- \alpha \, \frac{|\overrightarrow{p_{T_{e+\mu}}}|}{|\vec{\not{E}}_T|} \, \widehat{p_{T_{e+\mu}}} \cdot \hat{\zeta} \leq 0 \tag{5.4}$$

Combining the above with the fact that $-1 \leq \widehat{\not{E}_T} \cdot \hat{\zeta} \leq 1$ as it is the product of unit vectors we have that

$$\frac{D_\zeta}{|\vec{\not{E}}_T|} = \widehat{\not{E}} \cdot \hat{\zeta} - \alpha \, \frac{|\overrightarrow{p_{T_{e+\mu}}}|}{|\vec{\not{E}}_T|} \, \widehat{p_{T_{e+\mu}}} \cdot \hat{\zeta} \leq 1 \tag{5.5}$$

By subtracting 1 to each side of the inequality and multiplying by -1 one gets:

$$1 + \alpha \, \frac{|\overrightarrow{p_{T_{e+\mu}}}|}{|\vec{\not{E}}_T|} \, \widehat{p_{T_{e+\mu}}} \cdot \hat{\zeta} - \widehat{\not{E}} \cdot \hat{\zeta} \geq 0 \tag{5.6}$$

The logarithm of this expression is the definition for E_ζ

$$E_\zeta = \log_{10} \left(1 + \alpha \, \frac{|\overrightarrow{p_{T_{e+\mu}}}|}{|\vec{\not{E}}_T|} \, \widehat{p_{T_{e+\mu}}} \cdot \hat{\zeta} - \widehat{\not{E}} \cdot \hat{\zeta} \right) \tag{5.7}$$

$\phi \to \tau\tau$ events tend to cluster around $E_\zeta = 0$, as can be seen in the plot in figure 5.3. This is expected, from the way the variable was constructed from D_ζ and the fact that D_ζ itself clusters around 0. To simplify the cuts on the variable and make them one-sided, it is possible to take the modulus $|E_\zeta|$ as the discriminant, and select events with small values of $|E_\zeta|$. The distribution of this variable is shown in the last histogram in figure 5.3. A similar optimization to the case of D_ζ has been performed for the α parameter, finding the best value to be $\alpha = 1.33$.

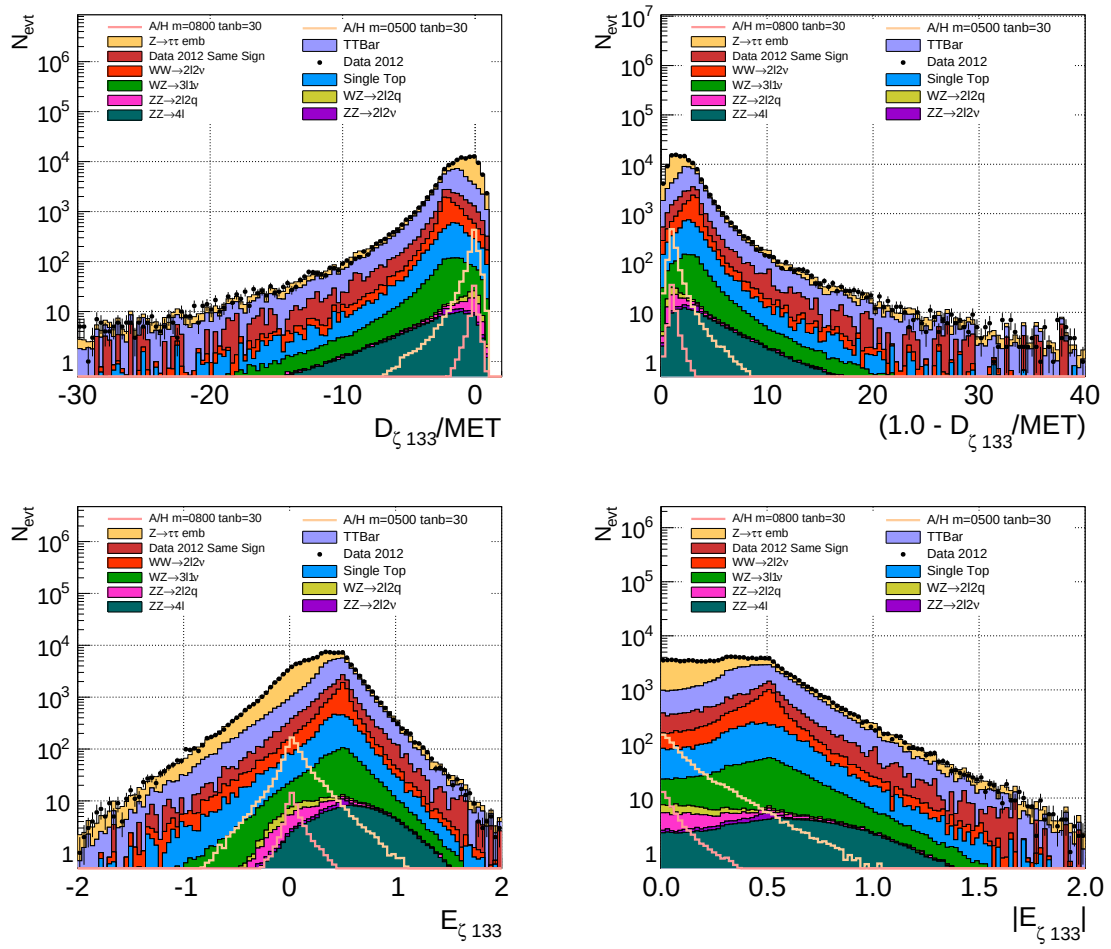


Figure 5.3.: Histograms for the inclusive category showing the distributions: $\frac{D_\zeta}{|\not{E}_T|}$ (top left), $1.0 - \frac{D_\zeta}{|\not{E}|}$ (top right), E_ζ (bottom left) and $|E_\zeta|$ (bottom right).

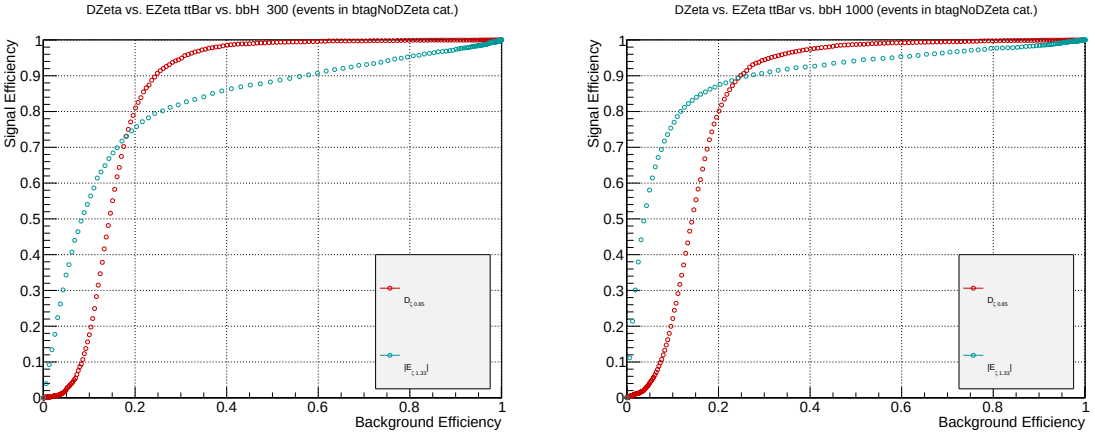


Figure 5.4.: Comparison of D_ζ and $|E_\zeta|$ separation power against $t\bar{t}$ background for different Higgs boson masses (left $m_\phi = 300$ GeV/c^2, right $m_\phi = 1000$ GeV/c^2.

The separation power of $|E_\zeta|$ is represented as a ROC curve and compared with D_ζ in figure 5.4. From this comparison it is evident that, differently from D_ζ, the separation of the E_ζ discriminant is not uniform as a function of mass. Indeed, a high separation is obtained mainly at high mass, while at lower Higgs mass the performance is worse than D_ζ.

We want to keep the sensitivity for heavy Higgs bosons as flat as possible across the interval of masses. It is therefore preferable that separation power of the chosen discriminant has a small correlation with the mass of the Higgs, since the estimator for the latter ($m_{\tau\tau}$) will be used for statistical inference of the limits. Following this point of view, and the considerations in the previous paragraph, $|E_\zeta|$ is not attractive compared to D_ζ in the selection of the events, due to the strong dependence of its separation power on the Higgs mass.

5.1.3. The κ_1 Discriminant

The observable κ_1 is defined as the ratio between the p_T of the dilepton system and the vector sum of the p_T of the individual leptons. An example of this variable in the btag category is shown in the top left plot in figure 5.5

$$\kappa_1 = \frac{|\overrightarrow{p_{Te}} + \overrightarrow{p_{T\mu}}|}{|\overrightarrow{p_{Te}}| + |\overrightarrow{p_{T\mu}}|} \tag{5.8}$$

The variable quantifies the extent to which the leptons emerge from the collision back-to-back in the transverse plane. Values close to 0 indicate that the leptons are back-to-back and have a similar transverse momentum, while values close to 1 indicate that the leptons emerge from the collision with a relatively small opening angle in the transverse plane.

As can be seen in figure 5.5, the observable κ_1 has little separation power between signal and WW background, but can separate signal events from $t\bar{t}$ ones. For that reason κ_1 has been included in the boosted decision tree which will be discussed in the following section.

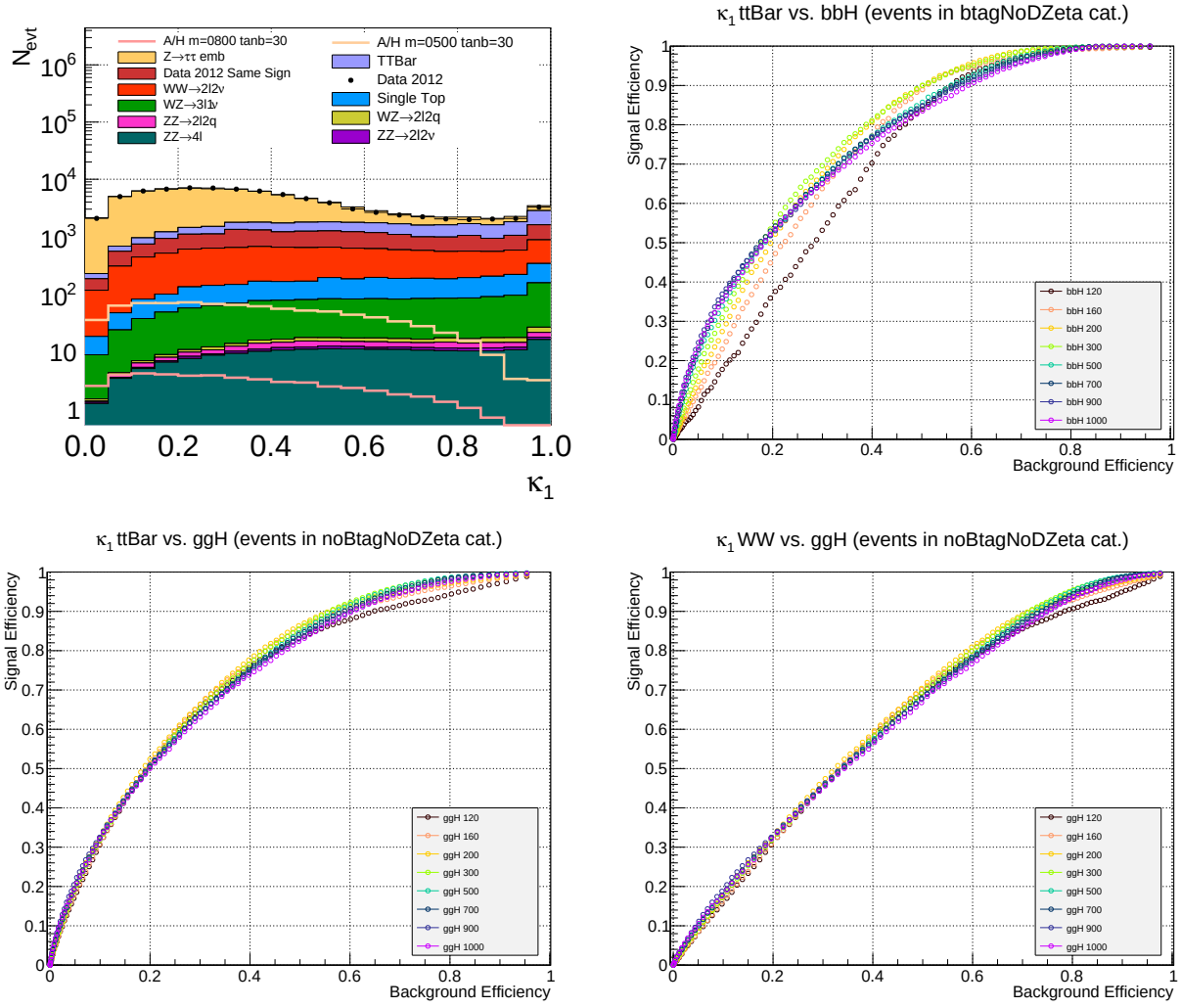


Figure 5.5.: Distribution of κ_1 in the nobtag category (top left). ROC curves showing the separation power of κ_1 against $t\bar{t}$ in the btag category (bottom left), $t\bar{t}$ in the nobtag category (top right) and WW in the nobtag category (bottom right).

5.2. Combination into Boosted Decision Trees

The previous paragraphs explored a series of observables that have some separation power in distinguishing signal and background. Each of these variables gives insight in one aspect that characterizes the signature of the events, an aspect that in general is not an optimal discriminator when taken alone. The combination of the variables into a multi variate analysis discriminator allows to distinguish events efficiently. In this analysis the combination is performed using boosted decision trees, which have been introduced in section 3.1.

Each of the two signal categories used in this analysis is optimized to select signal events produced by a specific process. The btag category selects preferentially events in which the Higgs boson is produced via b-associated production ($bb\phi$), while the nobtag category selects preferentially events in which the Higgs boson is produced via gluon fusion ($gg\phi$).

The Feynman diagram for the gluon fusion process is shown in figure 5.6. The process involves a loop at leading order and, therefore, both the transverse momentum spectrum of the Higgs boson and that of possible radiated jets depend significantly on the theoretical assumptions. To reduce the dependence on these assumptions it is preferable not to select events with cuts that have a dependence on these transverse momenta. In contrast, in the btag category this problem poses a minor issue, as the leading processes are described by tree-level diagrams and processes involving loop diagrams are important only from the point of view of corrections. Also, due to the selection cuts, in the btag category all events are guaranteed to have at least one jet which is b-tagged, so variables involving the properties of the b-jet can be used for selection.

Following these considerations it is convenient to select events in the two signal categories with separate BDTs: one for the nobtag category with a reduced number of input variables and one for the btag category using more variables involving also the b-jet. Comparing the performance in each category, with the help of the ROC curves that will be shown in the following, two configurations for the BDT were observed to yield the optimal performance: one is referred to in TMVA documentation simply as "BDT" and uses the adaptive boost method ("AdaBoost") during the training phase, the other is referred to as "BDTG", and involves the use of the gradient boost method. The latter has been chosen as the configuration used in this analysis.

The two BDTs are trained against an expected background mix of $t\bar{t}$ and WW events. The relative fraction in the number of $t\bar{t}$ and WW events in the training samples reflects the fractions of events predicted by Monte Carlo simulation for each category. The training for signal events has been performed on a mix of events representing, with equal weights, the range of Higgs boson masses of interest to this analysis and the two possible production channels.

A uniform distribution of masses is chosen in order to provide for equal sensitivity as a function of the mass. Variations in sensitivity can lead to a bias in the measurement of the Higgs boson mass in case of a possible discovery. The mass points considered are the following, in units of GeV/c^2: 300, 350, 400, 450, 500, 600, 700, 800, 900 and 1000. The large number of points guarantees a smooth transition between the medium

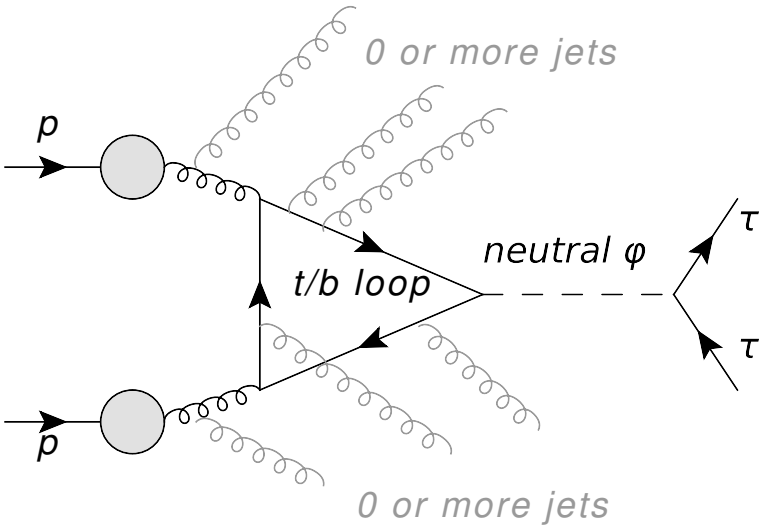


Figure 5.6.: Feynman diagram for the gluon fusion process in MSSM. In grey, possible gluon radiation diagrams are shown.

and the high mass regime.

5.2.1. Event Selection for the btag Category

Top quark pair events are expected to produce two hard b-jets, as opposed to the two soft b-jets produced in Higgs events. The presence of two b-tagged jet above 20 GeV (not exceeding the cut of not more than one jet above 30 GeV) in an event is a strong indication that the event is a background $t\bar{t}$ event.

The CSV b-tagging algorithm, though, is optimized for purity and is not fully efficient at its tight working point: in most of the cases the number of b-jets in the btag category is only one, since the other jet does not pass the tight CSV cut.

The number of jets passing the **medium** working point of the CSV algorithm shows, instead, a difference between signal and background. This difference can be used for discrimination.

A set of eight variables – whose distributions are shown in figure 5.7 – is used for the training of the BDT and later for the classification of the events in the btag category. These variables are

- The D_ζ discriminant, introduced previously, with a value of the parameter $\alpha = 0.85$
- The distance of closest approach between the electron and the muon tracks in the transverse plane ("2D DCA"). This variable can distinguish between prompt leptons and leptons emerging from the decay of a particle which lives long enough to travel a macroscopic distance, such as in the case of the τ lepton.
- The κ_1 discriminant introduced in the previous paragraphs.
- The p_T of the hardest lepton in the pair.
- The p_T of the softest lepton in the pair.

