

Contents

1	The Higgs Bosons and the MSSM	5
1.1	The Standard Model of Particle Physics	6
1.1.1	Introduction	6
1.1.2	Precision Test and Limitation of the SM	7
1.2	The Minimal Supersymmetric Standard Model	9
1.2.1	Introduction to the MSSM	9
1.2.2	The Higgs Sector in the MSSM	12
1.3	Neutral Higgs Bosons Phenomenology in the MSSM	13
1.3.1	MSSM Higgs Couplings with SM Particles	13
1.3.2	MSSM Higgs Benchmark Scenarios	15
1.3.3	Neutral MSSM Higgs Bosons Production and Decay at LHC	16
1.3.4	Current Status of the Search for Neutral MSSM Higgs Bosons	20
2	The ATLAS Detector at the LHC	23
2.1	The Large Hadron Collider	24
2.2	The ATLAS Detector	25
2.2.1	The ATLAS coordinate system	25
2.2.2	The Inner Detector	28
2.2.3	The Calorimeter System	30
2.2.4	The Muon Spectrometer	30
2.2.5	The Trigger System	32
2.2.6	Luminosity Measurement	32

3 Reconstruction of Physics Objects	35
3.1 Reconstruction of Charged Particle Tracks	36
3.2 Vertex Reconstruction	36
3.3 Electron Reconstruction and Identification	37
3.4 Muon Reconstruction	39
3.5 Jet Reconstruction and Energy Calibration	40
3.6 Identification of Jets from b-quarks	42
3.7 Tau-Jet Reconstruction	44
3.8 Missing Transverse Energy	44
3.9 Overlap Removal	45
3.10 Trigger	45
3.11 Truth Particles	46
4 Neutral MSSM Higgs Bosons Search	47
4.1 Introduction	49
4.1.1 The Higgs Sector in the MSSM	49
4.1.2 Signal and Background Processes	51
4.1.3 Analysis Strategy	53
4.1.4 Data and Simulated Event Samples	56
4.2 Event Selection and Categorization	57
4.2.1 The Common Selection Criteria	57
4.2.2 b-vetoed Event Category	58
4.2.3 b-tagged Event Category	59
4.2.4 Mass Reconstruction with MMC Technique	62
4.3 Background Prediction and Validation	65
4.3.1 Validation of the $t\bar{t}$ Background Prediction	65
4.3.2 Multi-jet Background Measurement	67
4.3.3 $Z \rightarrow \tau\tau + \text{Jets}$ Background Measurement	74
4.4 Systematic Uncertainties	77
4.4.1 Detector-related Systematics Uncertainties	78
4.4.2 Theoretical Uncertainties	81
4.4.3 Systematic Uncertainties of the $Z/\gamma^* \rightarrow \tau\tau$ Embedded Sample	85

<i>CONTENTS</i>	3
4.4.4 QCD Multi-Jet Systematic Uncertainties	86
4.5 Results	89
4.5.1 Statistical Procedure	89
4.5.2 Exclusion Limits on the Signal Production	93
5 Prospects For Higgs Search Improvements	99
5.1 Track-based Jets	101
5.2 Performance of the Track-based Jets	103
5.2.1 Track-based Jets Reconstruction	103
5.2.2 B-tagging with Track-Based Jets	105
5.2.3 Employment of Track-jet for the MSSM Higgs Boson Search	106
5.3 Systematic Uncertainties of Track-jet Reconstruction	111
5.3.1 Material Budget Uncertainty on Track-Based Jets Recon- struction	112
5.3.2 Validation of the Track Subtraction Method	114
A QCD Trigger bias	137
B QCD Additional Plots	139
C Limits cross checks and additional plots	141
C.1 The ABCD Method	141
C.2 Additional Limit Checks	142
C.3 Shape Systematics	155

Chapter 1

The Higgs Bosons and the MSSM

This chapter is devoted to introduce the theoretical background to the experimental search presented in this thesis. A brief overview of the Standard Model of particle physics is given in Section 1.1 based on Reference [1]. Among all the extension of the Standard Model, the Minimal Supersymmetric extension of the Standard Model (MSSM) is a theoretically favoured scenario as one of the most predictive framework beyond the Standard Model. The MSSM is introduced in Section 1.2 with focus on its Higgs sector, based on References [2, 3]. Finally, a review of the MSSM Higgs bosons phenomenological aspects which are relevant to the presented search is given in Section 1.3, based on Reference [4].

1.1 The Standard Model of Particle Physics

1.1.1 Introduction

A detailed description of the Standard Model of particle physics may be found in Reference [6], only a brief overview is given in what follows.

The Standard Model (SM) of particle physics is a theory aimed to describe and quantitatively predict the phenomenology of fundamental interactions. At “microscopic” level the spectrum of all interactions between matter and radiation can be understood in terms of three classes of fundamental forces: the strong, the electromagnetic and the weak forces. These interactions are described by a local relativistic quantum field theory, where to each particle is associated a field with suitable transformation properties under the Lorentz group. The theory is based on the principle of gauge invariance, which means invariance under a symmetry transformations that operates on basic internal degrees of freedom and depends on the space-time coordinate. The gravitational force is negligible in atomic and nuclear physics, in fact, quantum effects of gravity are expected at energies corresponding to the Planck mass $E \sim M_{\text{Planck}}c^2 \sim 10^{19}$ GeV.

The SM is a gauge field theory based on the symmetry group $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$. The group has $8 + 3 + 1 = 12$ generators with a non trivial commutator algebra. The electromagnetic and weak interactions [7–9] are described by the $SU(2)_L \otimes U(1)_Y$ symmetry group, while the $SU(3)_c$ is the colour group of the theory of strong interactions (QCD) [10]. To each generator of the symmetry group is associated a vector boson which act as mediator of the corresponding interactions. Eight gluons are associated to the $SU(3)_c$ colour generators, while four gauge bosons W^\pm , Z^0 and γ are associated to the generators of $SU(2)_L \otimes U(1)_Y$. Only the gluons and the photon are massless since the symmetry induced by the other three generators is spontaneously broken. In the SM the spontaneous symmetry breaking is realised by the Higgs mechanism [11–15]. The Higgs boson acts as mediator of a new class of interactions that, at tree level, are coupled in proportion to the particle masses. An Higgs boson with properties that resemble the one of the SM has recently been discovered at the LHC with $m_H \sim 126$ GeV [16, 17].

The fermionic matter fields of the SM are quarks and leptons. Quarks are subject to all SM interactions, each type of quark is a colour triplet and carries electroweak charges, in particular electric charges $+2/3$ for up-type quarks and $-1/3$ for down-type quarks. Leptons are colourless but have electroweak charges, in particular electric charge -1 for charged leptons e , μ and τ (opposite sign charge is intended for respective anti-particle) and charge 0 for neutrinos ν_e , ν_μ and ν_τ . Quarks and leptons are grouped in three “generations” with equal quantum numbers but different masses.

1.1.2 Precision Test and Limitation of the SM

The Standard Model has been successfully tested in a vast number of experiments over a wide range of energies during the last few decades. Precision tests of the electroweak theory performed at LEP, SLC and the Tevatron [20], has confirmed that the couplings of quark and leptons to the weak gauge bosons W^\pm and Z are indeed precisely those prescribed by the gauge symmetry. The accuracy of a few per-mille for these tests implies that, not only the tree level, but also the structure of quantum corrections has been verified. Several other experimental results [19] including rare decays of hadrons provides low-energies test of the Standard Model. The recent discovery of a Higgs boson is also another success of the SM, its mass, spin and couplings are in agreement with expected values from a global fit of electroweak constraints [21].

Among all the parameters of the Standard Model only few of them presents tension with experimental data, the most significant discrepancies are slightly above three standard deviations and are given by: the anomalous magnetic moment of the muon [22], a_μ and the forward-backward asymmetry of top quarks [23], $A_{FB}^{t\bar{t}}$.

In spite of this success, the Standard Model is conceptually unsatisfactory for quite few deficiencies and is widely believed to be an effective theory valid only at the present accessible energies. Beside the fact that it does not include gravitational force, it does not explain the pattern of fermion masses and in its simplest version does not include neutrino masses, it has at least other three conceptual problems which indicates the need for physics Beyond the Standard Model (BSM):

Hierarchy Problem Calculating the radiative correction to the Higgs boson mass, quadratic divergences occur of the order of the cut-off scale Λ , where Λ defines the energy beyond which the theory ceases to be valid and new physics should appear [24]. If the cut-off is chosen to be $\sim M_{Planck}$, then a fine tuning with an unnaturally high precision, $\mathcal{O}(10^{-30})$, should occur to cancel these divergences leaving the Higgs boson with a mass of the order of the electroweak breaking scale, M_{EW} . A question that has no satisfactory answer in the SM is how these cancellations can occur and why $\Lambda \gg M_{EW}$, these problems are called the fine-tuning and hierarchy problem [25–27].

Dark Matter The SM does not have a candidate which can explain the large contribution of non-barionic, non-luminous matter to the density of the Universe [28–30]. To be a Dark Matter candidate a particle should be stable, massive and should interact only via very weak interactions.