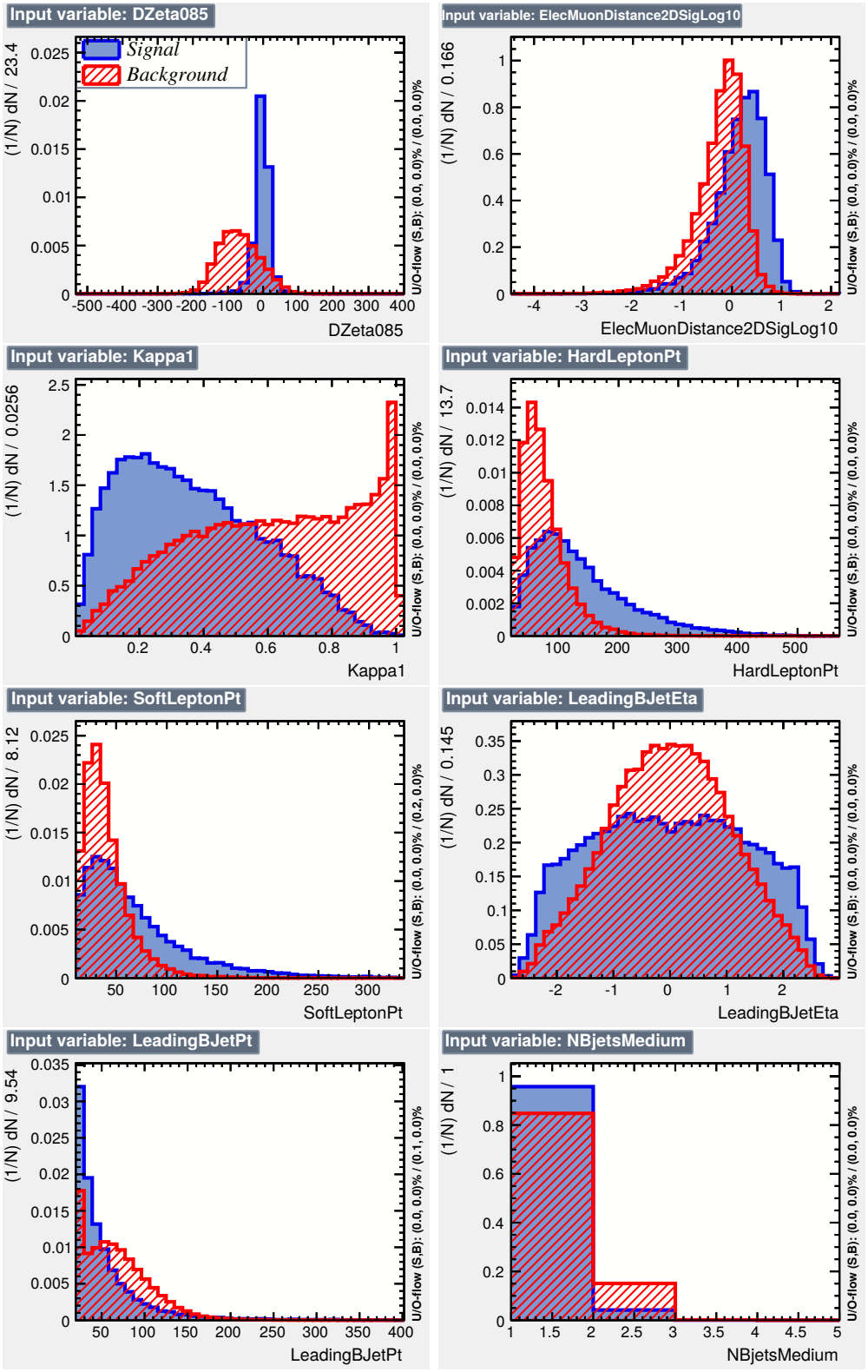


Figure 5.7.: Input variables for the btag BDT, background sample in red (hatched), signal sample in blue (full color). From left to right, line by line: $D_{\zeta\ 85}$, $\mathrm{DCA}_{2D\ e\mu}$, κ_1, hardest lepton p_T, softest lepton p_T, η and p_T of leading b-jet, number of b-jets passing the medium working point of the CSV discriminator.

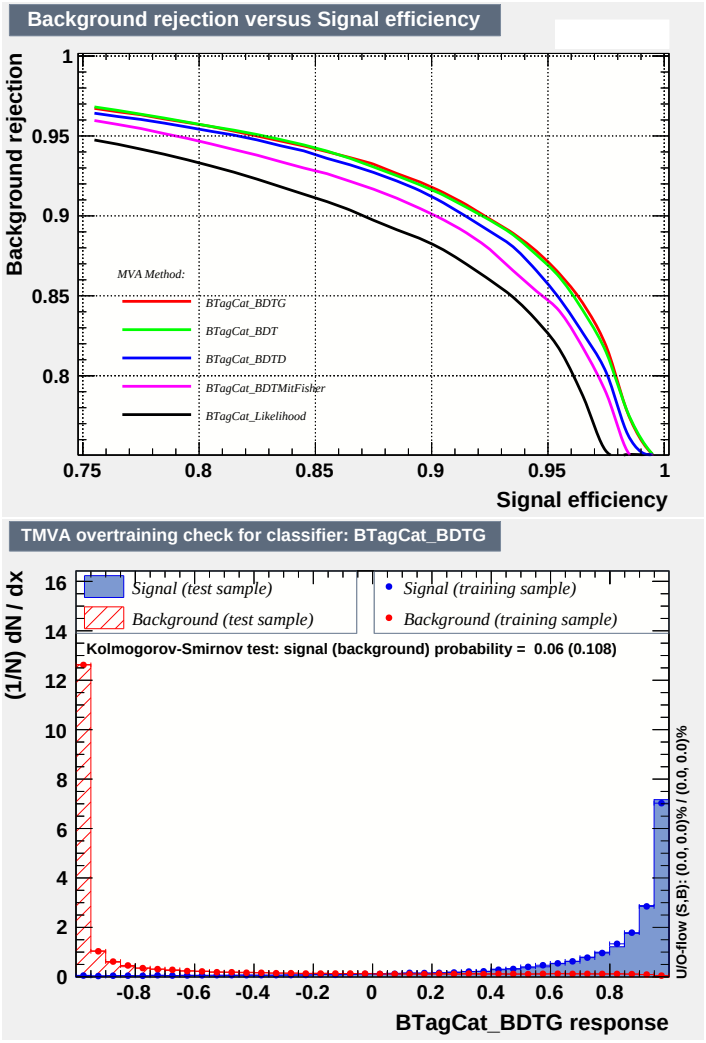


Figure 5.8.: Top: ROC curves of different BDT discriminators (colored) and a likelihood approach (black). Bottom: output distribution of the btag BDT showing signal (blue, full color) and background (red, hatched).

- The pseudorapidity of the hardest b-tagged jet.
- The p_T of the hardest b-tagged jet.
- The number of jets that pass the medium working point of the CSV b-tagging algorithm.

Some of the variables are correlated, as it can be seen in the linear correlation matrix in figure 5.9. The advantage of using a Boosted Decision Tree is that this kind of classifier is capable of effectively make use of correlations between the input variables.

The correlation matrix shows that the leptons are correlated in transverse momentum. This is expected, because a boost of the intermediate state (Higgs, $t\bar{t}$ or WW) leads to an increase in p_T for both of the leptons. Furthermore, the construction of κ_1 as a function of the transverse momenta of the leptons introduces an anticorrelation with these latter variables, because a larger boost of the intermediate state corresponds

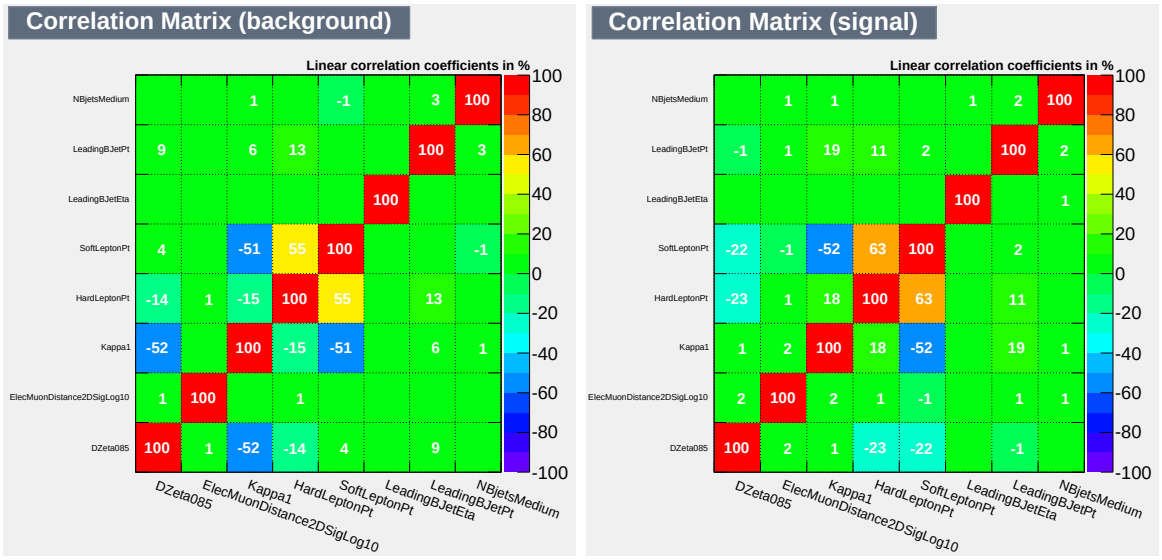


Figure 5.9.: Linear correlation matrix of the input variables in the background training sample (left) and in the signal training sample (right) for the btag BDT.

to a smaller opening angle and larger p_T of the leptons. The variable κ_1 is also anticorrelated with D_ζ in the background sample, because values of κ_1 close to the maximum of 1.0 imply that the lepton pair is boosted in the transverse plane. This in turns means that the negative term in the D_ζ expression – dependent on the p_T of the lepton pair – is large, hence D_ζ reaches lower values. In the signal sample this correlation is not present, because the values of D_ζ are always close to zero.

The gain resulting from the use of a BDT, and its correct treatment of the correlations, is evident in the ROC curve in figure 5.8. In this figure the performance of different BDT settings and that of a simple likelihood discriminant are compared. The BDT performance is superior, and this is represented by its curve being closer to the upper right corner of the plot, which corresponds to the point of maximum efficiency and purity.

5.2.2. Event Selection for the nobtag Category

The only input variables for the BDT in this category are the D_ζ discriminant and the $e\mu$ DCA, whose distributions are shown in figure 5.10. The linear correlation between these two variables is negligibly small. This lack of correlation is an expected feature, since D_ζ depends on the kinematics of the individual τ, top quark or W boson decay, while the DCA depends on the resolution of the detector and from the distance traveled by *both* the τ leptons before decaying. Having largely uncorrelated variables means that in this case the gain in performance from the use of a BDT instead of a likelihood is smaller compared to the case of the btag category. This can be seen in the performance ROC curve in figure 5.11, where the line corresponding to the likelihood and the ones corresponding to the BDT are almost on top of each other.

Figure 5.12 shows the distribution of the output values of the nobtag BDT. This

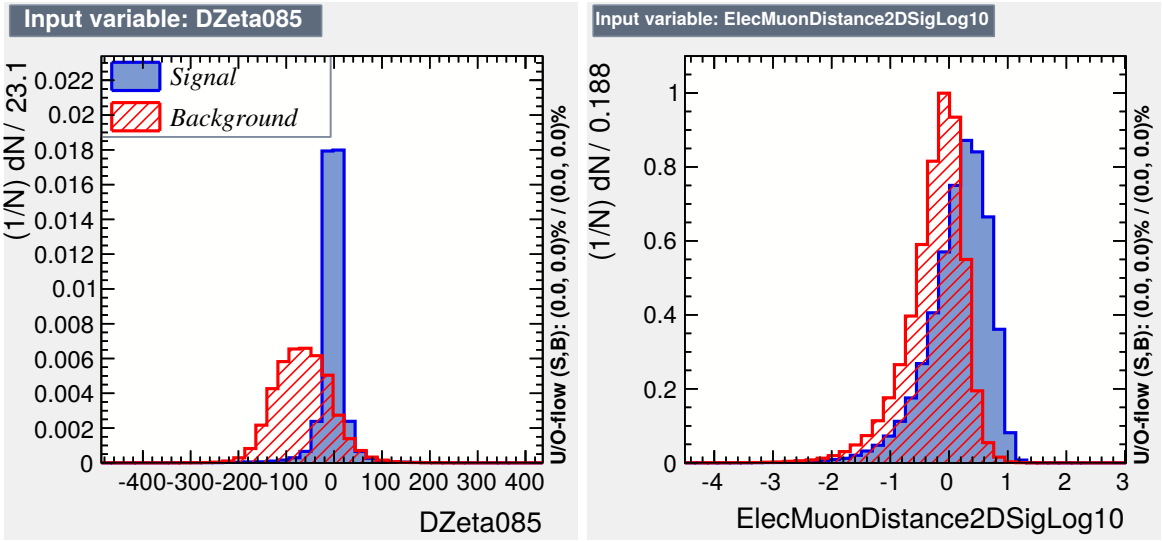


Figure 5.10.: The two input variables for the nobtag BDT: (left) $D_{\zeta\ 85}$, (right) the 2-dimensional $e\mu$ DCA. Signal is shown in blue (full color), background in red (hatched).

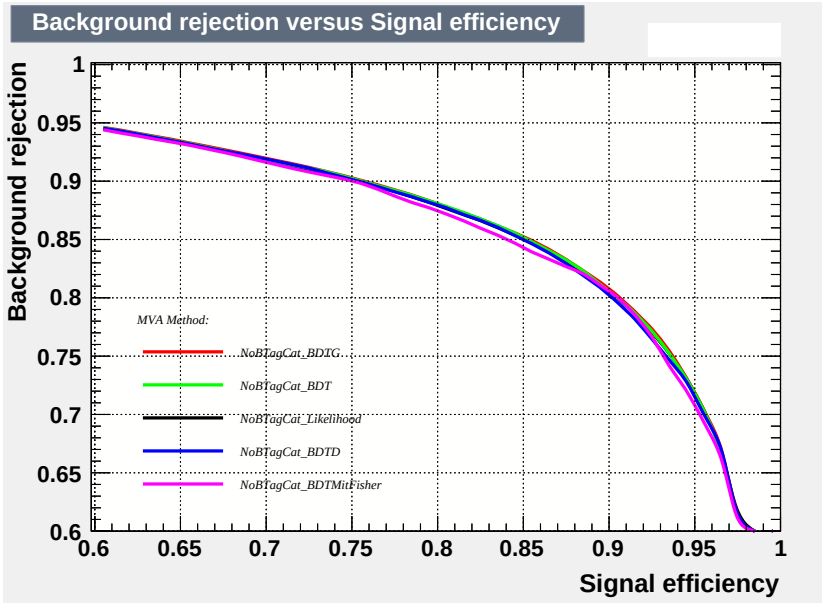


Figure 5.11.: ROC curve for various BDT (colored) and likelihood (black) discriminants for the nobtag category.

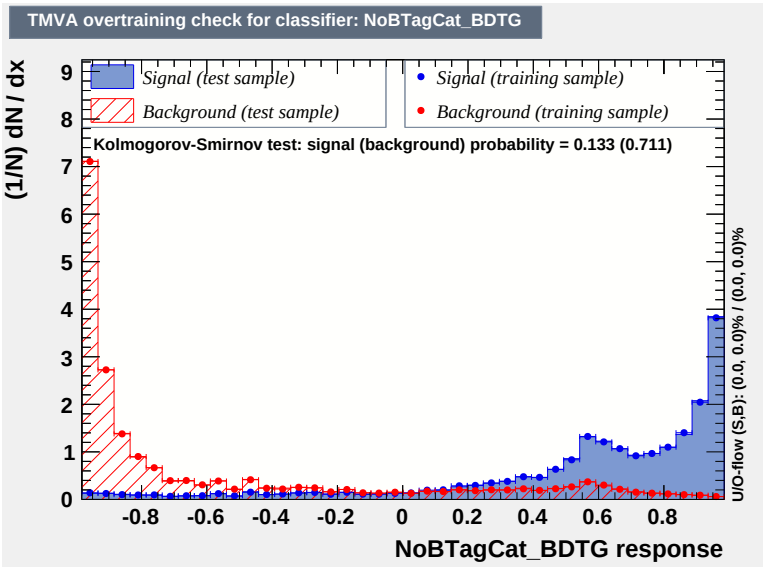


Figure 5.12.: Output of the nobtag BDT for signal (full blue) and background (hatched red).

distribution, in which the outputs of background and signal are "stretched" apart towards the boundaries of their ranges, is typical of the training performed by the gradient boost method

5.3. Comparison of BDT and Individual Discriminating Variables

In the following figures, the performance of the BDT is compared with the performance of the discriminators discussed in the previous paragraphs. At the top of figure 5.13 the performance in separation of the $t\bar{t}$ background from the Higgs boson signal is shown for the btag category. The BDT classifier is shown to bring a substantial improvement both at the lower edge of the mass interval being searched (400 GeV, left figure) and at the higher edge (900 GeV, right figure).

Similar ROC curves for the nobtag category are shown at the bottom of the same figure. In this case, a significant improvement in the rejection of $t\bar{t}$ events is observed when using a BDT. The rejection of the WW background, instead, does not show a significant improvement. However, considering that the dominant background is due to $t\bar{t}$ events the choice of using the BDT is justified.

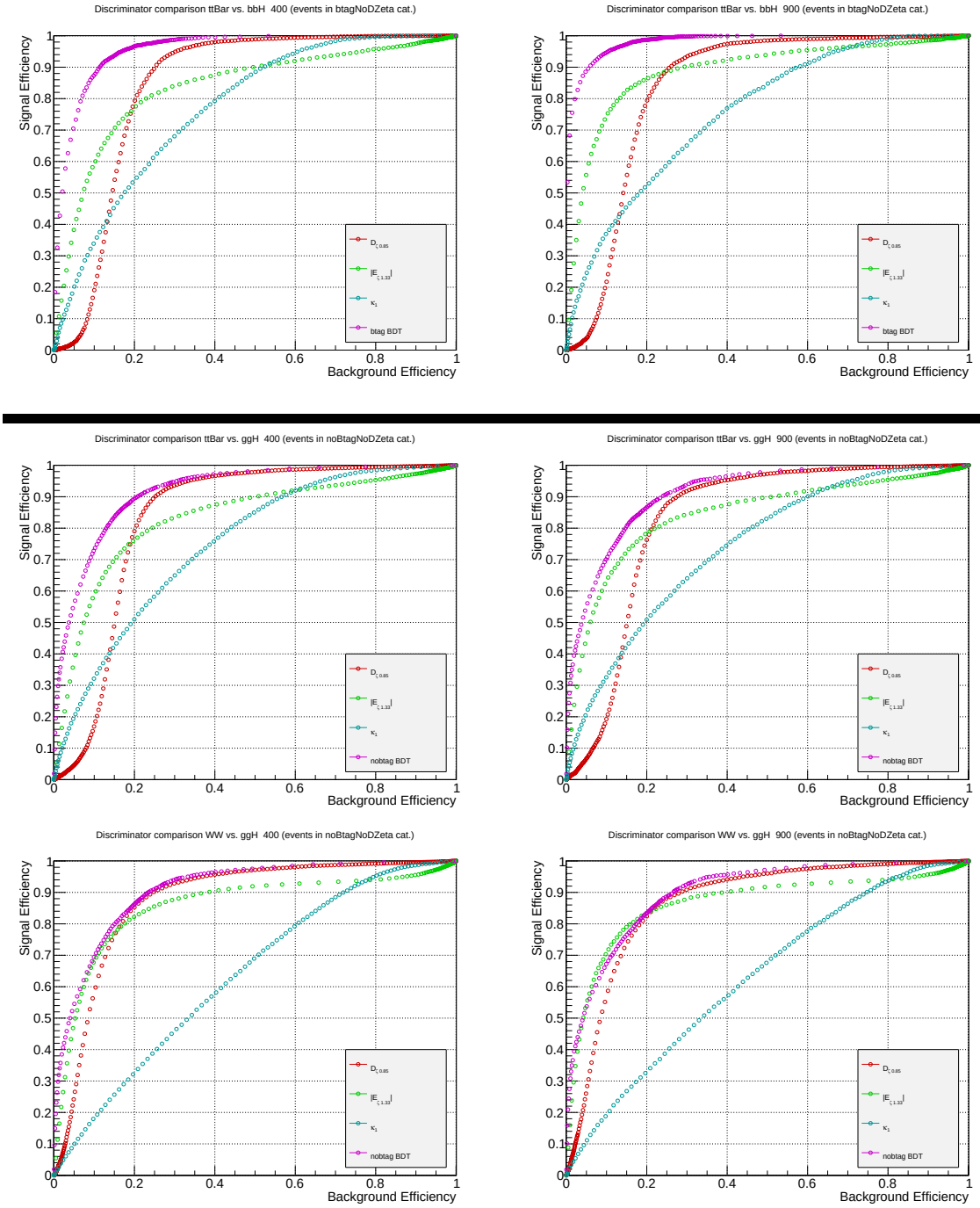


Figure 5.13.: ROC curves of different discriminants, comparing the separation between a signal at 400 GeV (left) and 900 GeV (right) and $t\bar{t}$ in the btag category (top row, above the line), $t\bar{t}$ (center) and WW (bottom) in the nobtag category. The performance of the respective BDT is shown with magenta circles.