Unification Problem Another unsatisfactory aspect of the SM is that does not provide the unification of the electroweak and strong interactions, their couplings do not meet at high energies. Considering the successful unification of electromagnetic and weak interaction, the existence of Grand Unified Theory (GUT) has been suggested [31, 32], which predicts the unification of all the three gauge coupling strength at the GUT energy scale, $\Lambda_{GUT} \simeq 10^{16}$ GeV and describes the three forces within a single gauge group with just one coupling constant.

Among all the extension of the SM, Supersymmetry is a theoretically favoured scenario as the most predictive framework beyond the Standard Model. As discussed in Section 1.2, it gives a natural answer to the hierarchy problem, provides a suitable candidate for Dark Matter and predicts unification of the three gauge couplings at GUT energy scale.

1.2 The Minimal Supersymmetric Standard Model

1.2.1 Introduction to the MSSM

Supersymmetry (SUSY) [33–35] was first introduced as a natural way to solve the hierarchy problem. In Supersymmetry a new symmetry that relates bosons to fermions is introduced. The SUSY generators \mathcal{Q} transforms fermion into bosons and vice versa:

$$\mathcal{Q}|\text{Fermion}\rangle = |\text{Boson}\rangle, \quad \mathcal{Q}|\text{Boson}\rangle = |\text{Fermion}\rangle \quad (1.1)$$

In a supersymmetric extension of the SM each of the known fundamental particles is in either a chiral or gauge *supermultiplet* and must have a superpartner with spin differing by 1/2 unit. SUSY naturally solve the hierarchy problem since the quadratic divergent loop contribution to the Higgs mass of the SM particles are cancelled by the loop contribution of the corresponding superpartners. The name of the superpartner of the quarks and leptons are made by adding an “s” to the SM name, standing for scalar. Accordingly, the gauge bosons related to the generator of the group $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$ should also have a spin 1/2 partner, whose name will be made by adding a “ino” at the end of the SM name. The symbol of superpartners is defined by adding a $(\tilde{\cdot})$ to the SM symbol. The SUSY particles share the same couplings with their SM partner, since the left-handed and right-handed components of fermions transform differently under gauge transformations also their superpartner present this feature.

The Minimal Supersymmetric extension of the Standard Model (MSSM) [36–41], is defined by requiring the minimal gauge group (i.e., the SM one) and the minimal particle content: three generation of fermions (without right-handed neutrinos), gauge bosons and two Higgs doublet, each with its own superpartner. Tables 1.1 and 1.2 summarise chiral and gauge supermultiplets in the MSSM. Among the gauge eigenstates summarised in these tables, the superpartner of the Higgs bosons, the *higgsinos* mix with the *wino* and *bino* to give the “ino” mass eigenstates: two charginos $\chi_{1,2}^{\pm}$ and four neutralinos $\chi_{1,2,3,4}^0$.

Names	Supermultiplets	Spin 1/2	Spin 0
quark, squarks $(\times 3 \text{ families})$	Q	$(u_L \ d_L)$	$(\tilde{u}_L \ \tilde{d}_L)$
	\bar{u}	u_R^\dagger	\tilde{u}_R^*
	\bar{d}	d_R^\dagger	\tilde{d}_R^*
leptons, sleptons $(\times 3 \text{ families})$	L	$(\nu \ e_L)$	$(\tilde{\nu} \ \tilde{e}_L)$
	\bar{e}	e_R^\dagger	\tilde{e}_R^*
higgsinos, Higgs	H_1	$(\tilde{H}_1^0 \ \tilde{H}_1^-)$	$(H_1^0 \ H_1^-)$
	H_1	$(\tilde{H}_2^+ \ \tilde{H}_2^0)$	$(H_2^+ \ H_2^0)$

Table 1.1: This table is based on Reference [2] and summarise the chiral supermultiplets in the Minimal Supersymmetric Standard Model. The spin-0 fields are complex scalars and the spin-1/2 are left-handed two-component Weyl fermions.

Names	Supermultiplets	Spin 1	Spin 1/2
gluons, gluinos	G_a ($a = 1, \dots, 8$)	g	\tilde{g}
W bosons, winos	W_a ($a = 1, \dots, 3$)	$W^\pm \ W^0$	$\tilde{W}^\pm \ \tilde{W}^0$
B boson, bino	B	B^0	\tilde{B}^0

Table 1.2: This table is based on Reference [2] and summarise the gauge supermultiplets in the Minimal Supersymmetric Standard Model.

R-parity conservation

The MSSM requires a discrete and multiplicative symmetry called *R*-parity [35], this symmetry assures barion and lepton number conservation and it is defined as follows:

$$R_p = (-1)^{2s+3B=L} \quad (1.2)$$

where L and B are lepton and barion numbers and s stands for the spin quantum number. The R-parity quantum number has value +1 for ordinary SM particles and -1 for their superpartners. This symmetry was first introduced as a simple way to overcome the problem of instability of the proton. Lepton and barion number violation leads, in many cases, to unstable proton with life-time shorter than the experimental lower limit. The conservation of *R*-parity has also other important phenomenological consequences: SUSY particles are always produced in pairs and decays always in an odd number of SUSY particles. Furthermore, the lightest SUSY particle, often chosen as one of the neutralinos, is stable and provides a suitable candidate for dark matter.

The Soft SUSY Breaking

In case Supersymmetry is an exact symmetry of nature, the SM particles and their relative superpartners should have the same mass and quantum numbers except for the spin. However, the particle spectrum of SUSY has not yet been observed, suggesting that, if these particles exist, they should have an higher mass than their SM superpartners. To achieve SUSY-breaking in a way which does not reintroduce the quadratic divergences to the Higgs mass squared, a so called “soft-SUSY-breaking” term is introduced to the SUSY Lagrangian [42, 43]. This term explicitly break SUSY introducing the mass terms for Higgs, gauginos and sfermions, as well as trilinear coupling terms between sfermions and Higgs bosons. In general, if intergenerational mixing and complex phases are allowed, the soft-SUSY-breaking terms will introduce a huge number of unknown parameters $\mathcal{O}(100)$ [44]. However, in absence of phases and mixing, and if the soft terms obey a set of boundary conditions [42, 43], only few new parameters are introduced

$\mathcal{O}(10)$.

1.2.2 The Higgs Sector in the MSSM

In the MSSM two doublets of complex scalar field of opposite hypercharge are required to break the electroweak symmetry, this requirement is necessary to generate masses separately for isospin up-type fermion and down-type fermions [34, 45, 46] and to cancel chiral anomalies that otherwise would spoil the renormalizability of the theory [47]. The two Higgs doublet are:

$$H_1 = \begin{pmatrix} H_1^0 \\ H_1^- \end{pmatrix} \text{ with } Y_{H_1} = -1, \quad H_2 = \begin{pmatrix} H_2^+ \\ H_2^0 \end{pmatrix} \text{ with } Y_{H_2} = +1 \quad (1.3)$$

In analogy with the SM, a similar Higgs mechanism is employed in the MSSM [36, 48] requiring that the minimum of the Higgs potential breaks $SU(2)_L \otimes U(1)_Y$ group while preserving the electromagnetic symmetry $U(1)_Q$. The neutral components of the two Higgs field acquire vacuum expectation values:

$$\langle H_1^0 \rangle = \frac{v_1}{\sqrt{2}}, \quad \langle H_2^0 \rangle = \frac{v_2}{\sqrt{2}} \quad (1.4)$$

Three of the original eight degrees of freedom of the scalar fields are absorbed by the W^\pm and Z bosons, which acquire their longitudinal polarisations and masses. The remaining degrees of freedom correspond to five scalar Higgs bosons: two CP-even and neutral h and H , a neutral CP-odd boson A and a pair of charged bosons H^\pm . Six parameters describes the MSSM Higgs sector: M_h , M_H , M_A , M_{H^\pm} , β and α , where the latter represents the mixing angle in the neutral CP-even sector, while $\tan \beta$ is equal to the ratio between the two vacuum expectation values $\tan \beta = v_1/v_2$. At tree level, only two of these parameters are actually independent, a common choice is to keep $\tan \beta$ and M_A as free the parameters of the Higgs sector. At tree level, the supersymmetric structure of the theory impose a strong hierarchical structure on the Higgs bosons mass spectrum: the h boson is the lightest with $M_h < M_Z$, while $M_A < M_H$ and $M_{H^\pm}^2 = M_A^2 M_W^2$. Furthermore, the following relation holds between the mixing angles:

$$\cos^2(\beta - \alpha) = \frac{M_h^2(M_Z^2 - M_h^2)}{M_A^2(M_H^2 - M_h^2)} \quad (1.5)$$

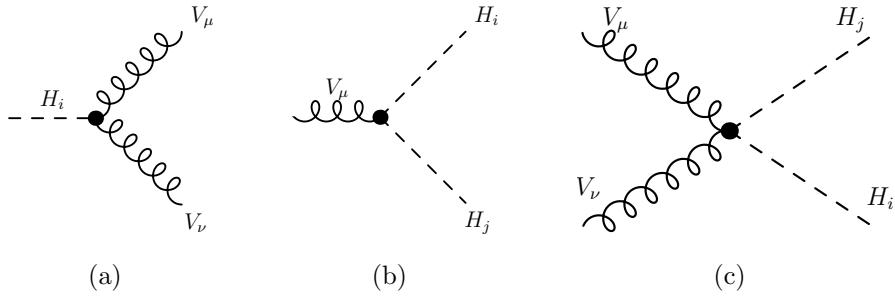


Figure 1.1: Feynman diagrams for the couplings between one Higgs boson and two gauge bosons (a), two Higgs bosons and one gauge boson (b) and two Higgs bosons and two gauge bosons (c). Based on [3].