5.4. Systematic Uncertainties

Various sources of systematic uncertainties can affect the number of events n_i counted in each bin of the mass histogram. Those which have a significant impact are implemented in the fit as nuisance parameters. They are extracted from data such that and their contribution to the uncertainty of the limit is taken into account.

The algorithm performing the fit has been described in section 3.10. The nuisance parameters that have been considered in this analysis are described in the following paragraphs, and the expected variance of their distribution is presented. All the systematic uncertainties, unless specified otherwise, are assumed to be distributed as a *log-normal* distribution, as recommended by CMS and the LHC Higgs Combination Group [127], and the systematic uncertainties – reported in tables A.1, A.2, A.3 and A.4 – are the fluctuations in sample yield induced by that specific nuisance parameter.

5.4.1. Experimental Systematic Uncertainties

Systematic uncertainties due to experimental sources have different origins:

- The use of parameters obtained from control regions or in datasets different from the one used for the analysis, e.g. the efficiency of the electron trigger. These uncertainties represent the systematic and statistical errors on these parameters.
- Approximations and assumptions made on the distribution of these parameters, for example on the assumed shape of the $m_{\tau\tau}$ distribution in W+jets and QCD events.
- Errors introduced from the analysis method. One example is the possible bias introduced by using the embedded sample to model the $Z \to \tau\tau$ background.

Luminosity

The exclusion on the existence of a MSSM Higgs boson is ultimately a statement on its cross section: if the upper limit on the cross section, corresponding to a previously chosen Confidence Level, is measured to be below the one predicted by the theory for a certain pair of parameters $(m_A, \tan\beta)$, then the validity of the theory is excluded for that specific combination of $(m_A, \tan\beta)$ values. Integrated luminosity $\mathscr{L}$ is a multiplicative term used to normalize all Monte Carlo samples, and therefore has a profound effect on the matching between data and simulation and on the estimation of cross sections and limits.

CMS can measure the instantaneous luminosity L at the collision point employing various kinds of detectors and techniques. The most precise and currently used one combines the measurement of the instantaneous visible event rate, through the measurement of the occupancy in the pixel detector at each bunch crossing, with the physical measurement of the beam size in the transverse plane through a *Van Der Meer scan* [128]. The measurement on the total integrated luminosity is assigned a systematic uncertainty of 2.6% [128]. This uncertainty affects all simulated samples **and** the embedded $Z \to \tau\tau$ sample (since it is normalized to the $Z \to \tau\tau$ Monte Carlo simulation), as these samples have been rescaled to the measured integrated luminosity for the data sample.

Electron efficiency

The electron identification and trigger efficiency are measured in bins of η and p_T with the help of the *Tag and Probe* method, which allows to select a pure and unbiased sample of electrons when they are produced by a known resonance, such as the Z boson or the charmonium states ($J/\psi, \psi', \ldots$). The scale factors between the efficiencies measured in data and those measured in Monte Carlo simulation are applied to the simulated samples and the embedded $Z \to \tau\tau$ sample (since the leptons in these samples are simulated). The systematic uncertainty on the electron efficiency is assigned a value of 2%.

Muon efficiency

In a similar way to the electron efficiency, muon identification and trigger efficiency are measured in bins of η and p_T via *Tag and Probe* and the corresponding data-to-Monte Carlo scale factors are applied to all simulated data and to the embedded $Z \to \tau\tau$ sample. The systematic uncertainty on the muon efficiency is assigned a value of 2%. The uncertainty is applied, again, to Monte Carlo simulations and the embedded $Z \to \tau\tau$ sample.

Electron energy scale

The excellent performance of the ECAL detector in CMS, paired with the efficient association of bremsstrahlung photons performed by the GSF reconstruction algorithm, allows precise reconstruction of the electron energy in each event. Corrections are applied to make the detector response uniform across its sensitive region. Their systematic uncertainty is modeled by defining two templates for the histogram being used for statistical inference. One template (the *down template*) is obtained by shifting the electron energy and momentum down by 2% if its supercluster centroid falls in the barrel, 5% if it falls in the endcaps. The other template (the *up template*) applies the same shifts in the opposite direction. The up and down templates define a fluctuation of 2σ in the scale of electron energy.

This method of providing systematic uncertainties is referred to as *shape* method and it allows to model systematic uncertainties which do not induce a uniform rescaling of the event yields across the mass range. The electron energy scale falls into this category, as it indirectly shifts the mass of the reconstructed $\tau\tau$ pair. This systematic uncertainty is applied to backgrounds that present a peak in the $m_{\tau\tau}$ spectrum, that is, the embedded $Z \to \tau\tau$ sample, the MSSM Higgs signal and SM Higgs simulated samples.

Jet energy scale

Systematic errors in the energy scale of the jets can lead to migrations of events in and out of the event categories. Two or more systematic uncertainty terms account for the jet energy scale (JES), and are tailored to the respective categories.

For the nobtag category, which has relatively simple cuts driven by 0 b-tagged jets with a $p_t > 20$ GeV/c requirement, the MSSM Higgs signal is assigned a systematic uncertainty of 1% on its normalization. The $t\bar{t}$ and electroweak (single top, dibosons)

samples are assigned a systematic uncertainty of 1%, with its nuisance parameter being anticorrelated with the one defined for the Higgs samples.

The btag category has a composite cut on the jets, and therefore the systematic effects are more complex to model. The MSSM Higgs is assigned a JES systematic uncertainty of 4%. An additional 1% term, correlated with the former, is assigned to the $bb\phi$ Monte Carlo sample.

The top quark pair Monte Carlo sample is assigned an uncertainty of 7% (this is large since the jet spectrum for $t\bar{t}$ changes significantly in the p_T region in which the category cut is imposed) and the electroweak sample is assigned an uncertainty of 3%. Both these nuisance parameters are defined to be anticorrelated with the ones of the Higgs signal sample.

Efficiency of b-tagging

The modeling of the b-tagging algorithm in the simulation is not completely identical to the one in data. Since the presence of b-tagged jets is a fundamental cut in the signal category selection, possible systematic uncertainties on the b-tagging efficiency are accounted for. In the nobtag category, $t\bar{t}$ background is assigned an uncertainty of 5% while MSSM Higgs and electroweak backgrounds are assigned an uncertainty of 2%.

The most common case for the b-tagging efficiency systematic error to have an impact is when a single jet in an event is ambiguously b-tagged. An improper modeling of the b-tagging efficiency can affect the estimation of the migration of events between the nobtag and btag categories. Due to the migration effect, the uncertainties in the btag category are set to be anticorrelated with the ones in the nobtag category. The SUSY $bb\phi$ process is assigned an uncertainty of 5%, the SUSY $gg\phi$ is assigned an uncertainty of 1%, $t\bar{t}$ is assigned an uncertainty of 2% and the electroweak samples an uncertainty of 3%.

Apart from real b-jets being subject to systematic uncertainties on their identification efficiencies, light jets can be mistagged as b-jets and the fake rate in simulation can be subject to systematic errors. The MSSM Higgs boson signal sample, as well as the electroweak and $t\bar{t}$ background samples are assigned an uncertainty of 2% in the nobtag category. This is assumed to be anticorrelated with the one in the btag category, which is 5% for the MSSM $gg\phi$ process and 3% for the electroweak ones. $t\bar{t}$ and $bb\phi$ events have always at least two real b-jets (not necessarily identified), so to first order the contribution of fake b-jets in the btag category is negligible.

Missing transverse energy scale

MET enters the event selection through the D_ζ variable (see section 4.3.2). A change in its scale can affect the distribution of D_ζ, and therefore the yields of the signal categories. This effect is estimated to be 1% for $t\bar{t}$ and 2% for MSSM in the nobtag category, and equal amounts for the btag category.

Bias introduced by embedding

The embedding technique used in the $Z \to \tau\tau$ sample involves cutting away particles contained in a cone around the original $Z \to \mu\mu$ muons and simulating inside this cone the propagation of the taus substituting the muons, their decay into leptons and the resulting detector response. The selection for the original $Z \to \mu\mu$ events, which are required to have a good quality, may introduce a bias in the representation of b-tagging performance. The cone cut may also introduce a bias, as the cone ends up being exceptionally free from background.

An effect of possible bias would be to shift a $Z \to \tau\tau$ event with an ambiguously b-tagged jet between the nobtag and btag categories. Since this effect is known to be small, only a small number of events (order of few %) could be affected. The nobtag category has a $Z \to \tau\tau$ yield which is of the order of 1000 times the one in the btag category, so an event shifting from the latter to the former has negligible effect. For this reason, a nuisance parameter is considered for the btag category only, and its fluctuation is estimated to be 1%. This value has been chosen by comparing events from Monte Carlo simulation to Monte Carlo events on which the embedding technique has been performed.

Shape of the spectrum in the fakes sample

The W+Jets/QCD fakes sample is obtained by selecting pairs of same-sign leptons (see section 3.15). Modeling the non-resonant part of a sample of opposite sign leptons with a sample of same-sign leptons may result in biases. Furthermore, the use of the reweighting procedure may introduce additional biases.

In order to account for these biases a shape uncertainty is defined. The upper fluctuation template is built from the same sign mass spectrum in which the ratio is not $m_{e\mu}$-dependent (this ratio is shown in equation 3.24) has been applied. The down fluctuation template is constructed so at each bin the central template is the average between the upper and lower template. Should one bin in the up/down templates assume a negative value, its contents are set to zero.

The denomination "upper"/"lower" does not necessarily represent the fact that the value in each bin of the template is higher/lower than that for the central template. The choice of the direction assigned to them is arbitrary, and does not influence the output of the limit, except for the sign of the coefficient representing the pull introduced by the systematic uncertainty.

5.4.2. Theoretical Systematic Uncertainties

The cross-sections for the processes included in this analysis are subject to systematic uncertainties. Major contributions are:

- The correct modeling of the associated jets and how these affect the acceptance of the events in the analysis.
- The description of the colliding partons inside the protons (the *parton distribution functions*, PDF).

Process	Systematic uncertainty on sample normalization
Embedded $Z \to \tau\tau$	3%
$t\bar{t}$	10%
Electroweak	15%
W+Jets/QCD fakes	30%

Table 5.1.: Values for the systematic uncertainty on normalization for the different samples.

- The choice of the renormalization scale when computing the cross section in processes involving the strong force with one or more particle loops, such as $gg\phi$.

Additional sources of systematic uncertainties are the use of methods to determine the proper normalization of the sample, such as in the case of the $Z \to \tau\tau$ embedded sample, whose yield is normalized to that of the $Z \to \tau\tau$ Monte Carlo simulation.

The uncertainties on the event yield for different samples are listed in table 5.1. They include the uncertainties on the production cross section. These normalization uncertainties are applied in the same way for nobtag and btag categories.

5.4.3. Tail Fitting and bin-by-bin Uncertainties

The likelihood used to build the test statistic (see section 3.10) takes into account possible fluctuations of the number of events in the data. These fluctuations are modeled by a Poisson distribution.

The Monte Carlo statistics used in modeling the signal and the various backgrounds are available in limited amounts, as simulations are computationally intensive. The consequence is that in the tail of the mass distribution the number of events in each bin are small, so the relative fluctuations are large. A possible solution, employed in all $H \to \tau\tau$ analyses in CMS, is to interpolate the Monte Carlo spectrum by modeling it with a function. This function is first fitted to the tail of the spectrum and then is used to generate a smoother spectrum, with the help of a random number generator. The function used for tail fitting has the form $exp(-P(m)/Q(m))$ where $P(m)$ and $Q(m)$ are first order polynomials in the mass variable m.

Systematic uncertainties ("bin-by-bin uncertainties") on the content of the bins in the tails of the Monte Carlo histograms are implemented as nuisance parameters to the fit. This is done in order to model possible systematic errors associated with fluctuations in the initial spectrum and with the extrapolation performed in the tail fitting procedure. Initially, one nuisance parameter is used for each bin and for each background distribution. To avoid adding too many nuisance parameters to the fit, that would make it slower and less stable when converging, a selection is performed so that only the significant bin-by-bin uncertainties are considered. A bin-by-bin uncertainty is considered significant if the difference in the value of the expected limit between a fit considering it and a fit not considering it is larger than 0.5%.

5.4.4. Fit on the Nuisance Parameters

The nuisance parameters are extracted by a maximum likelihood fit, as has been described in section 3.10. The full set of parameters resulting from the fit are shown in the appendices in tables A.1, A.2, A.3 and A.4.

In the tables the extracted values for the nuisance parameters are shown for the background-only hypothesis (i.e. signal strength fixed to $\mu = 0$), for a Higgs with a mass of $m_\phi = 400$ GeV/c^2 and for a mass $m_\phi = 700$ GeV/c^2. The first part of each list contains the systematic uncertainties described in the previous paragraphs, the second part contains the bin-by-bin uncertainties. The last number is the fitted signal strength μ, which is positive if the observed limit is above the expected limit line and negative if the observed limit is below.

5.5. Results

For each signal category, the $\tau\tau$ mass distribution is fed into the statistical analysis. In the latter, the methods described in section 3.10 are applied to estimate the limit on the cross section, separately for the two SUSY Higgs production processes.

In figure 5.14, the $m_{\tau\tau}$ distribution in data is compared to the Standard Model background and SUSY Higgs boson templates, with the normalization of these templates determined by the maximum likelihood fit. In the figure, the comparison is shown for the btag (histograms at the top) and nobtag (histograms at the bottom) category for the case where the Higgs signal strength μ is fixed to zero (background-only hypothesis), shown in the left histograms, and for the case where the signal strength is left floating in the fit, shown on the right histograms. The signal strength expresses the most likely observed cross section for the production of a Higgs boson, on the hypothesis that it exists, expressed in units of production cross section expected by theory.

The histograms showing the fit with a floating signal strength correspond to the case in which the Higgs boson has a mass of $m_\phi = 700$ GeV/c^2, which was taken as an example. This procedure is repeated for all mass points used in the search, so for each mass point a value $\mu(m_A)$ is found. Each of these values, when multiplied by the production cross section that corresponds to the Higgs boson template being considered, is the observed cross section in the 1-dimensional limits. These limits will be presented below.

1-dimensional limits

The 1-dimensional limits on the cross section for the $gg\phi$ and $bb\phi$ processes are shown as a function of Higgs boson mass in figure 5.15. The limits are relative to events selected using the BDT discussed in the previous sections. They can be compared to those obtained using the selection based on the D_ζ variable in the figure just below (figure 5.16).

While small differences are visible between the limits corresponding to the two selections, the overall trend is consistent. The use of the BDT improves the expected limit

in the high mass region by around 50%.

2-dimensional limits

The values of m_A and $\tan\beta$ modify the cross section for each of the two production processes, the relative strength between the processes and the shape of the mass distribution. It is therefore possible to perform a test for each pair of $(m_A, \tan\beta)$, by determining if the observed event yields exclude that specific pair of values. The limit is run in parallel on the two signal categories and presented in the $m_A - \tan\beta$ plane as in figure 5.17, which shows the limit for the events selected with the BDT. This limit can be compared with the one obtained from the events selected with the D_ζ discriminant in figure 5.18. Again as in the case of the 1D limits the trend is similar and a slight improvement of the expected limit is visible in the BDT case.

Use of $m_{e\mu}$ to derive the limits

Alternatively, limits are extracted from $m_{e\mu}$, the dilepton mass spectrum. In the high-mass regime, this observable turns out to provide similar performance to $m_{\tau\tau}$, with the benefit of being simpler. The observable $m_{e\mu}$ is simpler than $m_{\tau\tau}$ because it does not involve a kinematic fit and does not depend on the missing transverse energy. It is therefore a valid alternative to extract the limits and to cross-check the result obtained when using $m_{\tau\tau}$.

The dilepton mass spectrum, as determined by the maximum likelihood fit, for the btag and nobtag categories is shown in figure 5.19. The corresponding 1D cross section exclusion plots are shown in figure 5.20, where they are compared with the $m_{\tau\tau}$ case. While at lower masses there is an advantage in using $m_{\tau\tau}$, as the expected limit is slightly better, at high mass the difference becomes very small. Comparing the observed limits, the $m_{e\mu}$ one in the nobtag category is consistent with Standard Model predictions, which is not the case for the already mentioned $m_{\tau\tau}$ case.

The exclusion plot in the $m_A - \tan\beta$ plane, visible in figure 5.21, confirms what is observed in the 1-dimensional limits in Fig 5.20: the sensitivity resulting from the use of $m_{e\mu}$ as probe variable for the statistical analysis is slightly inferior to the $m_{\tau\tau}$ case. This is shown by the dashed black line in the figure, representing the expected limit, positioned more to the left towards higher masses in the $m_{\tau\tau}$ case. It is important to notice that the tools and methods used in this analysis were developed assuming that $m_{\tau\tau}$ was going to be used for extracting the limits. In case the analysis is made simpler by using the $m_{e\mu}$ variable, some conservative estimations on the systematic uncertainty could be tightened, with a corresponding benefit on the limits.

These conclusions suggest that it may be interesting to attempt a high mass search for a MSSM Higgs boson using just the $e\mu$ invariant mass, as this would reduce the systematic uncertainties through simplification of the analysis. The large volume of data that is expected to be collected during the second run of the LHC is an opportunity to take this route.

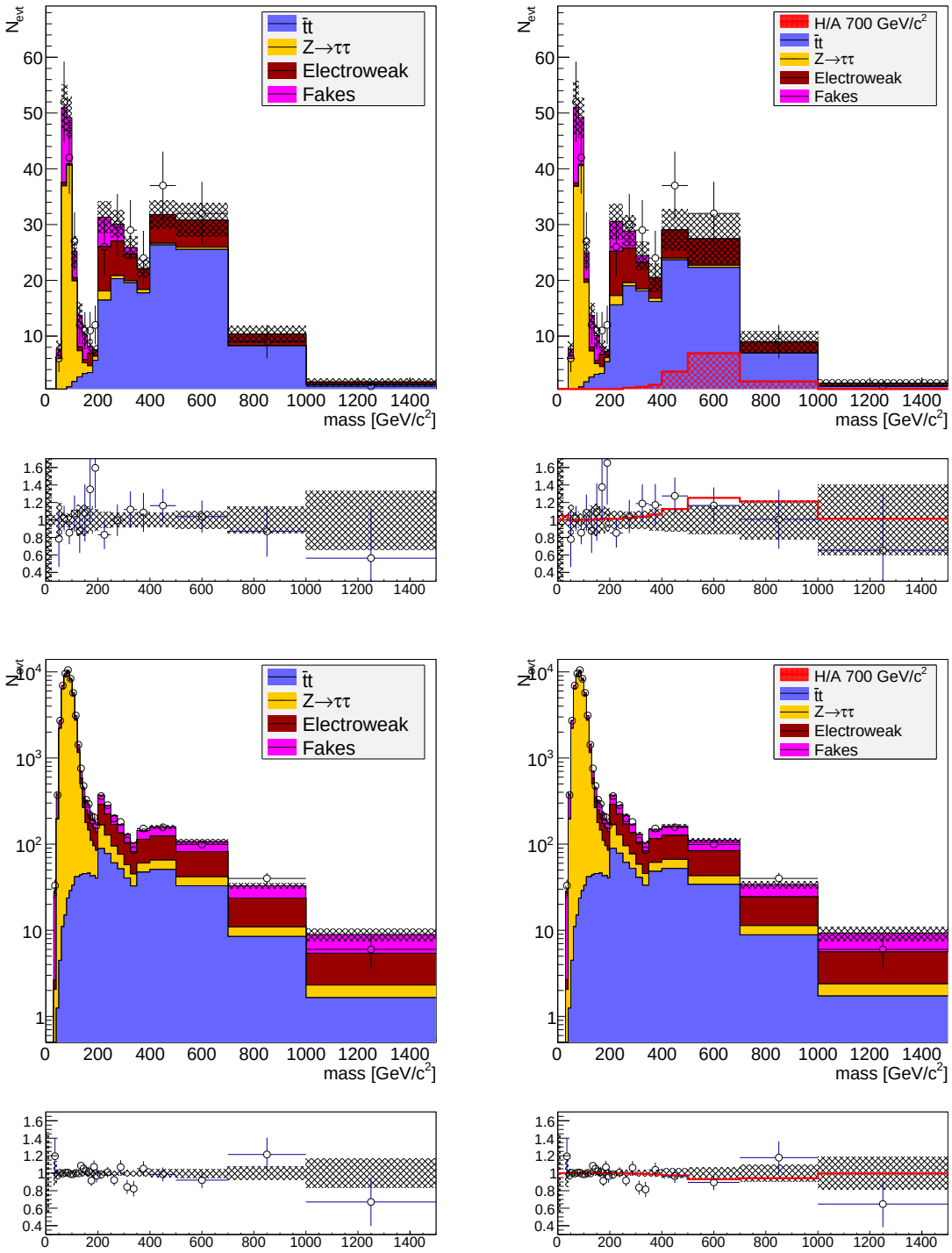


Figure 5.14.: Post-ML fit spectrum of $m_{\tau\tau}$ when selecting events using the BDTs. Top: btag category. Bottom: nobtag category. Left: background-only hypothesis. Right: $m_A = 700$ GeV/c^2 Higgs hypothesis.

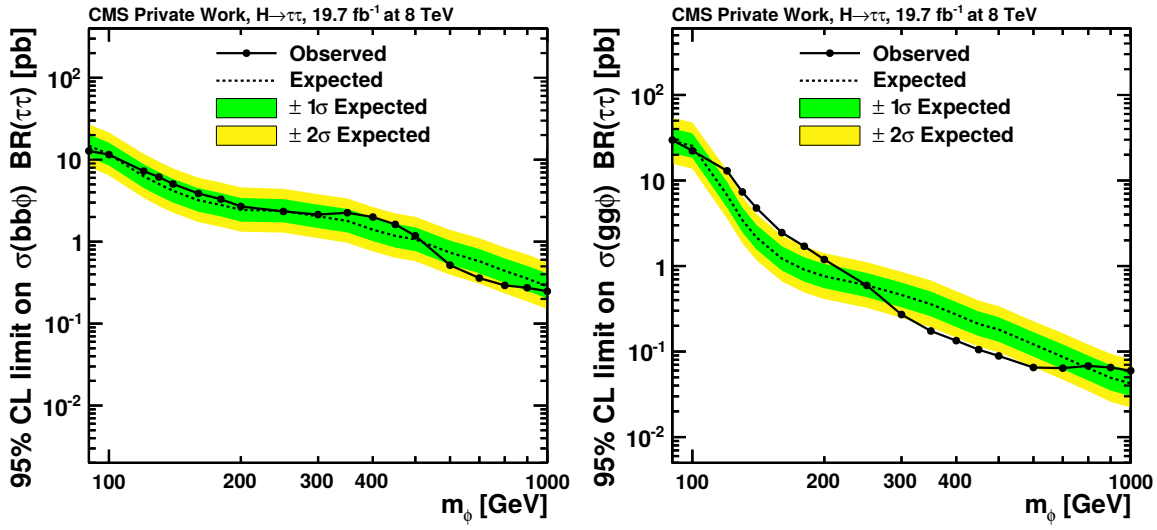


Figure 5.15.: Limits on $\sigma(pp \to \phi) \cdot BR(\phi \to \tau\tau)$ for the $bb\phi$ (left) and the $gg\phi$ (right) production processes deduced from the $m_{\tau\tau}$ spectrum with selection based on BDT.

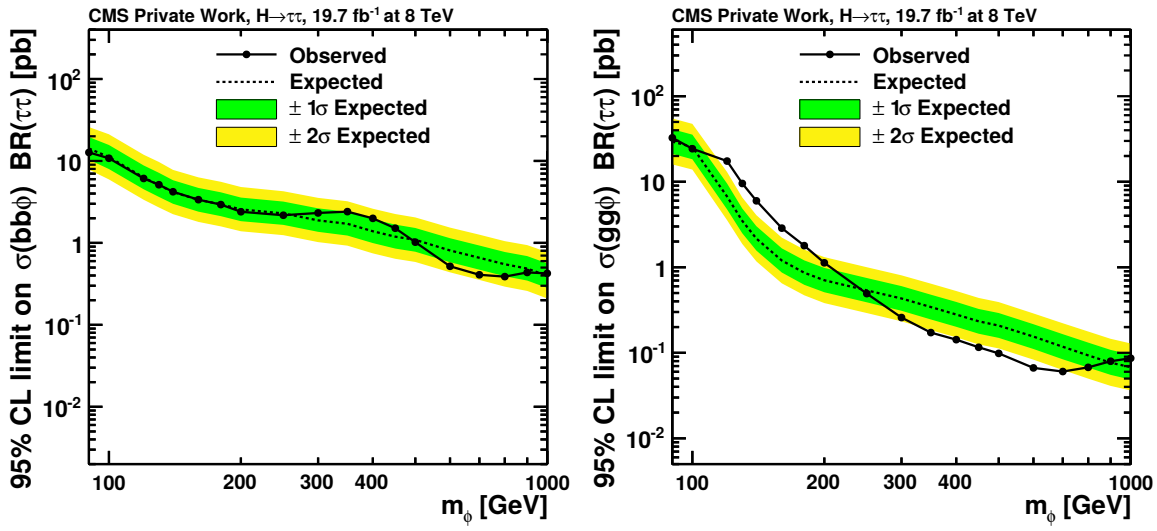


Figure 5.16.: Limits on $\sigma(pp \to \phi) \cdot BR(\phi \to \tau\tau)$ for the $bb\phi$ (left) and the $gg\phi$ (right) production processes deduced from the $m_{\tau\tau}$ spectrum with selection based on D_ζ.

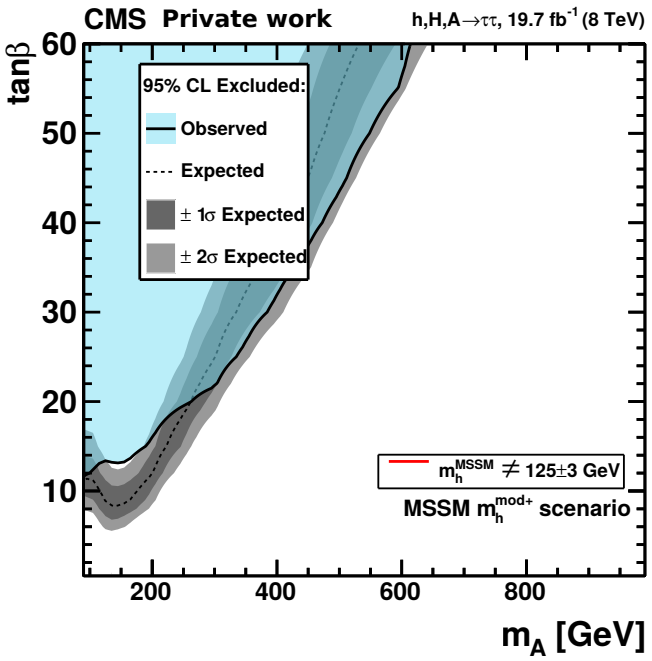


Figure 5.17.: MSSM Higgs exclusion limit in the $m_A - \tan\beta$ for events selected with the BDT discriminant.

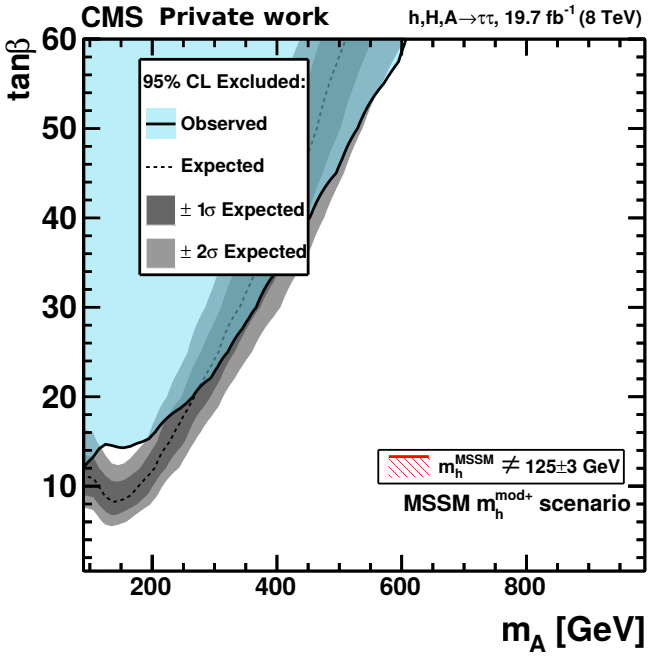


Figure 5.18.: MSSM Higgs exclusion limit in the $m_A - \tan\beta$ for events selected with the D_ζ discriminant.

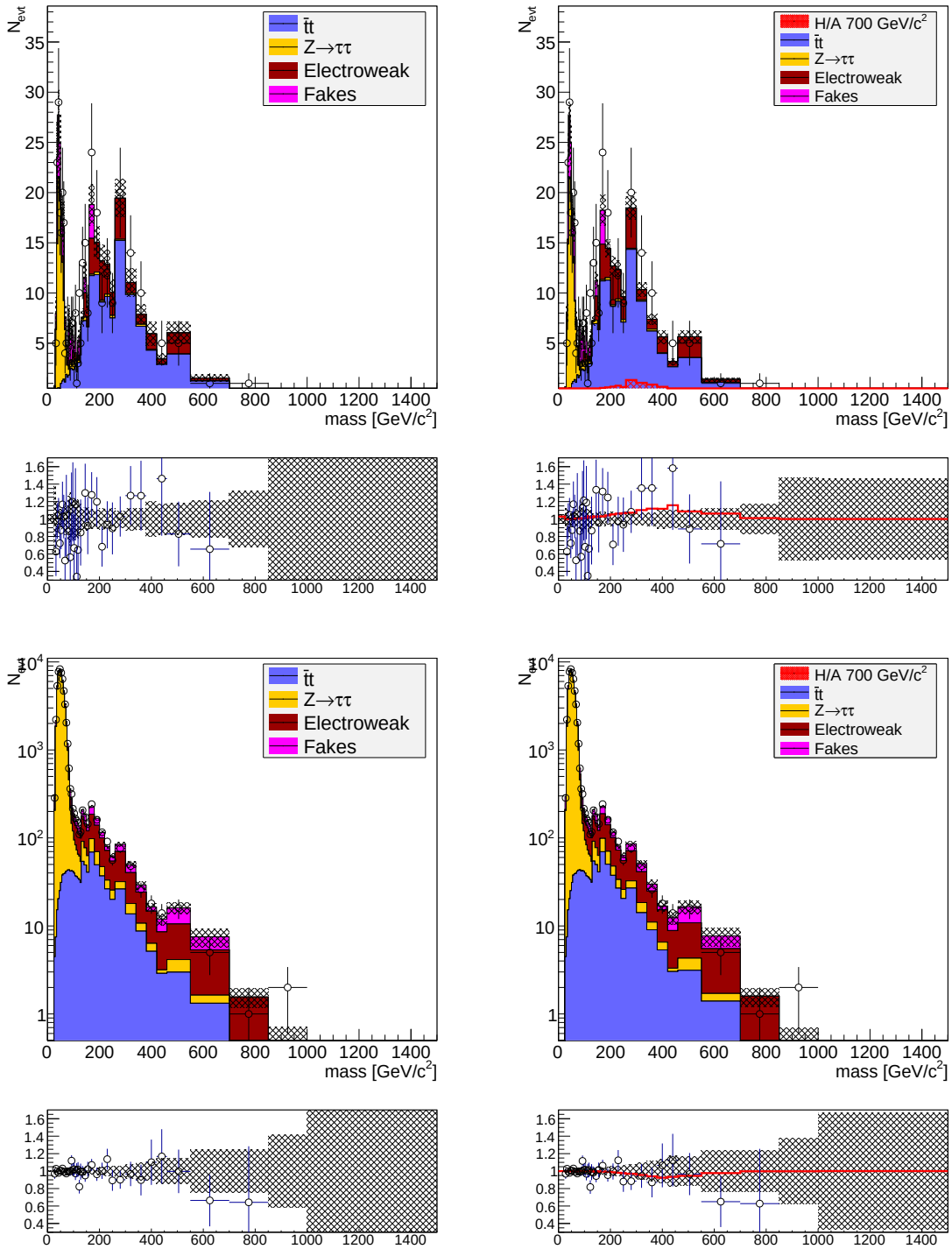


Figure 5.19.: Dilepton mass after the maximum likelihood fit. Top plots: btag category. Bottom plots: nobtag category. Left: background-only hypothesis. Right: $m_A = 700$ GeV/c^2 Higgs hypothesis.

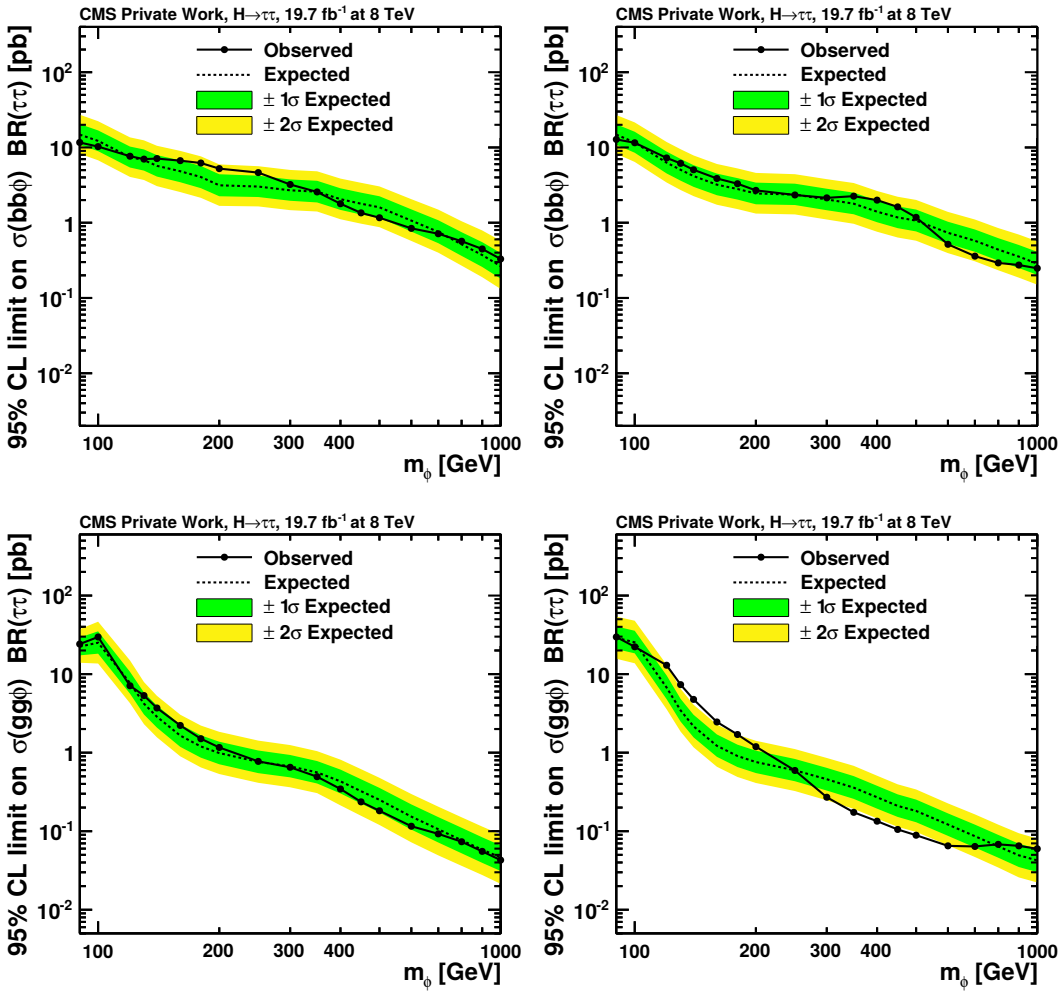


Figure 5.20.: Exclusion limit on cross section. Left: using $m_{e\mu}$ as test distribution. Right: using $m_{\tau\tau}$. Top: $bb\phi$ process. Bottom: $gg\phi$ process.

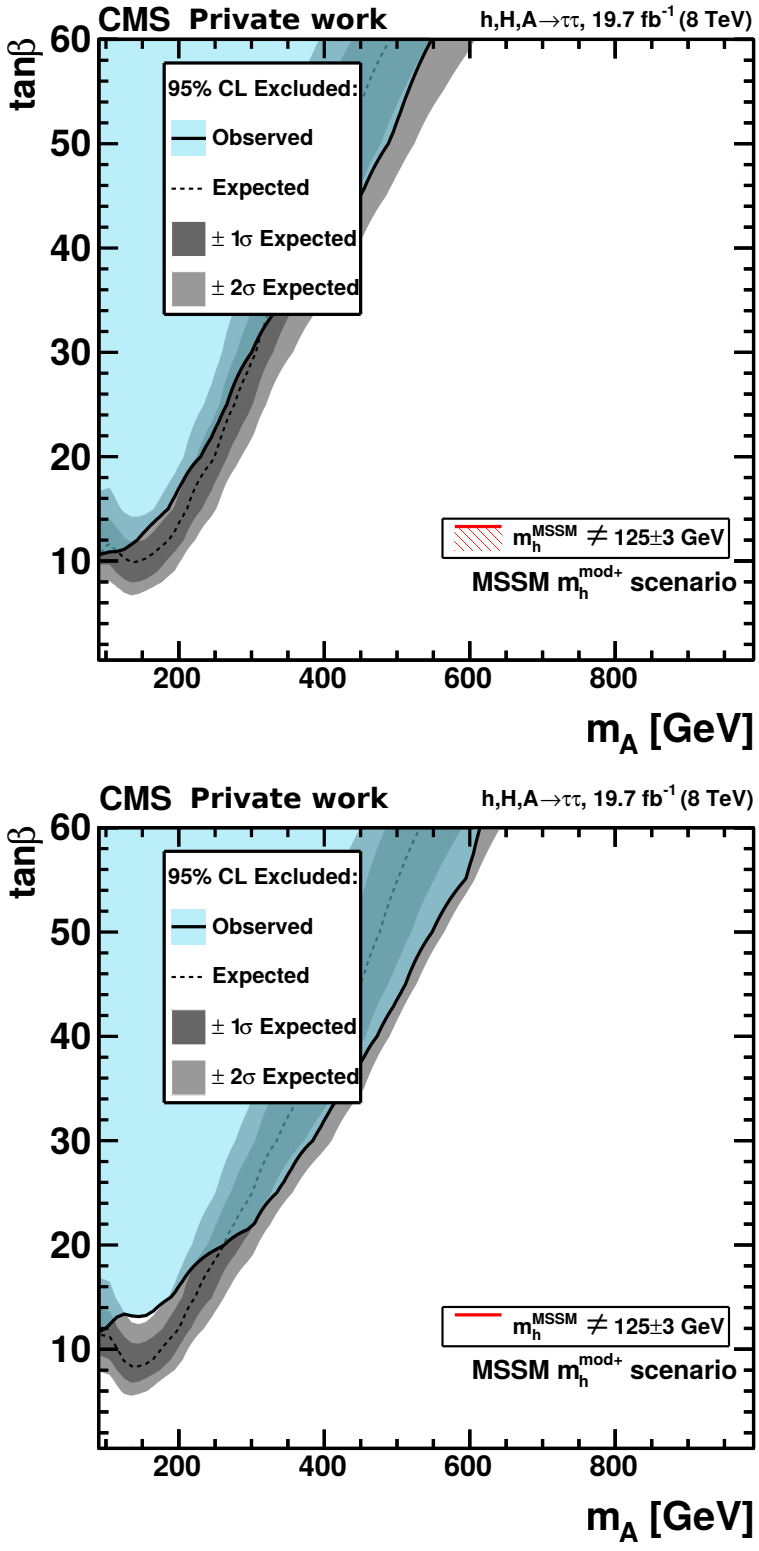


Figure 5.21.: Exclusion plot in the $m_A - tan\beta$ plane for a MSSM Higgs boson. Top: using the $m_{e\mu}$ observable. Bottom: using the $m_{\tau\tau}$ observable. Both cases refer to the BDT-based event selection.