These relations are broken by large radiative corrections to the Higgs bosons masses [49], which cause the constraint on the mass of h to move from the tree level value of M_Z to $M_h \lesssim 140$ GeV. Another restriction, coming from GUT assumptions gives $1 \lesssim \tan \beta \lesssim m_t/m_b$ [50].

1.3 Neutral Higgs Bosons Phenomenology in the MSSM

1.3.1 MSSM Higgs Couplings with SM Particles

The phenomenology of the MSSM Higgs bosons depends on their couplings with standard model and supersymmetric particles, a short overview of the former, based on the Reference [3], is given in this section.

The possible couplings between MSSM Higgs bosons and vector bosons are: the three-linear couplings $V_\mu V_\nu H_i$ among one Higgs boson and two gauge bosons and $V_\mu H_i H_j$ among one gauge boson and two Higgs bosons, as well as the couplings between two Higgs bosons and two gauge bosons $V_\mu V_\nu H_i H_j$. Figure 1.3.1 shows the Feynman diagram relative to these couplings. Among all of them, the most relevant for MSSM Higgs phenomenology are the trilinear couplings between two gauge bosons and one Higgs boson $V_\mu V_\nu H_i$. For this case, since the photon is massless,

there are no Higgs- $\gamma\gamma$ and Higgs- $Z\gamma$ couplings at tree level. CP-invariance also forbids WWA , ZZA and WZH^\pm couplings. Then, for the $V_\mu V_\nu H_i$ couplings, only the following possibilities remains:

$$Z_\mu Z_\nu h : ig_z M_Z \sin(\beta - \alpha) g_{\mu\nu}, \quad Z_\mu Z_\nu H : ig_z M_Z \cos(\beta - \alpha) g_{\mu\nu} \quad (1.6)$$

$$W_\mu^+ W_\nu^- h : ig_w M_W \sin(\beta - \alpha) g_{\mu\nu}, \quad W_\mu^+ W_\nu^- H : ig_w M_W \cos(\beta - \alpha) g_{\mu\nu} \quad (1.7)$$

The couplings of the neutral CP-even Higgs bosons h and H with pair of vector bosons are proportional to $\sin(\beta - \alpha)$ and $\cos(\beta - \alpha)$ respectively, where $\cos(\beta - \alpha)$ is fixed at tree level following equation (1.5). An interesting phenomenological consequence is that, calling G_{VVh} and G_{Vvh} the coupling between two generic vector bosons and one of the neutral CP-even Higgs bosons the following equation holds:

$$G_{VVh}^2 + G_{Vvh}^2 = g_{Vvh_{SM}}^2 \quad (1.8)$$

The equations (1.7) and (1.8) leads to the fact that the couplings with vector bosons for h (H) increase (decrease) with $\tan\beta$. For relatively large value of $\tan\beta$, h has SM-like couplings with vector bosons while H decouple from them. For an overview of all the other couplings between vector bosons and Higgs bosons, charged Higgs, trilinear and quartic coupling between Higgs bosons and couplings to SUSY particles refer to Reference [3].

The MSSM Higgs bosons couplings with isospin up-type u , and down-type d fermions also depend on $\tan\beta$ and may be written as follows:

$$\begin{aligned} G_{huu} &\propto m_u [\sin(\beta - \alpha) + \cot\beta \cos(\beta - \alpha)], & G_{hdd} &\propto m_u [\sin(\beta - \alpha) - \tan\beta \cos(\beta - \alpha)] \\ G_{Huu} &\propto m_u [\cos(\beta - \alpha) - \cot\beta \sin(\beta - \alpha)], & G_{Hdd} &\propto m_d [\cos(\beta - \alpha) + \tan\beta \sin(\beta - \alpha)] \\ G_{Auu} &\propto m_u \cot\beta, & G_{Add} &\propto m_d \tan\beta \end{aligned}$$

The couplings with down-type (up-type) fermions of either the h or H boson is enhanced (suppressed) by a factor $\tan\beta$, depending on the magnitude of $\cos(\beta - \alpha)$ or $\sin(\beta - \alpha)$, while the coupling of A boson with down-type (up-type) fermions are directly enhanced (suppressed) by $\tan\beta$.

1.3.2 MSSM Higgs Benchmark Scenarios

At tree level, the MSSM Higgs bosons masses, decay branching fraction and production cross section are all determined by two parameters, by convention chosen to be M_A and $\tan \beta$. As it has been pointed out in Section 1.2.2, radiative corrections contribute significantly to the MSSM Higgs bosons masses and the prediction of physics observables becomes dependent on several MSSM parameters [49]. The main corrections arises from the top-stop (s)quark sector and for large $\tan \beta$ also the bottom-sbottom (s)quark sector becomes increasingly important. Furthermore, these corrections are dependent on the SUSY-breaking scale M_{SUSY} , the trilinear Higgs-stop, Higgs-sbottom Yukawa couplings, the electroweak gaugino and gluino mass parameters.

Due to the large number of free parameters, a complete scan of the MSSM parameter space is impractical in experimental analysis and phenomenological studies. To cope with this difficulty several benchmark scenarios has been proposed [4, 52], where the SUSY parameters entering via radiative corrections are fixed to particular benchmark values which exhibit interesting features of the MSSM Higgs phenomenology, while the parameters M_A and $\tan \beta$ are left free to vary. Usually results are presented in a $M_A - \tan \beta$ plane.

The m_h^{\max} benchmark scenario [51] was used in the past searches for neutral MSSM Higgs bosons performed at LEP, Tevatron and LHC [64–67]. In this benchmark scenario the parameters that contributes to radiative corrections are fixed such that the mass of the light CP-even Higgs boson, M_h , is maximal under the variation of M_A and $\tan \beta$. The m_h^{\max} scenario allows to set conservative lower bounds on M_A , M_H^\pm and $\tan \beta$ [52]. However, given the recent discovery of a Higgs boson with mass ~ 126 GeV, this scenario tend to predict a too high mass for M_h , resulting to be, for large regions of the MSSM parameter space, inconsistent with this observation. This scenario is still currently used in the presented analysis since it offer the possibility to compare results with past experiments.

Recently, several benchmark scenarios has been updated [4] to accommodate the experimental constraints on past neutral MSSM Higgs searches and the observation of a SM-like Higgs boson. An interesting updated benchmark scenario is

the m_h^{mod} scenario, which has the feature to predict $M_h \simeq 125.5 \pm 3$ GeV for large region of MSSM parameter space. The m_h^{mod} scenario configuration is obtained by reducing the amount of mixing in the stop sector with respect to the m_h^{max} scenario. This can be done for both signs of the MSSM parameter that regulate the stop mixing X_t , giving rise to two complementary scenarios m_h^{mod+} and m_h^{mod-} . The difference between these two scenarios is found to be negligible for experimental searches, the m_h^{mod+} benchmark scenario has been used throughout this thesis as reference scenario.

Other interesting benchmark scenario are the light stop scenario and the light stau scenario. The first may lead to relevant modification of the gluon fusion production cross section, while the second leads to modification of the di-photon decay branching fraction of the light CP-even MSSM Higgs boson. For an overview of other relevant benchmark scenarios refer to Reference [4].

1.3.3 Neutral MSSM Higgs Bosons Production and Decay at LHC

For large region of the MSSM parameter space a SM-like Higgs boson is expected, this role is commonly played by the lightest CP-even Higgs boson, h . Given the Higgs bosons couplings discussed in Section 1.3.1 turns out that the MSSM Higgs bosons H and A tend to be degenerate in mass and decouple from gauge bosons. Furthermore the coupling of the latter two Higgs bosons with down (up) type fermions are enhanced (suppressed) by $\tan\beta$, therefore, for large $\tan\beta$ bottom-quark and τ lepton will play an important role for the Higgs bosons production and its decays.