Summary and outlook

A search for a heavy MSSM Higgs boson has been performed in the $\phi \to \tau\tau \to e\mu$ decay channel. In the mass range between 300 GeV and 1000 GeV, the results of the search are consistent with the expectations of the Standard Model of particle physics and do not indicate the existence of a MSSM Higgs boson at the current level of precision.

In the thesis, the reconstruction of events with a signature consistent with the production of a MSSM Higgs boson has been described in detail. The Higgs boson is reconstructed in the final state where the tau pair decays to an electron and a muon. One of the two leptons is required to have a $p_t > 20$ GeV/c, the other is required to have a $p_t > 10$ GeV/c. Additionally, the two production channels considered in this analysis have an event signature that is characterized by jets. The gluon fusion process has no b-tagged jets with a $p_t > 20$ GeV/c. The b-quark associated production process, instead, contains at least one b-jet with a $p_t > 20$ GeV/c and no more than one jet with $p_t > 30$ GeV/c.

The weak decay of each τ lepton into an electron or a muon produces two neutrinos which escape the experiment undetected, contributing to missing transverse energy (MET). The small rest energy of the τ, when compared to the boost in the laboratory frame, implies that the trajectory of the neutrinos and that of the charged lepton are very close in the solid angle, and that the momentum of the τ is approximately shared equally between the charged lepton and the neutrinos. Therefore, the MET and the p_T of the $e\mu$ pair are highly correlated in genuine $\tau\tau$ events, while this correlation is weak in $t\bar{t}$ and W^+W^- background. A discriminating variable D_ζ is constructed to make use of this property.

Furthermore, when discussing the identification of electrons, the training of a boosted decision tree with the purpose of efficiently identify electrons against background has been described. Then, the improved performance of the boosted decision tree has been compared with a previously-used cut-based selection. The combination of many variables into a boosted decision tree allows significant gains in electron identification efficiency, while keeping the level of background under control.

Furthermore, a set of new discriminating variables has been studied and their potential to discriminate signal and background events has been evaluated. One of the variables is κ, which is related to the opening angle and the p_t of the leptons and capable to discriminate $t\bar{t}$ background. Another variable is the distance of closest approach (2D DCA) of the electron and muon tracks, which is sensitive to the decay length of the τ leptons. These variables have then been combined with other kinematic variables such as the p_T of the leptons and that of the jets, using a boosted decision tree based on the TMVA package to produce an efficient global discriminating variable. The boosted decision tree has been trained using the gradient boost method. The output variable of the boosted decision tree has then been used to enhance the signal-to-background ratio in the selected data sample and the new selection has been shown to improve the limits on the production cross-section for a MSSM Higgs boson and on the allowed

parameters of the Higgs sector of this model.

The use of the invariant mass between the electron and the muon, $m_{e\mu}$, as the observable to use for statistical analysis has been investigated as an alternative to the $\tau\tau$ invariant mass, $m_{\tau\tau}$, normally used in Higgs $\phi \to \tau\tau$ searches. This variable has been shown to produce similar results, with a slight penalty in resolution, but with a reduced dependence on systematic uncertainties. The observable $m_{e\mu}$ is an attractive option for future searches of Higgs bosons at high mass at the LHC.

Ionizing radiation can affect the proper working of the CMS pixel detector, including the alteration of the logic state of the read-out chips. Such events are denominated single event upsets and can render the chip temporarily inefficient. The development of a software used for the detection of single event upsets in the read-out chips of the CMS pixel detector has been described in this thesis and the output of the software and its possible application in detector studies have been discussed briefly. This software is now in use in CMS and the analysis results helped to define the design of the read-out chip for the pixel detector upgrade.

At the end of the Spring of 2015, the Large Hadron Collider restarted its operations after a two year shutdown. This shutdown allowed the installation of improvements and consolidations to the machine, for operation at a center of mass energy of 13 TeV. The second run of the LHC (2015-2018) is expected to provide an integrated luminosity of the order of 100 fb^{-1}. The doubling of the center of mass energy of the proton-proton system will allow to significantly extend the range of particle searches into the multi-TeV region, allowing the discovery – or the exclusion – of a wide variety of models and associated particles. There is a reasonable belief that, if supersymmetry exists, some of the particles of the supersymmetric sector will be discovered in a few years of data taking at the LHC.

Appendices

A. Tables of Nuisance Parameters

Nuisance parameter	b only	b+s @ $m_\phi = 400 GeV/c^2$ (Correlation with μ)	b+s @ $m_\phi = 700 GeV/c^2$ (Correlation with μ)
CMS_eff_b_8TeV	$+0.02 \pm 0.98\sigma$	$-0.05 \pm 0.98\sigma$ (-0.12)	$+0.08 \pm 0.98\sigma$ $(+0.03)$
CMS_eff_e	$-0.03 \pm 0.95\sigma$	$-0.05 \pm 0.95\sigma$ (-0.05)	$+0.02 \pm 0.95\sigma$ (-0.00)
CMS_eff_m	$-0.03 \pm 0.95\sigma$	$-0.05 \pm 0.95\sigma$ (-0.05)	$+0.02 \pm 0.95\sigma$ (-0.00)
CMS_fake_b_8TeV	$-0.02 \pm 0.99\sigma$	$-0.04 \pm 0.99\sigma$ (-0.02)	$-0.00 \pm 0.99\sigma$ (-0.01)
CMS_htt_DiBosonNorm_8TeV	$-0.10 \pm 0.96\sigma$	$-0.18 \pm 0.96\sigma$ (-0.07)	$-0.01 \pm 0.97\sigma$ (-0.07)
CMS_htt_FakeShape_em_btag_8TeV	$-0.45 \pm 0.79\sigma$	$-0.39 \pm 0.78\sigma$ $(+0.08)$	$-0.52 \pm 0.82\sigma$ $(+0.04)$
CMS_htt_TTbarShape_em_btag_8TeV	$+0.06 \pm 0.80\sigma$	$+0.23 \pm 0.83\sigma$ $(+0.17)$	$-0.69 \pm 0.89\sigma$ $(+0.53)$
CMS_htt_extrap_ztt_em_btag_8TeV	$-0.01 \pm 0.99\sigma$	$+0.01 \pm 0.99\sigma$ $(+0.02)$	$-0.02 \pm 0.99\sigma$ $(+0.00)$
CMS_htt_fakes_em_8TeV	$-0.40 \pm 0.88\sigma$	$-0.33 \pm 0.88\sigma$ $(+0.07)$	$-0.48 \pm 0.88\sigma$ $(+0.05)$
CMS_htt_fakes_em_btag_8TeV	$-0.13 \pm 0.98\sigma$	$-0.11 \pm 0.98\sigma$ $(+0.02)$	$-0.16 \pm 0.98\sigma$ $(+0.02)$
CMS_htt_scale_met_8TeV	$+0.02 \pm 0.99\sigma$	$-0.01 \pm 0.99\sigma$ (-0.00)	$+0.04 \pm 0.99\sigma$ (-0.05)
CMS_htt_ttbarNorm_8TeV	$+0.21 \pm 0.75\sigma$	$-0.09 \pm 0.80\sigma$ (-0.32)	$+0.41 \pm 0.76\sigma$ (-0.18)
CMS_htt_ttbar_emb_8TeV	$+0.04 \pm 0.99\sigma$	$-0.02 \pm 0.99\sigma$ (-0.05)	$+0.08 \pm 0.99\sigma$ (-0.03)
CMS_htt_zttNorm_8TeV	$-0.04 \pm 0.93\sigma$	$+0.02 \pm 0.93\sigma$ $(+0.06)$	$-0.05 \pm 0.93\sigma$ $(+0.00)$
CMS_scale_e_8TeV	$+0.21 \pm 0.89\sigma$	$+0.22 \pm 0.85\sigma$ $(+0.01)$	$+0.12 \pm 0.95\sigma$ $(+0.06)$
CMS_scale_j_8TeV	$-0.13 \pm 0.83\sigma$	$+0.11 \pm 0.86\sigma$ $(+0.23)$	$-0.32 \pm 0.83\sigma$ $(+0.16)$
CMS_shift1_emu_btag_8TeV_EWK_fine_binning	$+0.36 \pm 0.93\sigma$	$+0.47 \pm 0.94\sigma$ $(+0.10)$	$+0.04 \pm 0.95\sigma$ $(+0.21)$
CMS_shift2_emu_btag_8TeV_EWK_fine_binning	$-0.04 \pm 0.96\sigma$	$-0.28 \pm 0.97\sigma$ (-0.21)	$-0.30 \pm 0.94\sigma$ $(+0.16)$
lumi_8TeV	$-0.02 \pm 0.99\sigma$	$-0.03 \pm 0.99\sigma$ (-0.04)	$-0.00 \pm 0.99\sigma$ $(+0.03)$
CMS_htt_em_emu_btag_8TeV_EWK_bin_1	$-0.58 \pm 1.00\sigma$	$-0.57 \pm 1.00\sigma$ $(+0.01)$	$-0.60 \pm 1.00\sigma$ $(+0.01)$
CMS_htt_em_emu_btag_8TeV_EWK_bin_2	$+0.01 \pm 0.99\sigma$	$+0.01 \pm 0.99\sigma$ $(+0.00)$	$+0.01 \pm 0.99\sigma$ $(+0.00)$
CMS_htt_em_emu_btag_8TeV_EWK_bin_3	$-0.11 \pm 0.99\sigma$	$-0.11 \pm 0.99\sigma$ $(+0.01)$	$-0.12 \pm 0.99\sigma$ $(+0.00)$
CMS_htt_em_emu_btag_8TeV_EWK_bin_4	$+0.01 \pm 0.99\sigma$	$+0.02 \pm 0.99\sigma$ $(+0.00)$	$+0.01 \pm 0.99\sigma$ $(+0.00)$
CMS_htt_em_emu_btag_8TeV_EWK_bin_5	$+0.07 \pm 0.99\sigma$	$+0.07 \pm 0.99\sigma$ $(+0.01)$	$+0.06 \pm 0.99\sigma$ $(+0.00)$
CMS_htt_em_emu_btag_8TeV_EWK_bin_6	$-0.12 \pm 0.98\sigma$	$-0.09 \pm 0.98\sigma$ $(+0.03)$	$-0.13 \pm 0.98\sigma$ $(+0.01)$
CMS_htt_em_emu_btag_8TeV_EWK_bin_7	$+0.14 \pm 0.98\sigma$	$+0.17 \pm 0.98\sigma$ $(+0.02)$	$+0.12 \pm 0.98\sigma$ $(+0.01)$
CMS_htt_em_emu_btag_8TeV_EWK_bin_8	$-0.03 \pm 0.97\sigma$	$+0.01 \pm 0.97\sigma$ $(+0.03)$	$-0.05 \pm 0.97\sigma$ $(+0.02)$
CMS_htt_em_emu_btag_8TeV_EWK_bin_9	$-0.01 \pm 0.97\sigma$	$+0.04 \pm 0.97\sigma$ $(+0.04)$	$-0.05 \pm 0.97\sigma$ $(+0.03)$
CMS_htt_em_emu_btag_8TeV_EWK_bin_10	$-0.02 \pm 0.96\sigma$	$+0.01 \pm 0.96\sigma$ $(+0.03)$	$-0.07 \pm 0.96\sigma$ $(+0.03)$
CMS_htt_em_emu_btag_8TeV_Fakes_bin_3	$-0.54 \pm 0.94\sigma$	$-0.56 \pm 0.93\sigma$ (-0.03)	$-0.50 \pm 0.94\sigma$ (-0.02)
CMS_htt_em_emu_btag_8TeV_Fakes_bin_4	$+0.11 \pm 0.95\sigma$	$+0.10 \pm 0.95\sigma$ (-0.01)	$+0.12 \pm 0.95\sigma$ (-0.01)
CMS_htt_em_emu_btag_8TeV_Fakes_bin_5	$+0.21 \pm 0.95\sigma$	$+0.22 \pm 0.95\sigma$ $(+0.01)$	$+0.22 \pm 0.96\sigma$ (-0.01)
CMS_htt_em_emu_btag_8TeV_Fakes_bin_6	$-0.13 \pm 0.98\sigma$	$-0.11 \pm 0.98\sigma$ $(+0.02)$	$-0.12 \pm 0.98\sigma$ (-0.01)
CMS_htt_em_emu_btag_8TeV_Fakes_bin_7	$+0.22 \pm 0.96\sigma$	$+0.26 \pm 0.95\sigma$ $(+0.04)$	$+0.20 \pm 0.96\sigma$ $(+0.01)$
CMS_htt_em_emu_btag_8TeV_Fakes_bin_8	$-0.00 \pm 0.96\sigma$	$+0.03 \pm 0.95\sigma$ $(+0.03)$	$-0.01 \pm 0.96\sigma$ $(+0.01)$
CMS_htt_em_emu_btag_8TeV_Fakes_bin_9	$+0.01 \pm 0.98\sigma$	$+0.04 \pm 0.97\sigma$ $(+0.03)$	$-0.01 \pm 0.98\sigma$ $(+0.01)$
CMS_htt_em_emu_btag_8TeV_Fakes_bin_10	$+0.00 \pm 0.99\sigma$	$+0.01 \pm 0.99\sigma$ $(+0.01)$	$-0.01 \pm 0.99\sigma$ $(+0.01)$
CMS_htt_em_emu_btag_8TeV_Fakes_bin_11	$-0.05 \pm 0.98\sigma$	$-0.07 \pm 0.98\sigma$ (-0.02)	$-0.08 \pm 0.98\sigma$ $(+0.02)$
CMS_htt_em_emu_btag_8TeV_Fakes_bin_12	$+0.09 \pm 0.99\sigma$	$+0.05 \pm 0.99\sigma$ (-0.04)	$+0.06 \pm 0.99\sigma$ $(+0.02)$
CMS_htt_em_emu_btag_8TeV_Fakes_bin_13	$+0.16 \pm 0.98\sigma$	$+0.08 \pm 0.99\sigma$ (-0.07)	$+0.13 \pm 0.98\sigma$ $(+0.02)$
CMS_htt_em_emu_btag_8TeV_Fakes_bin_15	$+0.01 \pm 0.98\sigma$	$-0.01 \pm 0.98\sigma$ (-0.02)	$+0.08 \pm 0.98\sigma$ (-0.04)
CMS_htt_em_emu_btag_8TeV_Ztt_bin_2	$+0.05 \pm 0.99\sigma$	$+0.06 \pm 0.99\sigma$ $(+0.00)$	$+0.05 \pm 0.99\sigma$ $(+0.00)$
CMS_htt_em_emu_btag_8TeV_Ztt_bin_8	$-0.01 \pm 0.99\sigma$	$-0.00 \pm 0.99\sigma$ $(+0.00)$	$-0.01 \pm 0.99\sigma$ $(+0.00)$
CMS_htt_em_emu_btag_8TeV_Ztt_bin_9	$-0.00 \pm 0.99\sigma$	$+0.00 \pm 0.99\sigma$ $(+0.00)$	$-0.01 \pm 0.99\sigma$ $(+0.00)$
CMS_htt_em_emu_btag_8TeV_Ztt_bin_10	$-0.00 \pm 0.99\sigma$	$-0.00 \pm 0.99\sigma$ $(+0.00)$	$-0.01 \pm 0.99\sigma$ $(+0.00)$
CMS_htt_em_emu_btag_8TeV_Ztt_bin_11	$-0.02 \pm 0.99\sigma$	$-0.02 \pm 0.99\sigma$ (-0.00)	$-0.02 \pm 0.99\sigma$ $(+0.01)$
CMS_htt_em_emu_btag_8TeV_Ztt_bin_12	$+0.02 \pm 0.99\sigma$	$+0.01 \pm 0.99\sigma$ (-0.01)	$+0.01 \pm 0.99\sigma$ $(+0.00)$
CMS_htt_em_emu_btag_8TeV_Ztt_bin_13	$+0.03 \pm 0.99\sigma$	$+0.01 \pm 0.99\sigma$ (-0.01)	$+0.02 \pm 0.99\sigma$ $(+0.00)$
CMS_htt_em_emu_btag_8TeV_Ztt_bin_14	$+0.01 \pm 0.99\sigma$	$-0.01 \pm 0.99\sigma$ (-0.01)	$+0.01 \pm 0.99\sigma$ $(+0.00)$
CMS_htt_em_emu_btag_8TeV_Ztt_bin_15	$-0.00 \pm 0.99\sigma$	$-0.00 \pm 0.99\sigma$ (-0.00)	$+0.01 \pm 0.99\sigma$ (-0.01)
CMS_htt_em_emu_btag_8TeV_Ztt_bin_16	$-0.04 \pm 0.99\sigma$	$-0.04 \pm 0.99\sigma$ $(+0.01)$	$-0.02 \pm 0.99\sigma$ (-0.02)
CMS_htt_em_emu_btag_8TeV_Ztt_bin_17	$+0.02 \pm 0.99\sigma$	$+0.02 \pm 0.99\sigma$ $(+0.00)$	$+0.02 \pm 0.99\sigma$ (-0.00)
CMS_htt_em_emu_btag_8TeV_Ztt_bin_18	$+0.02 \pm 0.99\sigma$	$+0.03 \pm 0.99\sigma$ $(+0.00)$	$+0.02 \pm 0.99\sigma$ $(+0.01)$
CMS_htt_em_emu_btag_8TeV_ttbar_bin_1	$-0.13 \pm 0.99\sigma$	$-0.12 \pm 0.99\sigma$ $(+0.01)$	$-0.13 \pm 0.99\sigma$ $(+0.00)$
CMS_htt_em_emu_btag_8TeV_ttbar_bin_2	$+0.18 \pm 0.96\sigma$	$+0.18 \pm 0.96\sigma$ $(+0.00)$	$+0.18 \pm 0.96\sigma$ (-0.00)
CMS_htt_em_emu_btag_8TeV_ttbar_bin_3	$-0.10 \pm 0.99\sigma$	$-0.09 \pm 0.99\sigma$ $(+0.01)$	$-0.11 \pm 0.99\sigma$ $(+0.00)$
μ	(μ fixed to 0)	$+0.78 \pm 0.69$	-0.56 ± 0.34

Table A.1.: The output of the maximum likelihood fit for the nuisance parameters θ_i and signal strength μ on events selected with D_ζ in the **btag** category.

Nuisance parameter	b only	b+s @ $m_\phi = 400 GeV/c^2$ (Correlation with μ)	b+s @ $m_\phi = 700 GeV/c^2$ (Correlation with μ)
CMS_eff_b_8TeV	$+0.86 \pm 0.93\sigma$	$+0.53 \pm 0.94\sigma(+0.13)$	$+0.68 \pm 0.93\sigma(+0.06)$
CMS_eff_e	$-0.28 \pm 0.87\sigma$	$-0.20 \pm 0.87\sigma(-0.02)$	$-0.24 \pm 0.87\sigma(-0.00)$
CMS_eff_m	$-0.28 \pm 0.87\sigma$	$-0.20 \pm 0.87\sigma(-0.02)$	$-0.24 \pm 0.87\sigma(-0.00)$
CMS_fake_b_8TeV	$+0.44 \pm 0.98\sigma$	$+0.27 \pm 0.98\sigma(+0.04)$	$+0.34 \pm 0.98\sigma(+0.01)$
CMS_htt_DiBosonNorm_8TeV	$-1.11 \pm 0.74\sigma$	$-0.70 \pm 0.77\sigma(-0.27)$	$-0.87 \pm 0.75\sigma(-0.17)$
CMS_htt_FakeShape_em_nobtag_8TeV	$+0.40 \pm 0.33\sigma$	$+0.32 \pm 0.34\sigma(+0.11)$	$+0.48 \pm 0.32\sigma(-0.14)$
CMS_htt_TTbarShape_em_nobtag_8TeV	$+1.04 \pm 0.88\sigma$	$+0.85 \pm 0.89\sigma(+0.12)$	$+0.22 \pm 0.97\sigma(+0.42)$
CMS_htt_fakes_em_8TeV	$+0.93 \pm 0.44\sigma$	$+0.95 \pm 0.44\sigma(-0.03)$	$+0.83 \pm 0.45\sigma(+0.10)$
CMS_htt_scale_met_8TeV	$+0.14 \pm 0.99\sigma$	$+0.08 \pm 0.99\sigma(-0.02)$	$+0.11 \pm 0.99\sigma(-0.03)$
CMS_htt_ttbarNorm_8TeV	$-1.30 \pm 0.84\sigma$	$-0.80 \pm 0.87\sigma(-0.28)$	$-1.02 \pm 0.84\sigma(-0.18)$
CMS_htt_zttNorm_8TeV	$+0.11 \pm 0.68\sigma$	$-0.01 \pm 0.68\sigma(+0.05)$	$+0.05 \pm 0.68\sigma(+0.01)$
CMS_scale_e_8TeV	$+0.36 \pm 0.15\sigma$	$+0.32 \pm 0.16\sigma(+0.11)$	$+0.33 \pm 0.16\sigma(+0.07)$
CMS_scale_j_8TeV	$+0.22 \pm 0.99\sigma$	$+0.13 \pm 0.99\sigma(+0.06)$	$+0.17 \pm 0.99\sigma(+0.05)$
CMS_shift1_emu_nobtag_8TeV_EWK_fine_binning	$+0.57 \pm 0.91\sigma$	$+0.23 \pm 0.92\sigma(+0.18)$	$+0.24 \pm 0.95\sigma(+0.17)$
CMS_shift1_emu_nobtag_8TeV_Fakes_fine_binning	$+0.59 \pm 0.90\sigma$	$+0.13 \pm 0.94\sigma(+0.21)$	$+0.33 \pm 0.96\sigma(+0.13)$
CMS_shift2_emu_nobtag_8TeV_EWK_fine_binning	$-0.46 \pm 0.92\sigma$	$-0.16 \pm 0.92\sigma(-0.15)$	$-0.91 \pm 0.94\sigma(+0.23)$
CMS_shift2_emu_nobtag_8TeV_Fakes_fine_binning	$-0.38 \pm 0.93\sigma$	$-0.13 \pm 0.93\sigma(-0.12)$	$-0.82 \pm 0.94\sigma(+0.21)$
lumi_8TeV	$-0.20 \pm 0.99\sigma$	$-0.13 \pm 0.99\sigma(+0.02)$	$-0.16 \pm 0.99\sigma(+0.03)$
CMS_htt_em_emu_nobtag_8TeV_EWK_bin_1	$-0.10 \pm 0.99\sigma$	$-0.11 \pm 0.99\sigma(+0.00)$	$-0.11 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_EWK_bin_3	$-0.05 \pm 0.99\sigma$	$-0.06 \pm 0.99\sigma(+0.00)$	$-0.06 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_EWK_bin_4	$+0.02 \pm 0.99\sigma$	$+0.02 \pm 0.99\sigma(-0.00)$	$+0.02 \pm 0.99\sigma(-0.00)$
CMS_htt_em_emu_nobtag_8TeV_EWK_bin_5	$-0.01 \pm 0.99\sigma$	$-0.01 \pm 0.99\sigma(+0.00)$	$-0.01 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_4	$+0.61 \pm 0.66\sigma$	$+0.61 \pm 0.67\sigma(+0.00)$	$+0.61 \pm 0.67\sigma(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_5	$-0.29 \pm 0.82\sigma$	$-0.28 \pm 0.82\sigma(+0.00)$	$-0.27 \pm 0.82\sigma(-0.00)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_6	$+0.12 \pm 0.90\sigma$	$+0.11 \pm 0.90\sigma(+0.01)$	$+0.12 \pm 0.90\sigma(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_7	$-0.13 \pm 0.94\sigma$	$-0.15 \pm 0.94\sigma(+0.01)$	$-0.14 \pm 0.94\sigma(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_8	$+0.12 \pm 0.96\sigma$	$+0.10 \pm 0.96\sigma(+0.01)$	$+0.12 \pm 0.96\sigma(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_9	$+0.24 \pm 0.96\sigma$	$+0.23 \pm 0.96\sigma(+0.00)$	$+0.24 \pm 0.96\sigma(-0.00)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_10	$-0.25 \pm 0.96\sigma$	$-0.25 \pm 0.96\sigma(-0.00)$	$-0.24 \pm 0.96\sigma(-0.01)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_11	$-0.27 \pm 0.96\sigma$	$-0.25 \pm 0.96\sigma(-0.01)$	$-0.25 \pm 0.96\sigma(-0.01)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_12	$-0.00 \pm 0.94\sigma$	$+0.00 \pm 0.95\sigma(-0.00)$	$+0.02 \pm 0.94\sigma(-0.01)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_13	$-0.15 \pm 0.91\sigma$	$-0.18 \pm 0.92\sigma(+0.02)$	$-0.15 \pm 0.91\sigma(-0.00)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_14	$+1.34 \pm 0.89\sigma$	$+1.25 \pm 0.89\sigma(+0.05)$	$+1.31 \pm 0.89\sigma(+0.01)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_15	$+0.97 \pm 0.85\sigma$	$+0.84 \pm 0.86\sigma(+0.07)$	$+0.91 \pm 0.85\sigma(+0.03)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_16	$+0.56 \pm 0.83\sigma$	$+0.40 \pm 0.84\sigma(+0.09)$	$+0.48 \pm 0.84\sigma(+0.04)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_17	$+0.09 \pm 0.83\sigma$	$-0.06 \pm 0.84\sigma(+0.08)$	$-0.00 \pm 0.84\sigma(+0.05)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_18	$-0.60 \pm 0.83\sigma$	$-0.72 \pm 0.84\sigma(+0.07)$	$-0.68 \pm 0.83\sigma(+0.04)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_19	$+0.48 \pm 0.87\sigma$	$+0.35 \pm 0.87\sigma(+0.07)$	$+0.38 \pm 0.87\sigma(+0.05)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_20	$-0.06 \pm 0.86\sigma$	$-0.15 \pm 0.87\sigma(+0.05)$	$-0.15 \pm 0.87\sigma(+0.05)$
CMS_htt_em_emu_nobtag_8TeV_Ztt_bin_1	$-0.04 \pm 0.99\sigma$	$-0.04 \pm 0.99\sigma(-0.00)$	$-0.04 \pm 0.99\sigma(-0.00)$
CMS_htt_em_emu_nobtag_8TeV_Ztt_bin_4	$+0.02 \pm 0.99\sigma$	$+0.02 \pm 0.99\sigma(+0.00)$	$+0.02 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_Ztt_bin_24	$+0.11 \pm 0.99\sigma$	$+0.16 \pm 0.99\sigma(-0.02)$	$+0.09 \pm 0.99\sigma(+0.01)$
CMS_htt_em_emu_nobtag_8TeV_Ztt_bin_25	$-0.17 \pm 0.99\sigma$	$-0.13 \pm 0.99\sigma(-0.02)$	$-0.18 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_Ztt_bin_26	$-0.17 \pm 0.99\sigma$	$-0.11 \pm 0.99\sigma(-0.03)$	$-0.18 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_Ztt_bin_27	$+0.05 \pm 0.99\sigma$	$+0.13 \pm 0.99\sigma(-0.04)$	$+0.05 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_Ztt_bin_28	$-0.03 \pm 0.99\sigma$	$-0.01 \pm 0.99\sigma(-0.01)$	$-0.00 \pm 0.99\sigma(-0.01)$
CMS_htt_em_emu_nobtag_8TeV_Ztt_bin_29	$-0.13 \pm 0.99\sigma$	$-0.17 \pm 0.99\sigma(+0.02)$	$-0.02 \pm 0.99\sigma(-0.05)$
CMS_htt_em_emu_nobtag_8TeV_Ztt_bin_30	$+0.08 \pm 0.99\sigma$	$+0.05 \pm 0.99\sigma(+0.01)$	$+0.10 \pm 0.99\sigma(-0.01)$
CMS_htt_em_emu_nobtag_8TeV_Ztt_bin_31	$+0.02 \pm 0.99\sigma$	$+0.02 \pm 0.99\sigma(+0.00)$	$+0.01 \pm 0.99\sigma(+0.01)$
CMS_htt_em_emu_nobtag_8TeV_ttbar_bin_1	$-0.10 \pm 0.99\sigma$	$-0.11 \pm 0.99\sigma(+0.00)$	$-0.11 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_ttbar_bin_3	$-0.04 \pm 0.99\sigma$	$-0.05 \pm 0.99\sigma(+0.00)$	$-0.04 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_ttbar_bin_4	$+0.02 \pm 0.99\sigma$	$+0.02 \pm 0.99\sigma(-0.00)$	$+0.02 \pm 0.99\sigma(-0.00)$
CMS_htt_em_emu_nobtag_8TeV_ttbar_bin_5	$-0.00 \pm 0.99\sigma$	$-0.01 \pm 0.99\sigma(+0.00)$	$-0.00 \pm 0.99\sigma(+0.00)$
μ	(μ fixed to 0)	-0.30 ± 0.14	-0.13 ± 0.06

Table A.2.: The output of the maximum likelihood fit for the nuisance parameters θ_i and signal strength μ on events selected with D_ζ in the **nobtag** category.

Nuisance parameter	b only	b+s @ $m_\phi = 400 GeV/c^2$ (Correlation with μ)	b+s @ $m_\phi = 700 GeV/c^2$ (Correlation with μ)
CMS_eff_b_8TeV	$+0.03 \pm 0.98\sigma$	$-0.04 \pm 0.98\sigma(-0.12)$	$+0.07 \pm 0.98\sigma(+0.03)$
CMS_eff_e	$-0.04 \pm 0.95\sigma$	$-0.08 \pm 0.95\sigma(-0.06)$	$-0.00 \pm 0.95\sigma(-0.00)$
CMS_eff_m	$-0.04 \pm 0.95\sigma$	$-0.08 \pm 0.95\sigma(-0.06)$	$-0.00 \pm 0.95\sigma(-0.00)$
CMS_fake_b_8TeV	$-0.02 \pm 0.99\sigma$	$-0.04 \pm 0.99\sigma(-0.02)$	$+0.01 \pm 0.99\sigma(-0.01)$
CMS_htt_DiBosonNorm_8TeV	$-0.09 \pm 0.95\sigma$	$-0.16 \pm 0.95\sigma(-0.07)$	$+0.05 \pm 0.97\sigma(-0.10)$
CMS_htt_FakeShape_em_btag_8TeV	$-0.31 \pm 0.76\sigma$	$-0.28 \pm 0.76\sigma(+0.05)$	$-0.32 \pm 0.75\sigma(+0.00)$
CMS_htt_TTbarShape_em_btag_8TeV	$+0.44 \pm 0.68\sigma$	$+0.64 \pm 0.74\sigma(+0.25)$	$-0.21 \pm 0.84\sigma(+0.64)$
CMS_htt_extrap_ztt_em_btag_8TeV	$-0.03 \pm 0.99\sigma$	$-0.02 \pm 0.99\sigma(+0.01)$	$-0.03 \pm 0.99\sigma(-0.00)$
CMS_htt_fakes_em_8TeV	$-0.21 \pm 0.87\sigma$	$-0.18 \pm 0.87\sigma(+0.04)$	$-0.25 \pm 0.86\sigma(+0.02)$
CMS_htt_fakes_em_btag_8TeV	$-0.07 \pm 0.98\sigma$	$-0.06 \pm 0.98\sigma(+0.01)$	$-0.08 \pm 0.98\sigma(+0.00)$
CMS_htt_scale_met_8TeV	$+0.02 \pm 0.99\sigma$	$-0.00 \pm 0.99\sigma(-0.01)$	$+0.03 \pm 0.99\sigma(-0.03)$
CMS_htt_ttbarNorm_8TeV	$+0.23 \pm 0.76\sigma$	$-0.04 \pm 0.81\sigma(-0.31)$	$+0.30 \pm 0.76\sigma(-0.05)$
CMS_htt_ttbar_emb_8TeV	$+0.05 \pm 0.99\sigma$	$-0.01 \pm 0.99\sigma(-0.05)$	$+0.07 \pm 0.99\sigma(-0.01)$
CMS_htt_zttNorm_8TeV	$-0.09 \pm 0.94\sigma$	$-0.05 \pm 0.94\sigma(+0.03)$	$-0.08 \pm 0.94\sigma(-0.01)$
CMS_scale_e_8TeV	$+0.41 \pm 0.84\sigma$	$+0.43 \pm 0.83\sigma(+0.02)$	$+0.43 \pm 0.84\sigma(-0.01)$
CMS_scale_j_8TeV	$-0.16 \pm 0.83\sigma$	$+0.06 \pm 0.86\sigma(+0.23)$	$-0.25 \pm 0.83\sigma(+0.08)$
CMS_shift1_emu_btag_8TeV_EWK_fine_binning	$+0.01 \pm 0.91\sigma$	$+0.07 \pm 0.94\sigma(+0.06)$	$-0.20 \pm 0.93\sigma(+0.17)$
CMS_shift2_emu_btag_8TeV_EWK_fine_binning	$+0.22 \pm 0.92\sigma$	$+0.01 \pm 0.95\sigma(-0.21)$	$-0.06 \pm 0.92\sigma(+0.20)$
lumi_8TeV	$-0.02 \pm 0.99\sigma$	$-0.03 \pm 0.99\sigma(-0.04)$	$+0.01 \pm 0.99\sigma(+0.02)$
CMS_htt_em_emu_btag_8TeV_EWK_bin_1	$-0.58 \pm 1.00\sigma$	$-0.57 \pm 1.00\sigma(+0.01)$	$-0.60 \pm 1.00\sigma(+0.01)$
CMS_htt_em_emu_btag_8TeV_EWK_bin_3	$-0.02 \pm 0.99\sigma$	$-0.02 \pm 0.99\sigma(-0.00)$	$-0.02 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_btag_8TeV_EWK_bin_4	$-0.05 \pm 0.99\sigma$	$-0.04 \pm 0.99\sigma(+0.00)$	$-0.05 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_btag_8TeV_EWK_bin_5	$+0.05 \pm 0.99\sigma$	$+0.05 \pm 0.99\sigma(+0.00)$	$+0.04 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_btag_8TeV_EWK_bin_6	$-0.01 \pm 0.98\sigma$	$+0.01 \pm 0.98\sigma(+0.02)$	$-0.02 \pm 0.98\sigma(+0.00)$
CMS_htt_em_emu_btag_8TeV_EWK_bin_7	$+0.16 \pm 0.97\sigma$	$+0.18 \pm 0.97\sigma(+0.02)$	$+0.16 \pm 0.97\sigma(-0.00)$
CMS_htt_em_emu_btag_8TeV_EWK_bin_8	$+0.04 \pm 0.97\sigma$	$+0.07 \pm 0.97\sigma(+0.03)$	$+0.04 \pm 0.97\sigma(+0.00)$
CMS_htt_em_emu_btag_8TeV_EWK_bin_9	$-0.01 \pm 0.96\sigma$	$+0.03 \pm 0.96\sigma(+0.04)$	$-0.03 \pm 0.96\sigma(+0.01)$
CMS_htt_em_emu_btag_8TeV_EWK_bin_10	$+0.08 \pm 0.96\sigma$	$+0.12 \pm 0.96\sigma(+0.04)$	$+0.06 \pm 0.96\sigma(+0.01)$
CMS_htt_em_emu_btag_8TeV_Fakes_bin_3	$-0.14 \pm 0.89\sigma$	$-0.16 \pm 0.89\sigma(-0.02)$	$-0.13 \pm 0.89\sigma(-0.01)$
CMS_htt_em_emu_btag_8TeV_Fakes_bin_4	$-0.30 \pm 0.96\sigma$	$-0.31 \pm 0.95\sigma(-0.02)$	$-0.28 \pm 0.96\sigma(-0.01)$
CMS_htt_em_emu_btag_8TeV_Fakes_bin_5	$+0.18 \pm 0.94\sigma$	$+0.19 \pm 0.94\sigma(+0.01)$	$+0.19 \pm 0.95\sigma(-0.00)$
CMS_htt_em_emu_btag_8TeV_Fakes_bin_6	$-0.01 \pm 0.96\sigma$	$+0.01 \pm 0.96\sigma(+0.02)$	$+0.00 \pm 0.96\sigma(-0.01)$
CMS_htt_em_emu_btag_8TeV_Fakes_bin_7	$+0.26 \pm 0.94\sigma$	$+0.29 \pm 0.94\sigma(+0.04)$	$+0.27 \pm 0.94\sigma(-0.01)$
CMS_htt_em_emu_btag_8TeV_Fakes_bin_8	$+0.07 \pm 0.94\sigma$	$+0.11 \pm 0.94\sigma(+0.04)$	$+0.08 \pm 0.94\sigma(-0.01)$
CMS_htt_em_emu_btag_8TeV_Fakes_bin_9	$-0.00 \pm 0.98\sigma$	$+0.03 \pm 0.98\sigma(+0.03)$	$-0.01 \pm 0.98\sigma(+0.00)$
CMS_htt_em_emu_btag_8TeV_Fakes_bin_10	$+0.04 \pm 0.99\sigma$	$+0.06 \pm 0.99\sigma(+0.02)$	$+0.04 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_btag_8TeV_Fakes_bin_11	$-0.05 \pm 0.98\sigma$	$-0.05 \pm 0.98\sigma(+0.01)$	$-0.08 \pm 0.98\sigma(+0.02)$
CMS_htt_em_emu_btag_8TeV_Fakes_bin_12	$-0.05 \pm 0.98\sigma$	$-0.09 \pm 0.98\sigma(-0.04)$	$-0.08 \pm 0.98\sigma(+0.03)$
CMS_htt_em_emu_btag_8TeV_Fakes_bin_13	$+0.14 \pm 0.98\sigma$	$+0.06 \pm 0.99\sigma(-0.07)$	$+0.10 \pm 0.99\sigma(+0.03)$
CMS_htt_em_emu_btag_8TeV_Fakes_bin_14	$+0.02 \pm 0.98\sigma$	$-0.05 \pm 0.99\sigma(-0.07)$	$-0.00 \pm 0.98\sigma(+0.02)$
CMS_htt_em_emu_btag_8TeV_Fakes_bin_15	$+0.13 \pm 0.98\sigma$	$+0.12 \pm 0.98\sigma(-0.01)$	$+0.15 \pm 0.98\sigma(-0.01)$
CMS_htt_em_emu_btag_8TeV_Ztt_bin_2	$-0.01 \pm 0.99\sigma$	$-0.01 \pm 0.99\sigma(+0.00)$	$-0.01 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_btag_8TeV_Ztt_bin_8	$+0.01 \pm 0.99\sigma$	$+0.01 \pm 0.99\sigma(+0.00)$	$+0.01 \pm 0.99\sigma(-0.00)$
CMS_htt_em_emu_btag_8TeV_Ztt_bin_9	$-0.00 \pm 0.99\sigma$	$+0.00 \pm 0.99\sigma(+0.01)$	$-0.00 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_btag_8TeV_Ztt_bin_10	$+0.01 \pm 0.99\sigma$	$+0.02 \pm 0.99\sigma(+0.00)$	$+0.01 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_btag_8TeV_Ztt_bin_11	$-0.01 \pm 0.99\sigma$	$-0.01 \pm 0.99\sigma(+0.00)$	$-0.01 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_btag_8TeV_Ztt_bin_12	$-0.01 \pm 0.99\sigma$	$-0.01 \pm 0.99\sigma(-0.01)$	$-0.01 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_btag_8TeV_Ztt_bin_13	$+0.03 \pm 0.99\sigma$	$+0.01 \pm 0.99\sigma(-0.01)$	$+0.02 \pm 0.99\sigma(+0.01)$
CMS_htt_em_emu_btag_8TeV_Ztt_bin_14	$+0.00 \pm 0.99\sigma$	$-0.01 \pm 0.99\sigma(-0.01)$	$-0.00 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_btag_8TeV_Ztt_bin_15	$+0.02 \pm 0.99\sigma$	$+0.02 \pm 0.99\sigma(-0.00)$	$+0.02 \pm 0.99\sigma(-0.00)$
CMS_htt_em_emu_btag_8TeV_Ztt_bin_16	$-0.03 \pm 0.99\sigma$	$-0.03 \pm 0.99\sigma(+0.01)$	$-0.01 \pm 0.99\sigma(-0.02)$
CMS_htt_em_emu_btag_8TeV_Ztt_bin_17	$-0.01 \pm 0.99\sigma$	$-0.00 \pm 0.99\sigma(+0.01)$	$-0.00 \pm 0.99\sigma(-0.00)$
CMS_htt_em_emu_btag_8TeV_Ztt_bin_18	$+0.01 \pm 0.99\sigma$	$+0.01 \pm 0.99\sigma(+0.00)$	$+0.00 \pm 0.99\sigma(+0.01)$
CMS_htt_em_emu_btag_8TeV_ttbar_bin_1	$-0.11 \pm 0.99\sigma$	$-0.10 \pm 0.99\sigma(+0.01)$	$-0.11 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_btag_8TeV_ttbar_bin_2	$-0.03 \pm 0.98\sigma$	$-0.03 \pm 0.98\sigma(+0.00)$	$-0.03 \pm 0.98\sigma(+0.00)$
CMS_htt_em_emu_btag_8TeV_ttbar_bin_3	$-0.01 \pm 0.99\sigma$	$-0.01 \pm 0.99\sigma(+0.00)$	$-0.01 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_btag_8TeV_ttbar_bin_4	$-0.03 \pm 0.99\sigma$	$-0.03 \pm 0.99\sigma(+0.00)$	$-0.03 \pm 0.99\sigma(+0.00)$
CMS_htt_em_emu_btag_8TeV_ttbar_bin_18	$+0.04 \pm 0.98\sigma$	$+0.06 \pm 0.98\sigma(+0.02)$	$-0.00 \pm 0.98\sigma(+0.03)$
μ	(μ fixed to 0)	$+0.73 \pm 0.70$	-0.40 ± 0.30

Table A.3.: The output of the maximum likelihood fit for the nuisance parameters θ_i and signal strength μ on events selected with the btag BDT in the **btag** category.