The production of the neutral CP -even MSSM Higgs bosons at hadron colliders proceeds via the same processes as for the SM Higgs production. The pseudoscalar A , instead, cannot be produced in association with gauge bosons or in vector boson fusion (VBF) processes at tree-level, as this coupling is forbidden due to CP -invariance. At the LHC one of the most relevant production mechanisms for the MSSM Higgs bosons is gluon fusion, $gg \rightarrow A/H/h$. In addition, the production in association with b -quarks becomes important for large value of $\tan\beta$. These are

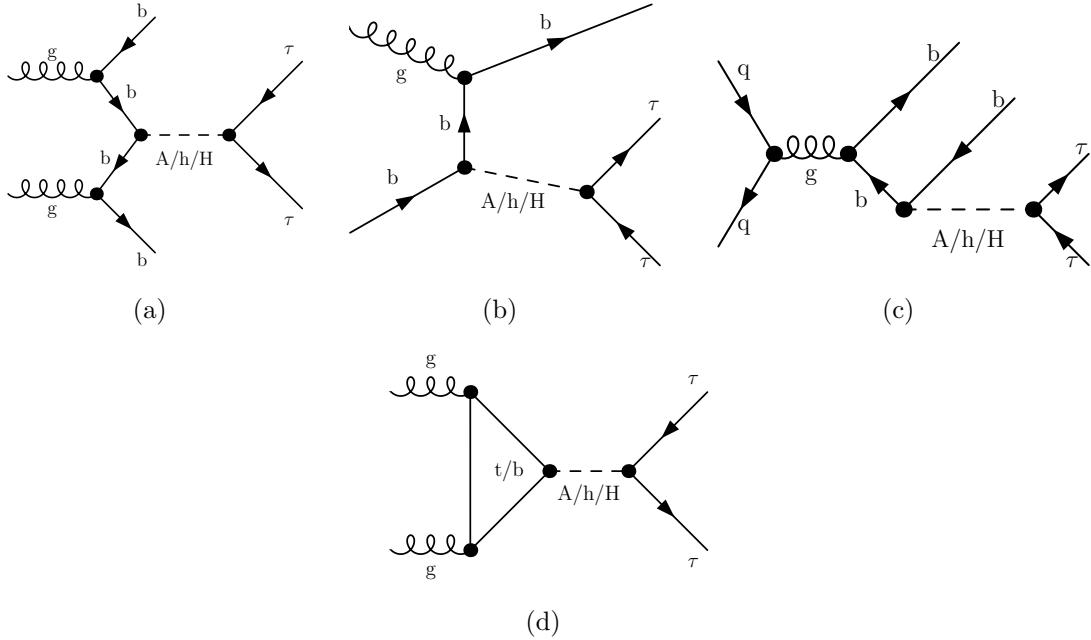


Figure 1.2: Feynman diagram for the production of the neutral MSSM Higgs bosons in association with b -quarks (a,b,c) and via gluon fusion (d) processes, subsequent decay in tau lepton pairs is considered.

the only two production mechanism that are considered in the presented analysis. Figure 1.2 shows the Feynman-diagram for these processes, while Figure 1.3 shows the production cross section of the neutral MSSM Higgs bosons via these two processes in the m_h^{max} scenario.

The decays of the neutral MSSM Higgs bosons (in the assumption that all supersymmetric particle are heavy enough) are the same as for the SM one with the already cited exception of A . Figure 1.3 shows the decay branching fractions in the m_h^{mod+} scenario for h , H and A as a function of the mass of A for two values of $\tan\beta$. The decay into tau lepton pairs is the most important after $b\bar{b}$ and the one used in this thesis.

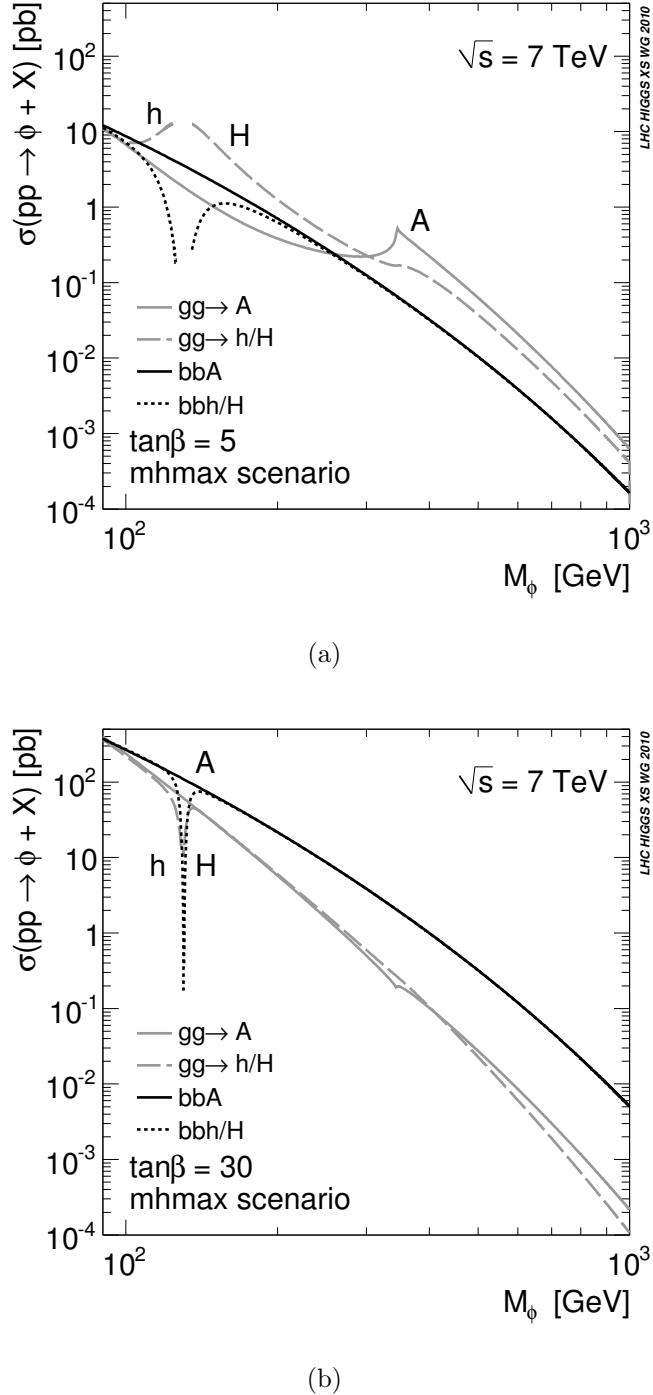


Figure 1.3: Central predictions for the total MSSM Higgs bosons production cross sections via gluon fusion and Higgs radiation off bottom quarks for $\sqrt{s} = 7$ TeV using NNLO and NLO MSTW2008 PDFs m_h^{max} scenario; (a) $\tan \beta = 5$, (b) $\tan \beta = 30$. Reference [5].

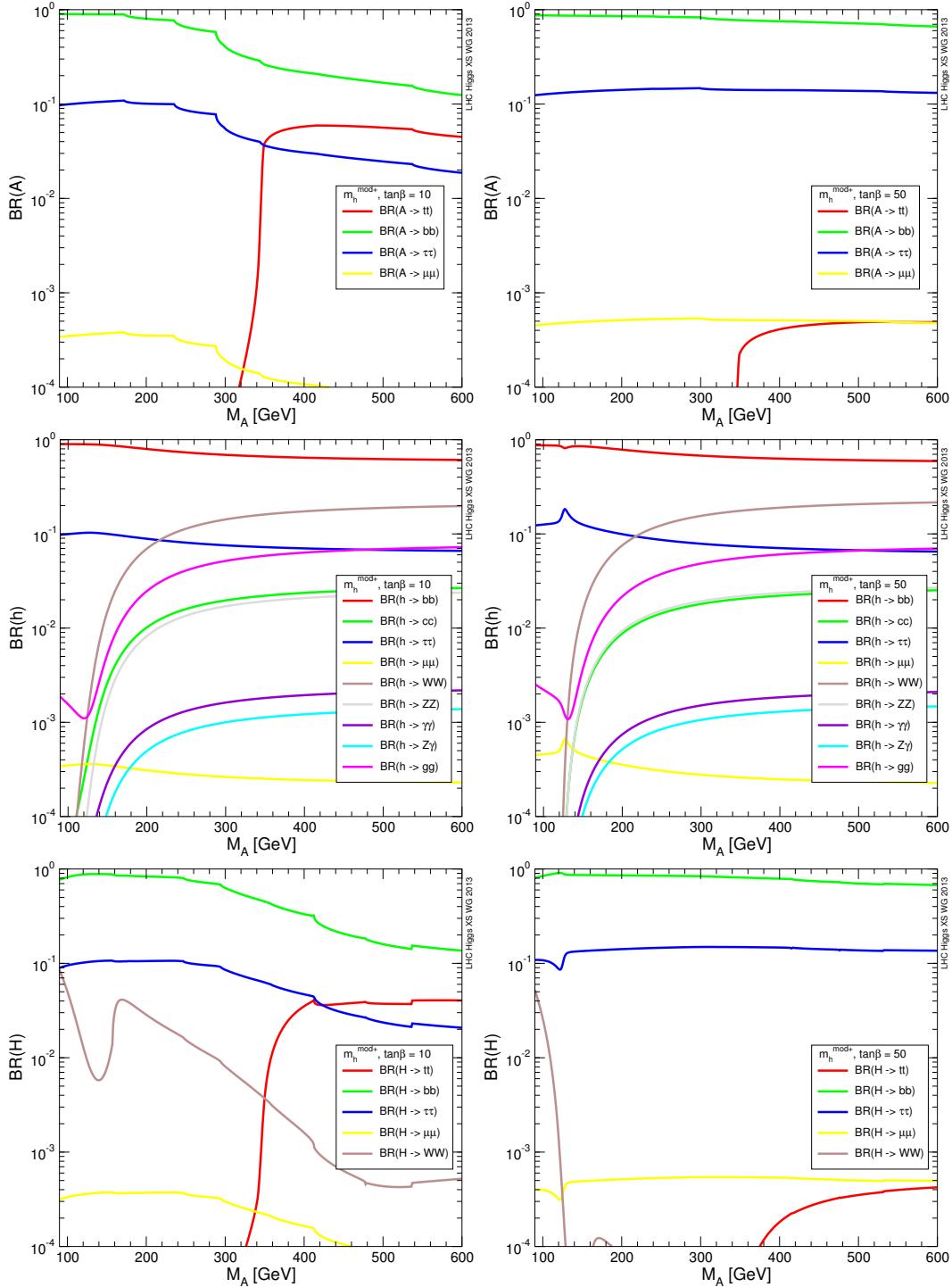


Figure 1.4: Branching fraction for the MSSM neutral Higgs bosons, $h/H/A$, in the m_h^{mod+} scenario for $\tan\beta = 10$ and $\tan\beta = 50$. Reference [4].