Nuisance parameter	b only	b+s @ $m_\phi = 400 GeV/c^2$ (Correlation with μ)	b+s @ $m_\phi = 700 GeV/c^2$ (Correlation with μ)
CMS_eff_b_8TeV	$+0.76 \pm 0.93\sigma$	$+0.34 \pm 0.95\sigma(+0.12)$	$+0.69 \pm 0.93(+0.04)$
CMS_eff_e	$-0.30 \pm 0.87\sigma$	$-0.19 \pm 0.87\sigma(-0.02)$	$-0.29 \pm 0.87(-0.00)$
CMS_eff_m	$-0.30 \pm 0.87\sigma$	$-0.19 \pm 0.87\sigma(-0.02)$	$-0.29 \pm 0.87(-0.00)$
CMS_fake_b_8TeV	$+0.40 \pm 0.98\sigma$	$+0.18 \pm 0.98\sigma(+0.03)$	$+0.35 \pm 0.98(+0.01)$
CMS_htt_DiBosonNorm_8TeV	$-1.11 \pm 0.71\sigma$	$-0.57 \pm 0.74\sigma(-0.29)$	$-0.95 \pm 0.72(-0.18)$
CMS_htt_FakeShape_em_nobtag_8TeV	$+0.40 \pm 0.33\sigma$	$+0.25 \pm 0.32\sigma(+0.14)$	$+0.46 \pm 0.33(-0.15)$
CMS_htt_TTbarShape_em_nobtag_8TeV	$+0.99 \pm 0.86\sigma$	$+0.51 \pm 0.87\sigma(+0.23)$	$+0.52 \pm 0.95(+0.42)$
CMS_htt_fakes_em_8TeV	$+1.13 \pm 0.39\sigma$	$+1.16 \pm 0.38\sigma(-0.03)$	$+1.07 \pm 0.40(+0.12)$
CMS_htt_scale_met_8TeV	$+0.12 \pm 0.99\sigma$	$+0.05 \pm 0.99\sigma(-0.03)$	$+0.11 \pm 0.99(-0.02)$
CMS_htt_ttbarNorm_8TeV	$-1.11 \pm 0.86\sigma$	$-0.48 \pm 0.90\sigma(-0.29)$	$-1.02 \pm 0.86(-0.10)$
CMS_htt_zttNorm_8TeV	$-0.00 \pm 0.68\sigma$	$-0.15 \pm 0.68\sigma(+0.04)$	$-0.03 \pm 0.68(+0.01)$
CMS_scale_e_8TeV	$+0.31 \pm 0.16\sigma$	$+0.25 \pm 0.16\sigma(+0.12)$	$+0.29 \pm 0.16(+0.07)$
CMS_scale_j_8TeV	$+0.20 \pm 0.99\sigma$	$+0.09 \pm 0.99\sigma(+0.07)$	$+0.18 \pm 0.99(+0.03)$
CMS_shift1_emu_nobtag_8TeV_EWK_fine_binning	$+0.91 \pm 0.89\sigma$	$+0.50 \pm 0.92\sigma(+0.18)$	$+0.83 \pm 0.89(+0.07)$
CMS_shift1_emu_nobtag_8TeV_Fakes_fine_binning	$+1.10 \pm 0.87\sigma$	$+0.49 \pm 0.92\sigma(+0.25)$	$+1.02 \pm 0.88(+0.06)$
CMS_shift2_emu_nobtag_8TeV_EWK_fine_binning	$-0.39 \pm 0.94\sigma$	$-0.20 \pm 0.93\sigma(-0.08)$	$-0.71 \pm 0.94(+0.28)$
CMS_shift2_emu_nobtag_8TeV_Fakes_fine_binning	$-0.35 \pm 0.93\sigma$	$-0.18 \pm 0.93\sigma(-0.06)$	$-0.68 \pm 0.93(+0.28)$
lumi_8TeV	$-0.20 \pm 0.99\sigma$	$-0.11 \pm 0.99\sigma(+0.03)$	$-0.18 \pm 0.99(+0.01)$
CMS_htt_em_emu_nobtag_8TeV_EWK_bin_1	$-0.00 \pm 0.99\sigma$	$-0.00 \pm 0.99\sigma(+0.00)$	$-0.00 \pm 0.99(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_EWK_bin_3	$+0.00 \pm 0.99\sigma$	$-0.00 \pm 0.98\sigma(+0.00)$	$+0.00 \pm 0.98(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_EWK_bin_4	$+0.06 \pm 0.99\sigma$	$+0.07 \pm 0.99\sigma(-0.00)$	$+0.06 \pm 0.99(-0.00)$
CMS_htt_em_emu_nobtag_8TeV_EWK_bin_5	$-0.01 \pm 0.99\sigma$	$-0.01 \pm 0.99\sigma(+0.00)$	$-0.01 \pm 0.99(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_3	$+0.01 \pm 0.66\sigma$	$-0.01 \pm 0.67\sigma(+0.01)$	$+0.00 \pm 0.66(+0.01)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_4	$+1.67 \pm 0.68\sigma$	$+1.69 \pm 0.68\sigma(-0.00)$	$+1.67 \pm 0.68(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_5	$-0.36 \pm 0.80\sigma$	$-0.32 \pm 0.81\sigma(-0.01)$	$-0.36 \pm 0.80(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_6	$-0.17 \pm 0.89\sigma$	$-0.16 \pm 0.89\sigma(-0.00)$	$-0.18 \pm 0.89(+0.01)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_7	$-0.10 \pm 0.93\sigma$	$-0.12 \pm 0.93\sigma(+0.01)$	$-0.10 \pm 0.93(+0.01)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_8	$+0.06 \pm 0.95\sigma$	$+0.03 \pm 0.95\sigma(+0.01)$	$+0.06 \pm 0.95(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_9	$+0.41 \pm 0.96\sigma$	$+0.39 \pm 0.96\sigma(+0.00)$	$+0.41 \pm 0.96(-0.00)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_10	$-0.29 \pm 0.95\sigma$	$-0.28 \pm 0.96\sigma(-0.01)$	$-0.28 \pm 0.95(-0.01)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_11	$-0.19 \pm 0.95\sigma$	$-0.16 \pm 0.95\sigma(-0.01)$	$-0.17 \pm 0.95(-0.01)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_12	$+0.07 \pm 0.93\sigma$	$+0.08 \pm 0.94\sigma(-0.01)$	$+0.07 \pm 0.93(-0.01)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_13	$-0.27 \pm 0.89\sigma$	$-0.31 \pm 0.90\sigma(+0.02)$	$-0.27 \pm 0.89(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_14	$+1.27 \pm 0.86\sigma$	$+1.14 \pm 0.87\sigma(+0.05)$	$+1.25 \pm 0.86(+0.01)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_15	$+0.97 \pm 0.83\sigma$	$+0.77 \pm 0.84\sigma(+0.09)$	$+0.93 \pm 0.83(+0.03)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_16	$+0.18 \pm 0.82\sigma$	$-0.05 \pm 0.84\sigma(+0.10)$	$+0.14 \pm 0.82(+0.04)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_17	$-0.24 \pm 0.82\sigma$	$-0.45 \pm 0.83\sigma(+0.09)$	$-0.28 \pm 0.82(+0.04)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_18	$-0.97 \pm 0.81\sigma$	$-1.14 \pm 0.82\sigma(+0.08)$	$-1.01 \pm 0.81(+0.04)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_19	$+0.85 \pm 0.86\sigma$	$+0.64 \pm 0.87\sigma(+0.09)$	$+0.79 \pm 0.86(+0.05)$
CMS_htt_em_emu_nobtag_8TeV_Fakes_bin_20	$+0.19 \pm 0.85\sigma$	$+0.03 \pm 0.86\sigma(+0.07)$	$+0.14 \pm 0.85(+0.04)$
CMS_htt_em_emu_nobtag_8TeV_Ztt_bin_1	$-0.04 \pm 0.99\sigma$	$-0.04 \pm 0.99\sigma(-0.00)$	$-0.04 \pm 0.99(-0.00)$
CMS_htt_em_emu_nobtag_8TeV_Ztt_bin_3	$+0.00 \pm 0.99\sigma$	$-0.00 \pm 0.99\sigma(+0.00)$	$+0.00 \pm 0.99(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_Ztt_bin_24	$+0.06 \pm 0.99\sigma$	$+0.12 \pm 0.99\sigma(-0.02)$	$+0.05 \pm 0.99(+0.01)$
CMS_htt_em_emu_nobtag_8TeV_Ztt_bin_25	$-0.16 \pm 0.99\sigma$	$-0.11 \pm 0.99\sigma(-0.02)$	$-0.16 \pm 0.99(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_Ztt_bin_26	$-0.19 \pm 0.99\sigma$	$-0.12 \pm 0.99\sigma(-0.03)$	$-0.20 \pm 0.99(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_Ztt_bin_27	$+0.07 \pm 0.99\sigma$	$+0.18 \pm 0.99\sigma(-0.04)$	$+0.06 \pm 0.99(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_Ztt_bin_28	$-0.09 \pm 0.99\sigma$	$-0.05 \pm 0.99\sigma(-0.01)$	$-0.08 \pm 0.99(-0.00)$
CMS_htt_em_emu_nobtag_8TeV_Ztt_bin_29	$-0.08 \pm 0.99\sigma$	$-0.14 \pm 0.99\sigma(+0.02)$	$-0.01 \pm 0.99(-0.06)$
CMS_htt_em_emu_nobtag_8TeV_Ztt_bin_30	$+0.10 \pm 0.99\sigma$	$+0.07 \pm 0.99\sigma(+0.01)$	$+0.12 \pm 0.99(-0.02)$
CMS_htt_em_emu_nobtag_8TeV_Ztt_bin_31	$+0.05 \pm 0.99\sigma$	$+0.06 \pm 0.99\sigma(-0.00)$	$+0.04 \pm 0.99(+0.01)$
CMS_htt_em_emu_nobtag_8TeV_ttbar_bin_1	$-0.06 \pm 0.99\sigma$	$-0.06 \pm 0.99\sigma(+0.00)$	$-0.06 \pm 0.99(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_ttbar_bin_3	$+0.00 \pm 0.99\sigma$	$-0.00 \pm 0.99\sigma(+0.00)$	$+0.00 \pm 0.99(+0.00)$
CMS_htt_em_emu_nobtag_8TeV_ttbar_bin_4	$+0.07 \pm 0.99\sigma$	$+0.08 \pm 0.99\sigma(-0.00)$	$+0.07 \pm 0.99(-0.00)$
CMS_htt_em_emu_nobtag_8TeV_ttbar_bin_31	$+0.10 \pm 0.99\sigma$	$+0.11 \pm 0.99\sigma(-0.01)$	$+0.08 \pm 0.99(+0.02)$
μ	(μ fixed to 0)	-0.35 ± 0.13	-0.06 ± 0.04

Table A.4.: The output of the maximum likelihood fit for the nuisance parameters θ_i and signal strength μ on events selected with the btag BDT in the **nobtag** category.

References

[1] Ian J R Aitchison and Anthony J G Hey. *Gauge theories in particle physics: a practical introduction; 4th ed.* Vol. 2. Boca Raton, FL: CRC Press, 2013. URL: http://cds.cern.ch/record/1507184.

[2] Ian J R Aitchison and Anthony J G Hey. *Gauge theories in particle physics: a practical introduction; 4th ed.* Vol. 2. Boca Raton, FL: CRC Press, 2013. URL: http://cds.cern.ch/record/1507184.

[3] Michael E Peskin and Daniel V Schroeder. *An Introduction to Quantum Field Theory; 1995 ed.* Includes exercises. Boulder, CO: Westview, 1995. URL: http://cds.cern.ch/record/257493.

[4] Claude Itzykson and Jean-Bernard Zuber. *Quantum field theory.* International series in pure and applied physics. Also a reprint ed.: Mineola, Dover, 2005. New York, NY: McGraw-Hill, 1980. URL: http://cds.cern.ch/record/100772.

[5] D.J. Bird et al. "Detection of a cosmic ray with measured energy well beyond the expected spectral cutoff due to cosmic microwave radiation". In: *Astrophys.J.* 441 (1995), pp. 144–150. DOI: 10.1086/175344.

[6] Emmy Noether. "Invariant Variation Problems". In: *Gott.Nachr.* 1918 (1918), pp. 235–257. DOI: 10.1080/00411457108231446. arXiv: physics/0503066 [physics].

[7] Geoffrey T J Arnison et al. "Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540GeV$". In: *Phys. Lett. B* 122.CERN-EP-83-13 (Jan. 1983), 103–116. 31 p. URL: http://cds.cern.ch/record/142059.

[8] Geoffrey T J Arnison et al. "Observation of the muonic decay of the charged intermediate vector boson". In: *Phys. Lett. B* 134.CERN-EP-83-199 (Dec. 1983), 469–476. 16 p. URL: http://cds.cern.ch/record/149156.

[9] Geoffrey T J Arnison et al. "Observation of muonic Z^0 decay at the $\overline{p}p$ collider". In: *Phys. Lett. B* 147.CERN-EP-84-100 (Aug. 1984), 241–248. 18 p. URL: http://cds.cern.ch/record/154095.

[10] Geoffrey T J Arnison et al. "Experimental observation of lepton pairs of invariant mass around $95GeV/c^2$ at the CERN SPS collider". In: *Phys. Lett. B* 126.CERN-EP-83-073 (June 1985), 398–410. 17 p. URL: http://cds.cern.ch/record/163857.

[11] K.A. Olive et al. "Review of Particle Physics". In: *Chin.Phys.* C38 (2014), p. 090001. DOI: 10.1088/1674-1137/38/9/090001.

[12] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys.Rev.Lett.* 13 (1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321.

[13] Peter W. Higgs. "Broken symmetries, massless particles and gauge fields". In: *Phys.Lett.* 12 (1964), pp. 132–133. DOI: 10.1016/0031-9163(64)91136-9.

[14] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys.Rev.Lett.* 13 (1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.

[15] Peter W. Higgs. "Spontaneous Symmetry Breakdown without Massless Bosons". In: *Phys.Rev.* 145 (1966), pp. 1156–1163. DOI: 10.1103/PhysRev.145.1156.

[16] T.W.B. Kibble. "Symmetry breaking in nonAbelian gauge theories". In: *Phys.Rev.* 155 (1967), pp. 1554–1561. DOI: 10.1103/PhysRev.155.1554.

[17] G.S. Guralnik, C.R. Hagen, and T.W.B. Kibble. "Global Conservation Laws and Massless Particles". In: *Phys.Rev.Lett.* 13 (1964), pp. 585–587. DOI: 10.1103/PhysRevLett.13.585.

[18] F. Abe et al. "Observation of top quark production in $\bar{p}p$ collisions". In: *Phys.Rev.Lett.* 74 (1995), pp. 2626–2631. DOI: 10.1103/PhysRevLett.74.2626. arXiv: hep-ex/9503002 [hep-ex].

[19] S. Abachi et al. "Observation of the top quark". In: *Phys.Rev.Lett.* 74 (1995), pp. 2632–2637. DOI: 10.1103/PhysRevLett.74.2632. arXiv: hep-ex/9503003 [hep-ex].

[20] Serguei Chatrchyan et al. "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV". In: *JHEP* 1306 (2013), p. 081. DOI: 10.1007/JHEP06(2013)081. arXiv: 1303.4571 [hep-ex].

[21] Georges Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys.Lett.* B716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[22] Serguei Chatrchyan et al. "Measurement of the properties of a Higgs boson in the four-lepton final state". In: *Phys.Rev.* D89.9 (2014), p. 092007. DOI: 10.1103/PhysRevD.89.092007. arXiv: 1312.5353 [hep-ex].

[23] Vardan Khachatryan et al. "Observation of the diphoton decay of the Higgs boson and measurement of its properties". In: *Eur.Phys.J.* C74.10 (2014), p. 3076. DOI: 10.1140/epjc/s10052-014-3076-z. arXiv: 1407.0558 [hep-ex].

[24] Serguei et al. Chatrchyan. "Measurement of Higgs boson production and properties in the WW decay channel with leptonic final states". In: *J. High Energy Phys.* 01.arXiv:1312.1129. CMS-HIG-13-023. CERN-PH-EP-2013-221 (Dec. 2013). Replaced with published version. Added journal reference and DOI, 096. 83 p. URL: http://cds.cern.ch/record/1633401.

[25] Serguei et al. Chatrchyan. "Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks". In: *Phys. Rev. D* 89.arXiv:1310.3687. CMS-HIG-13-012. CERN-PH-EP-2013-188 (Oct. 2013). Comments: Replaced with published version. Added journal reference and DOI, 012003. 49 p. URL: http://cds.cern.ch/record/1610290.

[26] Vardan Khachatryan et al. "Search for the standard model Higgs boson produced through vector boson fusion and decaying to b bbar". In: (2015). arXiv: 1506.01010 [hep-ex].

[27] "Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons". In: *J. High Energy Phys.* 05.arXiv:1401.5041. CMS-HIG-13-004. CERN-PH-EP-2014-001 (Jan. 2014). Comments: Published in JHEP, 104. 73 p. URL: http://cds.cern.ch/record/1643937.