1.3.4 Current Status of the Search for Neutral MSSM Higgs Bosons

The measure of the couplings of the observed SM-like Higgs boson can shed light on the Higgs sector and determine if this boson is fully responsible for the generation of all the SM particles masses. There are two approaches to explore the Higgs sector: one, is to use the measured Higgs couplings with SM particles to set constraint on new physics, while the other is to directly search for additional Higgs bosons in a well defined model.

In case the SM-like Higgs boson is interpreted as the light CP-even Higgs boson of the MSSM, the couplings of the Higgs boson to vector bosons (k_V), up-type fermions (k_u) and down-type fermions (k_d), can be expressed as a function of m_A and $\tan \beta$ and this allow to set exclusion limits in the $m_A - \tan \beta$ plane [53]. Figure 1.5 shows the exclusion limits in a “simplified MSSM” model [54, 55] via fits to the measured rates of Higgs boson production and decay.

The current latest constraint on $m_A - \tan \beta$ by direct search of neutral MSSM Higgs bosons [] are shown in Figure 1.6 and are part of the work of this thesis.

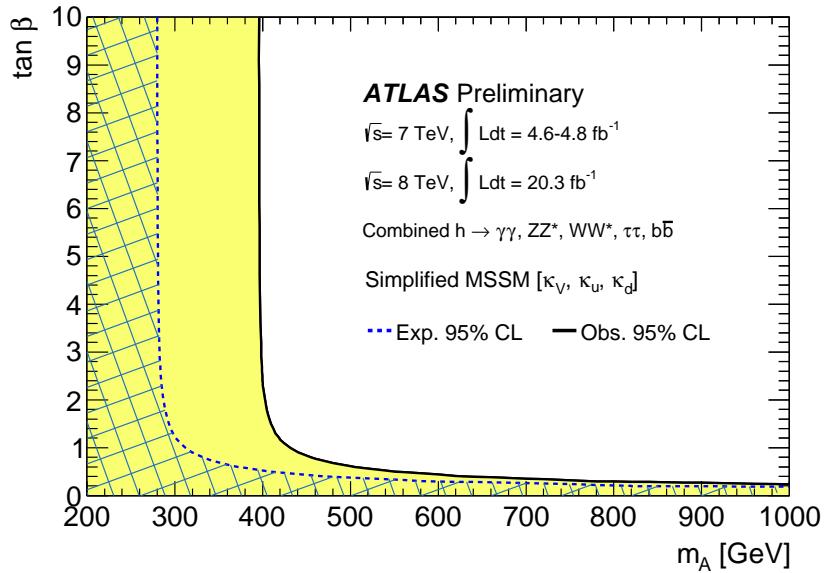


Figure 1.5: Regions of the $m_A - \tan \beta$ plane excluded in a simplified MSSM model via fits to the measured rates of Higgs boson production and decays. The likelihood contours where $2 \ln \Lambda = 6.0$, corresponding approximately to 95% CL (2σ), are indicated for the data and expectation assuming the SM Higgs sector. The light shaded and hashed regions indicate the observed and expected exclusions, respectively. The SM decoupling limit is $m_A \rightarrow \infty$. See Reference [53].

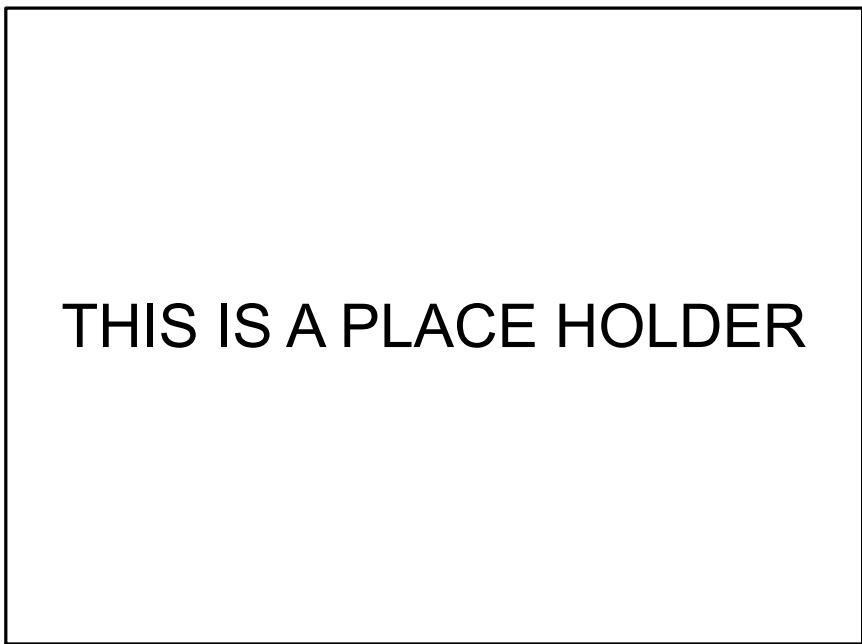


Figure 1.6: Limit CMS or ATLAS depending if we manage to publish in time

Chapter 2

The ATLAS Detector at the LHC

The Large Hadron Collider (LHC) located at the European Organization for Nuclear Research (CERN) is the largest particle collider facility in the world. The ATLAS experiment is one of the several experiments situated at the LHC, it is a general-purpose detector dedicated to explore a wide range of physics topics, from precision measurements of known Standard Model processes to the search for physics beyond the Standard Model. Proton-proton collision data recorded by the ATLAS experiment has been used in this thesis for the search for the neutral MSSM Higgs bosons.

This chapter is organised as follows: the design and performance of the LHC are summarised in Section 2.1 (based on Reference [56]), while a description of the ATLAS detector is given in Section 2.2 (based on Reference [57]).

2.1 The Large Hadron Collider

The LHC is a superconducting hadron synchrotron collider. It is installed in the tunnel of the former LEP electron-positron collider and has ~ 27 km circumference. LHC is designed to collide proton beams with a centre-of-mass energy of 14 TeV and an unprecedented luminosity of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$. It can also collide heavy ions (lead) with an energy of 2.8 TeV per nucleon and a peak luminosity of $10^{27} \text{ cm}^{-2}\text{s}^{-1}$.

The LHC provides proton-proton collision for several experiments, the most relevant for the physics program are the ATLAS, CMS [59], LHCb [60] and ALICE [61] experiments. Figure 2.1 shows the layout of the CERN accelerator complex, the protons follows several acceleration steps before injection in the LHC machine. First a linac accelerator (*Linac 2*) accelerates the protons to an energy of 50 MeV, then they are injected in the *booster* which accelerates them further to 1.4 GeV. The energy is increased to 25 GeV and successively to 450 GeV by means of two synchrotron accelerators, the *Proton Synchrotron* (PS) and the *Super Proton Synchrotron* (SPS). Finally, two proton beams are injected with opposite direction into the LHC where they reach their final energy.

The LHC beams are constituted by bunches of protons which are housed in two separate vacuum pipes. Radiofrequency cavities are employed to accelerate the protons, while superconducting magnets bends and focuses the bunches. The LHC is designed to accelerate up to 2835 bunches per beam, each of them containing $\sim 10^{11}$ protons. The nominal bunch spacing allows collision every 25 ns representing a real challenge for any read-out electronics.

In 2010, first collisions took place at the LHC between proton beams of energy of 3.5 TeV. The LHC was successfully in operation during years 2011 and 2012, beam of protons were initially delivered with energies of 3.5 TeV which was increased to 4 TeV in 2012. Peak luminosities of about $4 \times 10^{33} \text{ cm}^{-2}\text{s}^{-2}$ and $8 \times 10^{33} \text{ cm}^{-2}\text{s}^{-2}$ have been reached during years 2011 and 2012 respectively. The ATLAS experiment recorded in fully operational conditions an integrated luminosity of 4.57 fb^{-1} during year 2011, while an integrated luminosity of 20.3 fb^{-1} was recorded during year 2012. Data recorded during these two years led to one

of the major milestone in particle physics, the discovery in 2012 of a Higgs boson.