[28] Georges Aad et al. "Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments". In: *Phys.Rev.Lett.* 114 (2015), p. 191803. DOI: 10.1103/PhysRevLett.114.191803. arXiv: 1503.07589 [hep-ex].

[29] G.W. Bennett et al. "Final Report of the Muon E821 Anomalous Magnetic Moment Measurement at BNL". In: *Phys.Rev.* D73 (2006), p. 072003. DOI: 10.1103/PhysRevD.73.072003. arXiv: hep-ex/0602035 [hep-ex].

[30] M.J.G. Veltman. "Second Threshold in Weak Interactions". In: *Acta Phys.Polon.* B8 (1977), p. 475.

[31] Steven Weinberg. "Gauge Hierarchies". In: *Phys.Lett.* B82 (1979), p. 387. DOI: 10.1016/0370-2693(79)90248-X.

[32] M.J.G. Veltman. "The Infrared - Ultraviolet Connection". In: *Acta Phys.Polon.* B12 (1981), p. 437.

[33] C.H. Llewellyn Smith and Graham G. Ross. "The Real Gauge Hierarchy Problem". In: *Phys.Lett.* B105 (1981), p. 38. DOI: 10.1016/0370-2693(81)90035-6.

[34] F. Zwicky. "Die Rotverschiebung von extragalaktischen Nebeln". In: *Helv.Phys.Acta* 6 (1933), pp. 110–127.

[35] Edward W. Kolb and Michael S. Turner. "The Early Universe". In: *Nature* 294 (1981), p. 521. DOI: 10.1038/294521a0.

[36] D.N. Spergel et al. "First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Determination of cosmological parameters". In: *Astrophys.J.Suppl.* 148 (2003), pp. 175–194. DOI: 10.1086/377226. arXiv: astro-ph/0302209 [astro-ph].

[37] P.A.R. Ade et al. "Planck 2015 results. XIII. Cosmological parameters". In: (2015). arXiv: 1502.01589 [astro-ph.CO].

[38] Pierre Fayet and S. Ferrara. "Supersymmetry". In: *Phys.Rept.* 32 (1977), pp. 249–334. DOI: 10.1016/0370-1573(77)90066-7.

[39] Pierre Fayet. "Spontaneously Broken Supersymmetric Theories of Weak, Electromagnetic and Strong Interactions". In: *Phys.Lett.* B69 (1977), p. 489. DOI: 10.1016/0370-2693(77)90852-8.

[40] Glennys R. Farrar and Pierre Fayet. "Phenomenology of the Production, Decay, and Detection of New Hadronic States Associated with Supersymmetry". In: *Phys.Lett.* B76 (1978), pp. 575–579. DOI: 10.1016/0370-2693(78)90858-4.

[41] Pierre Fayet. "Relations Between the Masses of the Superpartners of Leptons and Quarks, the Goldstino Couplings and the Neutral Currents". In: *Phys.Lett.* B84 (1979), p. 416. DOI: 10.1016/0370-2693(79)91229-2.

[42] Stephen P. Martin. "A Supersymmetry primer". In: *Adv.Ser.Direct.High Energy Phys.* 21 (2010), pp. 1–153. DOI: 10.1142/9789814307505_0001. arXiv: hep-ph/9709356 [hep-ph].

[43] A. Brignole, Luis E. Ibanez, and C. Munoz. "Soft supersymmetry breaking terms from supergravity and superstring models". In: *Adv.Ser.Direct.High Energy Phys.* 21 (2010), pp. 244–268. DOI: 10.1142/9789814307505_0004. arXiv: hep-ph/9707209 [hep-ph].

[44] Kenzo Inoue et al. "Aspects of Grand Unified Models with Softly Broken Supersymmetry". In: *Prog.Theor.Phys.* 68 (1982), p. 927. DOI: 10.1143/PTP.68.927.

[45] Edward Witten. "Constraints on Supersymmetry Breaking". In: *Nucl.Phys.* B202 (1982), p. 253. DOI: 10.1016/0550-3213(82)90071-2.

[46] Edward Witten. "Dynamical Breaking of Supersymmetry". In: *Nucl.Phys.* B188 (1981), p. 513. DOI: 10.1016/0550-3213(81)90006-7.

[47] N. Sakai. "Naturalness in Supersymmetric Guts". In: *Z.Phys.* C11 (1981), p. 153. DOI: 10.1007/BF01573998.

[48] John F. Donoghue and Ling Fong Li. "Properties of Charged Higgs Bosons". In: *Phys.Rev.* D19 (1979), p. 945. DOI: 10.1103/PhysRevD.19.945.

[49] J.F. Gunion and Howard E. Haber. "Higgs Bosons in Supersymmetric Models. 1." In: *Nucl.Phys.* B272 (1986), p. 1. DOI: 10.1016/0550-3213(86)90340-8.

[50] J.F. Gunion and Howard E. Haber. "Higgs Bosons in Supersymmetric Models. 2. Implications for Phenomenology". In: *Nucl.Phys.* B278 (1986), p. 449. DOI: 10.1016/0550-3213(86)90050-7.

[51] S. Heinemeyer, O. Stal, and G. Weiglein. "Interpreting the LHC Higgs Search Results in the MSSM". In: *Phys.Lett.* B710 (2012), pp. 201–206. DOI: 10.1016/j.physletb.2012.02.084. arXiv: 1112.3026 [hep-ph].

[52] S Heinemeyer et al. "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties". In: (2013). Ed. by S Heinemeyer. DOI: 10.5170/CERN-2013-004. arXiv: 1307.1347 [hep-ph].

[53] LHC Higgs Corss Section Working Group. *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*. https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCH
2014.

[54] Maximilien Brice. "First LHC magnets installed. Installation du premier aimant du LHC". Apr. 2005. URL: https://cds.cern.ch/record/834351.

[55] Julie Haffner. "The CERN accelerator complex. Complexe des accélérateurs du CERN". In: (Oct. 2013). General Photo. URL: https://cds.cern.ch/record/1621894.

[56] Oliver Sim Brüning et al. *LHC Design Report*. Geneva: CERN, 2004. URL: https://cds.cern.ch/record/782076.

[57] Lyndon Evans and Philip Bryant. "LHC Machine". In: *JINST* 3 (2008). Ed. by Lyndon Evans, S08001. DOI: 10.1088/1748-0221/3/08/S08001.

[58] K. Aamodt et al. “The ALICE experiment at the CERN LHC”. In: *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002.

[59] G. Aad et al. “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[60] Jr. Alves A. Augusto et al. “The LHCb Detector at the LHC”. In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.

[61] O. Adriani et al. “The LHCf detector at the CERN Large Hadron Collider”. In: *JINST* 3 (2008), S08006. DOI: 10.1088/1748-0221/3/08/S08006.

[62] James Pinfold et al. *Technical Design Report of the MoEDAL Experiment*. Tech. rep. CERN-LHCC-2009-006. MoEDAL-TDR-001. Geneva: CERN, June 2009. URL: https://cds.cern.ch/record/1181486.

[63] G. Anelli et al. “The TOTEM experiment at the CERN Large Hadron Collider”. In: *JINST* 3 (2008), S08007. DOI: 10.1088/1748-0221/3/08/S08007.

[64] S. Chatrchyan et al. “The CMS experiment at the CERN LHC”. In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004.

[65] “The construction of the CMS electromagnetic calorimeter: lead-tungstate (PWO) crystals. La construction du calorimetre electromagnetique de CMS: les cristaux de PWO”. CMS Collection. Feb. 2001. URL: https://cds.cern.ch/record/929101.

[66] S. Chatrchyan et al. *CMS Tracker Detector Performance Results*. https://twiki.cern.ch/twiki/bin/view/CMSPublic/DPGResultsTRK. Accessed: 21.03.2015.

[67] Serguei Chatrchyan et al. “Description and performance of track and primary-vertex reconstruction with the CMS tracker”. In: *JINST* 9.10 (2014), P10009. DOI: 10.1088/1748-0221/9/10/P10009. arXiv: 1405.6569 [physics.ins-det].

[68] M. Raymond et al. “The CMS tracker APV25 0.25-mu-m CMOS readout chip”. In: *Conf.Proc.* C00091111 (2000), pp. 130–134.

[69] Christopher M Bishop. *Pattern Recognition and Machine Learning*. Information Science and Statistics. New York, NY: Springer, 2006.

[70] RobertE. Schapire. “The strength of weak learnability”. English. In: *Machine Learning* 5.2 (1990), pp. 197–227. ISSN: 0885-6125. DOI: 10.1007/BF00116037. URL: http://dx.doi.org/10.1007/BF00116037.

[71] Xuchun Li, Lei Wang, and E. Sung. “A study of AdaBoost with SVM based weak learners”. In: *Neural Networks, 2005. IJCNN ’05. Proceedings. 2005 IEEE International Joint Conference on*. Vol. 1. July 2005, 196–201 vol. 1. DOI: 10.1109/IJCNN.2005.1555829.

[72] Robert E Schapire. “Using output codes to boost multiclass learning problems”. In: *ICML*. Vol. 97. 1997, pp. 313–321.

[73] Yoav Freund, Robert E Schapire, et al. “Experiments with a new boosting algorithm”. In: *ICML*. Vol. 96. 1996, pp. 148–156.

[74] Yoav Freund and Robert E Schapire. "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting". In: *Journal of Computer and System Sciences* 55.1 (1997), pp. 119–139. ISSN: 0022-0000. DOI: http://dx.doi.org/10.1006/jcss.1997.1504. URL: http://www.sciencedirect.com/science/article/pii/S002200009791504X.

[75] Jerome Friedman, Trevor Hastie, Robert Tibshirani, et al. "Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors)". In: *The annals of statistics* 28.2 (2000), pp. 337–407.

[76] Jerome H. Friedman. "Greedy Function Approximation: A Gradient Boosting Machine". English. In: *The Annals of Statistics* 29.5 (2001), pp. 1189–1232. ISSN: 00905364. URL: http://www.jstor.org/stable/2699986.

[77] A. Mayr et al. "The Evolution of Boosting Algorithms. From Machine Learning to Statistical Modelling". In: *Methods of Information in Medicine* 53.6 (2014), pp. 419–417. ISSN: 0026-1270. DOI: http://dx.doi.org/10.3414/ME13-01-0122. URL: http://methods.schattauer.de/en/contents/archivestandard/issue/2156/manuscript/22328.html.

[78] Andreas Hoecker et al. "TMVA: Toolkit for Multivariate Data Analysis". In: *PoS* ACAT (2007), p. 040. arXiv: physics/0703039.

[79] I. Antcheva et al. "ROOT: A C++ framework for petabyte data storage, statistical analysis and visualization". In: *Comput.Phys.Commun.* 182 (2011), pp. 1384–1385. DOI: 10.1016/j.cpc.2011.02.008.

[80] E. Chabanat and N. Estre. "Deterministic annealing for vertex finding at CMS". In: (2005), pp. 287–290.

[81] R. Fruhwirth et al. "New vertex reconstruction algorithms for CMS". In: *eConf* C0303241 (2003), TULT013. arXiv: physics/0306012 [physics].

[82] A Holzner. "78 reconstructed vertices in event from high-pileup run 198609". CMS Collection. Sept. 2012.

[83] *Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET*. Tech. rep. CMS-PAS-PFT-09-001. 2009. Geneva: CERN, Apr. 2009.

[84] *Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector*. Tech. rep. CMS-PAS-PFT-10-001. 2010.

[85] *Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV*. Tech. rep. CMS-PAS-PFT-10-002. Geneva: CERN, 2010.

[86] D. Buskulic et al. "Performance of the ALEPH detector at LEP". In: *Nucl.Instrum.Meth.* A360 (1995), pp. 481–506. DOI: 10.1016/0168-9002(95)00138-7.

[87] Claus Grupen et al. *The ALEPH "experience": 25 years of memories*. Geneva: CERN, 2005.

[88] Wolfgang Adam et al. *Track Reconstruction in the CMS tracker*. Tech. rep. CMS-NOTE-2006-041. Geneva: CERN, Dec. 2006.

[89] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti- k t jet clustering algorithm". In: *Journal of High Energy Physics* 2008.04 (2008), p. 063. URL: http://stacks.iop.org/1126-6708/2008/i=04/a=063.

[90] S Catani et al. "Longitudinally-invariant $k_\perp$-clustering algorithms for hadron-hadron collisions". In: *Nucl. Phys. B* 406.CERN-TH-6775-93. LU-TP-93-2 (Feb. 1993), 187–224. 38 p. URL: http://cds.cern.ch/record/246812.

[91] S D Ellis and Davison Eugene Soper. "Successive combination jet algorithm for hadron collisions". In: *Phys. Rev. D* 48.hep-ph/9305266. CERN-TH-6860-93 (Apr. 1993), 3160–3166. 15 p. URL: http://cds.cern.ch/record/248558.

[92] Yuri L. Dokshitzer et al. "Better jet clustering algorithms". In: *JHEP* 9708 (1997), p. 001. DOI: 10.1088/1126-6708/1997/08/001. arXiv: hep-ph/9707323 [hep-ph].

[93] M. Wobisch and T. Wengler. "Hadronization corrections to jet cross-sections in deep inelastic scattering". In: (1998). arXiv: hep-ph/9907280 [hep-ph].

[94] The CMS Collaboration. "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS". In: *J. Instrum.* 6.arXiv:1107.4277. CMS-JME-10-011. CERN-PH-EP-2011-102 (July 2011), P11002. 67 p. URL: http://cds.cern.ch/record/1369486.

[95] Matteo Cacciari and Gavin P Salam. "Pileup subtraction using jet areas". In: *Phys. Lett. B* 659.arXiv:0707.1378. LPTHE-07-01. 1-2 (July 2007). Comments: 10 pages, 6 figures, 119–126. 10 p. URL: http://cds.cern.ch/record/1046126.

[96] CMS Collaboration. "Pileup Jet Identification". In: (2013).

[97] G Antchev et al. "Luminosity-Independent Measurement of the Proton-Proton Total Cross Section at $\sqrt{s} = 8$ TeV". In: *Phys. Rev. Lett.* 111.TOTEM-2012-005. CERN-PH-EP-2012-354 (Nov. 2012), 012001. 8 p. URL: https://cds.cern.ch/record/1495764.

[98] The CMS Collaboration. "Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s}$=8 TeV". In: *J. High Energy Phys.* 02.arXiv:1312.7582. CMS-TOP-12-007. CERN-PH-EP-2013-234 (Dec. 2013). Comments: Replaced with published version. Added journal reference and DOI, 024. 29 p. URL: http://cds.cern.ch/record/1640470.

[99] Ivan Marchesini. *Identification of b-quark Jets in the CMS experiment*. Tech. rep. CMS-CR-2013-412. Geneva: CERN, Nov. 2013. URL: http://cds.cern.ch/record/1629516.

[100] *Performance of b tagging at sqrt(s)=8 TeV in multijet, ttbar and boosted topology events*. Tech. rep. CMS-PAS-BTV-13-001. Geneva: CERN, 2013. URL: http://cds.cern.ch/record/1581306.

[101] S. Baffioni et al. "Electron reconstruction in CMS". In: *Eur.Phys.J.* C49 (2007), pp. 1099–1116. DOI: 10.1140/epjc/s10052-006-0175-5.

[102] R Fruehwirth. "A Gaussian-mixture approximation of the Bethe-Heitler model of electron energy loss by bremsstrahlung". In: *Computer Physics Communications* 154.2 (2003), pp. 131–142. ISSN: 0010-4655. DOI: http://dx.doi.org/10.1016/S0010-4655(03)00292-3. URL: http://www.sciencedirect.com/science/article/pii/S0010465503002923.

[103] Wolfgang Adam et al. *Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC*. Tech. rep. CMS-NOTE-2005-001. Geneva: CERN, Jan. 2005.

[104] "Electron performance with 19.6 fb^{-1} of data collected at $\sqrt{s} = 8$ TeV with the CMS detector." In: (2013).

[105] *Performance of muon identification in pp collisions at s**0.5 = 7 TeV*. Tech. rep. CMS-PAS-MUO-10-002. 2010. Geneva: CERN, 2010. URL: http://cds.cern.ch/record/1279140.

[106] The CMS collaboration. "Missing transverse energy performance of the CMS detector". In: *Journal of Instrumentation* 6.09 (2011), P09001. URL: http://stacks.iop.org/1748-0221/6/i=09/a=P09001.

[107] *MET performance in 8 TeV data*. Tech. rep. CMS-PAS-JME-12-002. Geneva: CERN, 2013. URL: http://cds.cern.ch/record/1543527.

[108] Alexander L. Read. "Presentation of search results: The CL(s) technique". In: *J.Phys.* G28 (2002), pp. 2693–2704. DOI: 10.1088/0954-3899/28/10/313.

[109] Thomas Junk. "Confidence level computation for combining searches with small statistics". In: *Nucl.Instrum.Meth.* A434 (1999), pp. 435–443. DOI: 10.1016/S0168-9002(99)00498-2. arXiv: hep-ex/9902006 [hep-ex].

[110] "Procedure for the LHC Higgs boson search combination in summer 2011". In: (2011).

[111] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur.Phys.J.* C71 (2011), p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0,10.1140/epjc/s10052-013-2501-z. arXiv: 1007.1727 [physics.data-an].

[112] Lorenzo Moneta et al. "The RooStats Project". In: *PoS* ACAT2010 (2010), p. 057. arXiv: 1009.1003 [physics.data-an].

[113] J. Gaiser. "Charmonium Spectroscopy From Radiative Decays of the J/ψ and ψ'". In: (1982).

[114] Sergio Cittolin, Attila Rácz, and Paris Sphicas. *CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project*. Technical Design Report CMS. Geneva: CERN, 2002. URL: http://cds.cern.ch/record/578006.

[115] Federico De Guio. "The CMS data quality monitoring software: experience and future prospects". In: *J.Phys.Conf.Ser.* 513 (2014), p. 032024. DOI: 10.1088/1742-6596/513/3/032024.

[116] Stefan Höche. "Introduction to parton-shower event generators". In: (2014). arXiv: 1411.4085 [hep-ph].

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

[118] John Allison et al. “Geant4 developments and applications”. In: *IEEE Trans.Nucl.Sci.* 53 (2006), p. 270. DOI: 10.1109/TNS.2006.869826.

[119] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. “PYTHIA 6.4 Physics and Manual”. In: *JHEP* 0605 (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175 [hep-ph].

[120] Michiel Botje et al. “The PDF4LHC Working Group Interim Recommendations”. In: (2011). arXiv: 1101.0538 [hep-ph].

[121] N. Davidson et al. “Universal Interface of TAUOLA Technical and Physics Documentation”. In: *Comput.Phys.Commun.* 183 (2012), pp. 821–843. DOI: 10.1016/j.cpc.2011.12.009. arXiv: 1002.0543 [hep-ph].

[122] Johan Alwall et al. “MadGraph 5 : Going Beyond”. In: *JHEP* 1106 (2011), p. 128. DOI: 10.1007/JHEP06(2011)128. arXiv: 1106.0522 [hep-ph].

[123] Simone Alioli et al. “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”. In: *JHEP* 1006 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[124] Vardan Khachatryan et al. “Measurement of the differential cross section for top quark pair production in pp collisions at sqrt(s) = 8 TeV”. In: (2015). arXiv: 1505.04480 [hep-ex].

[125] *Search for a standard model Higgs bosons decaying to tau pairs in pp collisions.* Tech. rep. CMS-PAS-HIG-12-018. Geneva: CERN, 2012.

[126] J. Beringer et al. “The Review of Particle Physics 2012 and 2013 partial update for the 2014 edition”. In: *Phys. Rev. D* 86 (2012).

[127] *Procedure for the LHC Higgs boson search combination in Summer 2011.* Tech. rep. CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11. Geneva: CERN, Aug. 2011.

[128] *CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update.* Tech. rep. CMS-PAS-LUM-13-001. Geneva: CERN, 2013. URL: https://cds.cern.ch/record/1598864.