2.2 The ATLAS Detector

The ATLAS detector is a multi-purpose detector which aim to explore a wide range of physics topics. The physics programme of the ATLAS experiment ranges from precision measurements of known Standard Model processes to the search for physics beyond the Standard Model in a large variety of scenarios. The ATLAS detector is designed to satisfy the tight requirements on particle identification and measurements accuracy imposed by the physics goals. A schematic view of the ATLAS detector is shown in Figure 2.2. ATLAS is 44 m long and 25 m high, it has a typical design for a collider experiments with a forward-backward symmetry with respect the interaction point. It consist of four sub-detector which are built cylindrically around the beam pipe, going from the inside to outside they are: an Inner Detector tracker, an electromagnetic calorimeter, an hadronic calorimeter and finally a muon spectrometer surrounds the whole experiment. Each of these sub-detector is briefly described in what follows based on Reference [57].

2.2.1 The ATLAS coordinate system

The ATLAS coordinate system has its origin in the interaction region. The z -axis is pointing along the beam direction, the y -axis is pointing upwards and the x -axis towards the centre of the LHC ring. The angle ϕ is defined in the plane orthogonal to the beam axis, starting from the positive side of the x -axis. The angle θ is instead defined with respect to the z -axis. A commonly used spatial coordinate in experiments at collider is the rapidity y :

$$y = 1/2 \cdot \ln \left(\frac{E + p_z}{E - P_z} \right) \quad (2.1)$$

The difference in the rapidity of two particles is independent of Lorentz boosts along the beam axis. In the limit of β approaching to 1 or for massless particles

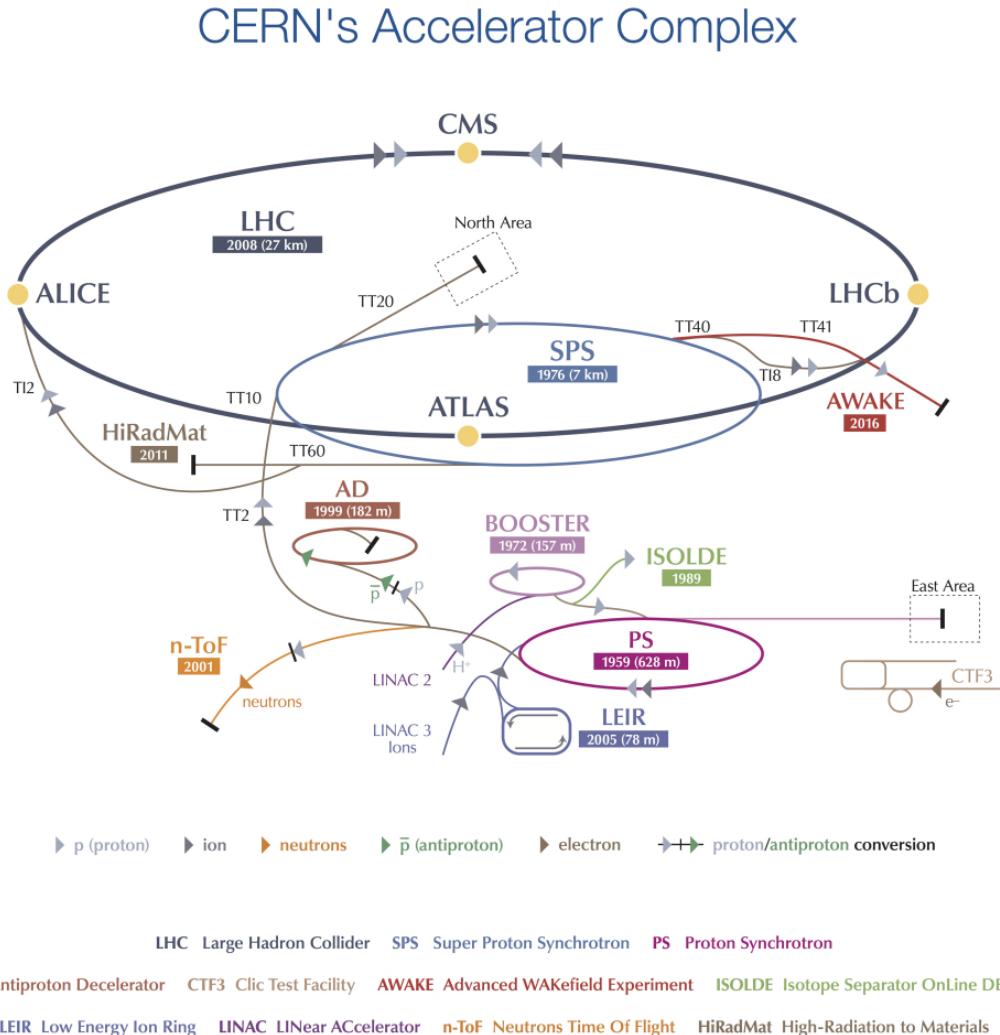


Figure 2.1: Illustration of the CERN accelerator complex [58]. The acceleration of protons starts with Linac2 followed by the acceleration in the Booster. The Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) accelerate further the protons before final injection into the LHC machine.

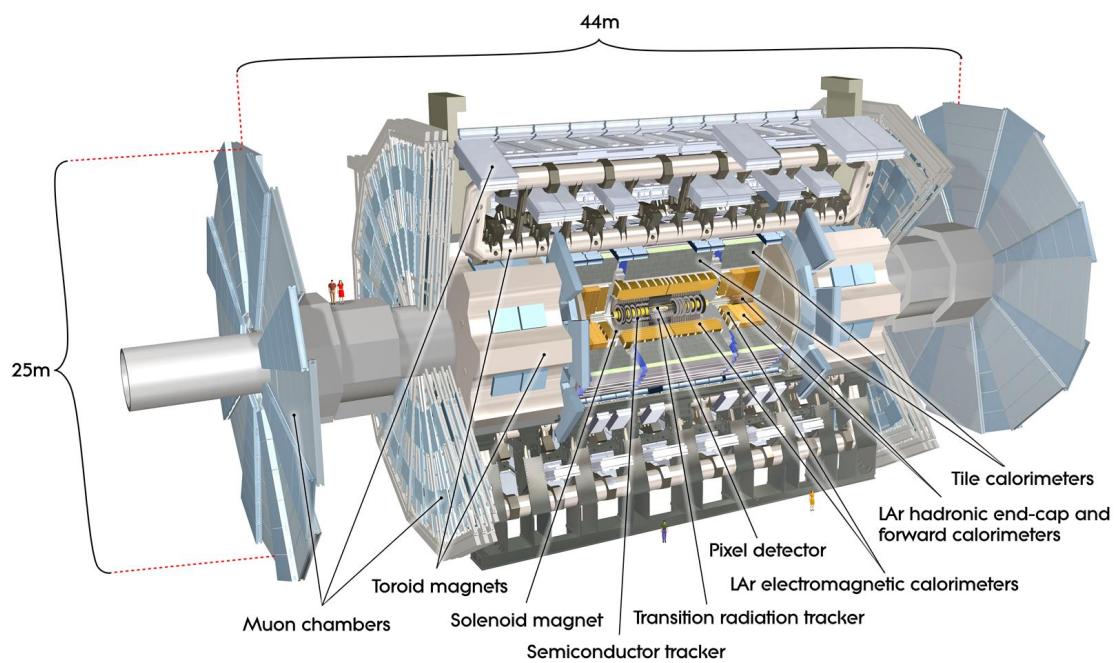


Figure 2.2: Cut-away view of the ATLAS detector with its sub-detectors system.
Reference [57].

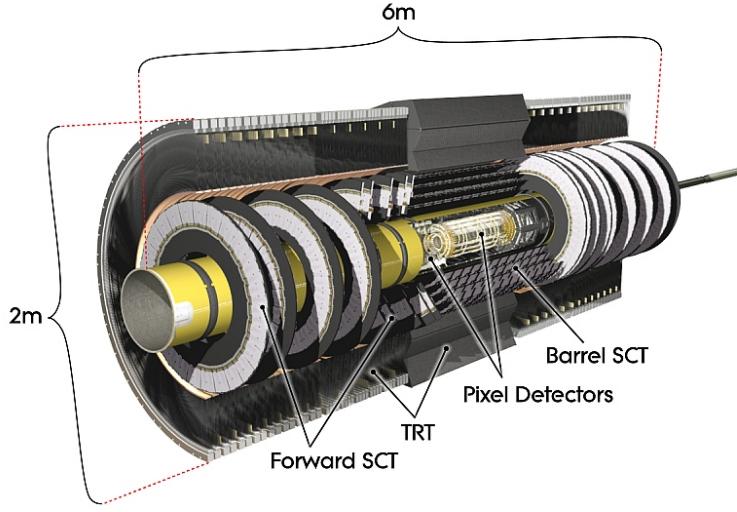


Figure 2.3: Cut-away view of the ATLAS Inner Detector tracker. Reference [57].

the rapidity tend to the pseudorapidity η :

$$\eta = 1/2 \cdot \ln \left(\frac{\theta}{2} \right) \quad (2.2)$$

Given the symmetry of the ATLAS detector with respect to the interaction point, the detector is divided in two regions called *barrel* for $|\eta| \lesssim 1.5$ (depending on the considered detector) and *endcap* for larger η . In ATLAS the angular separation between two particles is commonly measured by $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

2.2.2 The Inner Detector

The Inner Detector (ID) performs track reconstruction and momentum measurements of charged particles, it has a total length of 5.3 m and a diameter of 2.5 m . The momentum measurement is performed by measuring the tracks curvature in a 2 T magnetic field generated by a super-conducting solenoid. The layout of the Inner Detector is illustrated in Figure 2.3, it consist of three independent detector modules with fine granularity covering the region $|\eta| < 2.5$. The innermost of these

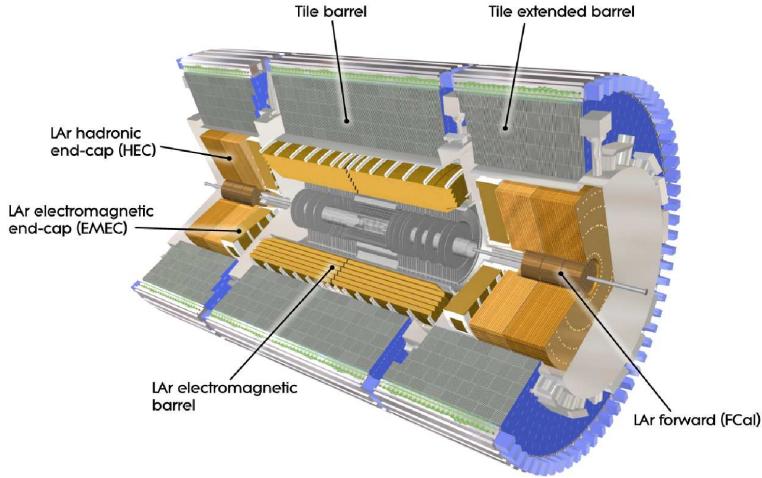


Figure 2.4: Cut-away view of the ATLAS calorimeter system. Reference [57].

detectors is the Pixel detector, which consist of three cylindrical layers of pixel silicon sensors in the barrel and three disks in the endcap region. The closest layer of pixels to the beam pipe is referred to as the B-layer. The spatial resolution of the pixel sensors is $10\ \mu m$ in the transverse and $115\ \mu m$ in the longitudinal direction with respect to the beam pipe.

The Semi-Conductor Tracker (SCT) surrounds the Pixel detector with four cylindrical layers of silicon microstrip sensor in the barrel and nine disks in the endcap region. The spatial resolution achieved by the SCT is of $17\ \mu m$ in the transverse and $590\ \mu m$ in the longitudinal direction respectively.

The outermost detector module is the Transition Radiation Tracker (TRT). It is composed of $4\ mm$ diameter Kapton straw tubes with a tungsten wire in their centre. The tube is filled with a gas mixture which allows the detection of transition radiation photons. This detector can only measure position along the transverse direction.

2.2.3 The Calorimeter System

An illustration of the ATLAS calorimeter system is shown in Figure 2.4, it consist of an electromagnetic calorimeter (EM) surrounded by an hadronic calorimeter. These calorimeters cover the range $|\eta| < 4.9$ using different techniques suited to the widely varying radiation environment over this large η -range. Both these calorimeters are sampling calorimeters, they are builded alternating active material which performs the detector response and a passive absorber. The total detector thickness at $\eta = 0$ is 9.7 interaction length.

The EM Lar calorimeter is ideally suited for precision measurements of electrons and photons. It uses lead as absorber material and liquid argon as active material. It extends up to $|\eta| < 3.2$.

The hadronic calorimeter has a coarser granularity with respect the EM calorimeter and it is suited for jet reconstruction and missing transverse energy measurements. The hadronic calorimeter is divided in three sub-detectors which make use of different technology to cope with the changing radiation environment as a function of η . The Tile calorimeter covers region in pseudorapidity up to $|\eta| < 1.7$, it uses scintillating tiles as active material and steel as absorber. In the forward region ATLAS is instrumented with a Lar hadronic endcap calorimeter (HEC), which extends up to $|\eta| < 3.2$ and uses argon as active material and copper as absorber. The region in $3.1 < |\eta| < 4.9$ is instrumented instead with a liquid argon Forward CALorimeter (FCAL), which is divided in three modules, the closest to the interaction point uses copper as absorber material, while the other two uses tungsten.

2.2.4 The Muon Spectrometer

The muon spectrometer is instrumented with separate high-precision tracking and trigger chambers. The measure of muon momenta is performed by reconstructing the track curvature in an intense magnetic field produced by the large superconducting air-core toroid magnets. The layout of the muon spectrometer is shown in Figure 2.5.

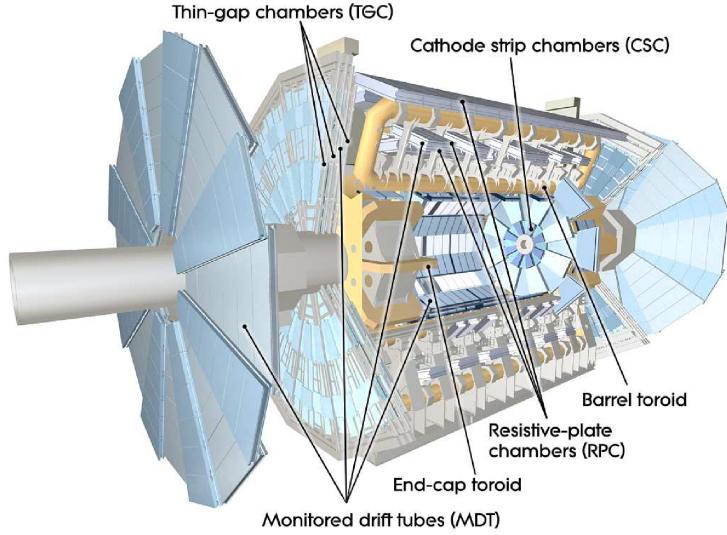


Figure 2.5: Cut-away view of the ATLAS muon spectrometer system. Reference [57].

Precision measurement of the track coordinates in the principal bending direction of the magnetic field is provided by Monitored Drift Tubes (MDTs) up to $|\eta| < 2.7$. Given the demanding rate and background conditions at large pseudo-rapidities, $2 < |\eta| < 2.7$, the innermost MDT layer is replaced by Cathode Strip Chambers (CSCs), which are multiwire proportional chambers with cathodes segmented into strips. Precise muon momentum measurement is achieved for muons with momenta up to 1 TeV. The best momentum resolution, 3-4%, is achieved for muons with transverse momenta ~ 100 GeV, while resolution of $\sim 10\%$ are reached for muons with momenta up to 1 TeV.

The trigger system covers the pseudorapidity range $|\eta| < 2.4$. Resistive Plate Chambers (RPCs) are used in the barrel and Thin Gap Chambers (TGCs) in the end-cap regions for the trigger information.

2.2.5 The Trigger System

The trigger system has three distinct levels: L1, L2, and the event filter (EF). Each trigger level refines the decisions made at the previous level and, where necessary, applies additional selection criteria.

The L1 trigger searches for high transverse-momentum muons, electrons, photons, jets, and τ leptons decaying into hadrons, as well as large missing and total transverse energy. Its selection is based on information from the set of detectors described previously. The L1 trigger defines in the interesting events one or more Regions-of-Interest (RoI), given by $\eta - \phi$ coordinates of interesting feature of the event.

The L2 selection is seeded by the RoI information provided by the L1 trigger, it uses the full granularity and precision of all the available detector data within the RoIs. The L2 triggers are designed to reduce the trigger rate to approximately 3.5 kHz.

The final stage of the event selection is carried out by the event filter, which reduces the event rate to roughly 200 Hz. Its selections are implemented using the offline analysis and reconstruction procedures described in Chapter 3.

2.2.6 Luminosity Measurement

A precise measurement of the recorded integrated luminosity is extremely important for all the physics measurements of the ATLAS physics program.

Several techniques are employed in ATLAS for the measure of the luminosity. The most relevant detectors that monitor the luminosity are the Inner Detector, the BMC and the LUCID detectors. For a detailed description of the ATLAS luminosity measurements and its performance see Reference [63]. The Inner Detector measures the luminosity by the average reconstructed proton-proton interaction per bunch crossing. The LUCID detector surrounds the beampipe on both sides of the interaction point at a distance of 17 m, providing with its Cherenkov detectors the measures of the particle flux from the interaction point in a very forward region. The BCM counts the number of collision per bunch crossing providing an

independent luminosity estimate.

Chapter 3

Reconstruction of Physics Objects

The raw ATLAS data, containing detectors signals of all read-out channels, need to undergo several reconstruction steps before they can be analyzed. Reconstruction software is developed within the object oriented ATLAS reconstruction software framework ATHENA [92], which allow for reconstruction and identification of various objects corresponding to physics particles traversing the detector.

This chapter briefly describes the ATLAS reconstruction strategies for physics objects relevant for the analysis presented in this thesis. For a detailed overview of the ATLAS detector reconstruction software see [93].

3.1 Reconstruction of Charged Particle Tracks

The reconstruction of charged particles tracks and interaction vertices is based on the measurements in the inner detector which allow for the reconstruction of tracks within the pseudorapidity range of $|\eta| < 2.5$. A track is characterized by its four-momentum vector and two impact parameters: d_0 , i.e., the distance of closest approach between the track and the interaction point in the transverse plane and z_0 , i.e. the z coordinate of the track calculated at the same point of closest approach.

Tracks are reconstructed by the inner detector track reconstruction software [94]. At first raw data from the pixel and SCT detectors are transformed in three-dimensional space points (so called “hits”), while the TRT detector information is translated into drift circles. Subsequently, track seeds are formed from a combination of space-points in the three pixel layers and the first SCT layer. These seeds are then extrapolated through the SCT to form track candidates from all hits on the track path. The track candidates are obtained by a fit trough all hits using a *Kalman filter* algorithm [95]. Ambiguities in the association of the hits to the track are resolved by this fitting procedure and tracks produced by a random association of hits are rejected. The selected tracks are then extrapolated to the TRT and finally refitted using the full information of all three tracking detectors. In order to improve the tracking efficiency for secondary tracks from photon conversion or decays of long-lived particles (like kaons), a complementary algorithm [94] searches for unassociated track segments in the TRT, these segments are then extrapolated towards the SCT and the pixel detector in a similar manner as in the default algorithm. All tracks with $p_T > 100$ MeV are considered for physics analysis.

3.2 Vertex Reconstruction

The vertex reconstruction algorithm and its performance are described in detail in [93,96] and only briefly summarized here. The vertex finding algorithm selects a set of well reconstructed tracks and generates a vertex seed according to the average value of the tracks z coordinate. The z coordinate of the tracks is computed relative

to the expected average position of the collision point. An *adaptive vertex fitting* algorithm [97] determines the vertex position based on the vertex seed and on the tracks around it via a χ^2 fit. Based on this fit, tracks that are incompatible with the found vertex by more than seven standard deviations are used to seed the next vertex. The procedure is performed iteratively until either all tracks are associated to a vertex or no additional vertex can be found. The performance of this procedure depends on the expected position of the average interaction point which is monitored during LHC data taking and is computed in intervals of a few minutes as described in [98].

The vertex with the largest sum of transverse momentum of all associated tracks is identified as the *primary vertex* (PV), corresponding to the interaction point of the hard scattering process in the event. All other vertices in the event are assumed to result from minimum bias interactions and are called *pile-up* vertices. In data recorded during 2012, there were on average 21 multiple interactions occurring per bunch crossing. Such a high vertex multiplicity strongly affects the ambient energy density in the event, such that an accurate pile-up description in simulation is crucial for the modelling of physics processes. In ATLAS, events are simulated assuming various pile-up conditions and weighted such to reproduce the observed average number of interactions per bunch crossing.

3.3 Electron Reconstruction and Identification

Electron are reconstructed and identified by combining EM calorimeter and inner detector measurements. The corresponding dedicated algorithm is described in [99]. The electron candidate is reconstructed as a clusters of EM calorimeter cells which is matched to a track in the inner detector. Special care during the matching is taken to account for Bremsstrahlung losses of the charged particle. The electron energy is computed as a weighted average between the cluster energy and the track momentum. Several corrections are applied to take into account energy losses in the material of the inner detector and effect of electromagnetic shower leakage. The electron direction is defined by the corresponding track parameters.

Further identification criteria are applied to electron candidates to reduce contaminating contribution of photon conversions and hadronic jets. Three different identification criteria are provided based on a multi-variate analysis program (TMVA [100]) and several selection criteria :

- Loose electron identification: variables related to the shape of the electromagnetic shower and to the amount of the hadronic leakage are used in a multi-variate analysis program.
- Medium electron identification: the total shower width and the difference between the largest and second largest energy deposit are considered in a multi-variate analysis program in addition to the loose variables. Furthermore stricter track matching requirements are imposed.
- Tight electron identification: in addition to medium requirements, converted photons are rejected by requiring a hit in the innermost layer of the inner detector. Furthermore, the number of TRT hits associated to the electron is employed as additional variable in the multi-variate analysis program.

The performances of the electron identification are measured with several calibration data samples (using events with leptonic decays of W , Z bosons and J/ψ meson) and compared to simulation [101]. Corresponding corrections of the simulated electron identification efficiency are measured and applied as p_T and η dependent weight to each simulated electron candidate. Additional corrections are applied to the energy scale and energy resolution of simulated electrons to match the one in data according to [102]. Systematic uncertainties on the measure of the identification efficiency ranges from 1-2% depending on the transverse momentum of the electron, while uncertainties on the measure of the energy scale and resolution range approximately from 0.3-3% depending on η . Finally, the electrons used in the presented analysis are rejected if they are detected in a region of the calorimeter with readout problems or suffering from high noise.

Prompt electrons, originating from the decay of a resonance like the Z^0 boson or the Higgs boson are very likely to be *isolated*, i.e. there is little particle activity expected in their surroundings. This is in contrast to electrons originating from

hadron decays, which instead will be likely to be surrounded by a jet of particles. Two isolation variables are defined to account for the activity in a cone of size $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ around the electron candidate:

- Track isolation, $p_T^{cone} = \sum_{\Delta R < 0.4} p_{T,i}$, is the scalar sum of the transverse momenta $p_{T,i}$ of all tracks i in a cone $\Delta R \leq 0.4$ around the electron direction. The electron track itself is not counted here.
- Calorimeter isolation, $E_T^{cone} = \sum_{\Delta R < 0.2} E_{T,i}$, is the scalar sum of transverse energies $E_{T,i}$ of each topological cluster i in a cone $\Delta R \leq 0.2$ around the electron direction. Clusters associated to the electron itself are not counted. The value of this variable is corrected as a function of the vertex multiplicity in the event in order to account for the pile-up effects and therefore to assure a constant electron selection efficiency for each event.

3.4 Muon Reconstruction

ATLAS employs a variety of strategies for the reconstruction and identification of muons, relying primarily on the tracking in the muon spectrometer and supplemented in most cases with the tracking in the inner detector and the energy deposit in the calorimeter. A detailed description of the muon reconstruction algorithms and their performance is reported in [93]. In the following only the muon reconstruction strategy relevant for this thesis is described.

The STACO *combined* muon algorithm [103] associates tracks found in the muon spectrometer with the corresponding inner detector track and calorimeter energy deposit. At first, track segments are reconstructed in each of the three muon stations and are linked together to form a track. The muon spectrometer track is extrapolated to the inner detector taking into account the energy loss and multiple scattering in the calorimeters. The extrapolated track is matched with an inner detector track via χ^2 -matching. Finally, a statistical combination of the inner detector and muon spectrometer tracks is performed to obtain a combined muon track.

Muon identification efficiency, momentum scale and momentum resolution are evaluated in [104] where performance is compared with prediction from simulation. A set of corrections on the muon momentum scale, resolution and identification efficiency is applied to simulation to ensure a good agreement with data. Uncertainties on these corrections are of the order of a fraction of percent.

Isolation variables, are derived and employed in a similar manner as for electrons. The only difference is the use of calorimeter clusters with fixed size (so-called towers) instead of the topological cells in the definition of E_T^{cone} . Pile-up corrections similar to those employed for electrons are used for muons as well.

3.5 Jet Reconstruction and Energy Calibration

Jets are reconstructed by means of the FastJet package [105], which provides a broad range of jet finding algorithms and analysis tools. In the following jet reconstruction methods relevant for the analysis presented in this theses are briefly described, for more detail see [93].

In general, jets may be reconstructed out of any set of four vector objects. In ATLAS, the jet reconstruction relies most commonly on energy deposit measured by the calorimeters. Calorimeter cells are grouped together by a clustering algorithm forming the so called *topological clusters* [106], i.e. three-dimensional clusters representing the energy depositions of the shower particles. The clustering procedure starts with seed calorimeter cells with a signal-to-noise ratio greater than a certain threshold. All nearby cells are combined with the seed cells if they pass a second, lower, signal-to-noise ratio threshold.

Each topological cluster is then used as input to the *anti- k_t* algorithm [107]. The algorithm defines a metric to assess distances between the clusters i and j :

$$d_{ij} = \min\left(\frac{1}{k_{t,i}^2}, \frac{1}{k_{t,j}^2}\right) \cdot \frac{\Delta R_{ij}^2}{R^2} \quad \text{and} \quad (3.1)$$

$$d_i = \frac{1}{k_{t,i}^2}, \quad (3.2)$$

where $k_{t,i}$ is the p_T of the cluster i and $\Delta R_{ij}^2 = \sqrt{\Delta\phi_{ij}^2 + \Delta\eta_{ij}^2}$ is the angular

distance between the two cluster i and j . For the presented analysis the distance parameter R is chosen to be $R = 0.4$. If the distance d_{ij} between two cluster i and j is smaller than d_i , the clusters are grouped together and their four momenta are summed. Otherwise they are kept as a single entity. The clustering procedure is iterated until no further cluster can be merged. The metric is designed such that high- p_T clusters will accumulate the soft activity surrounding them, therefore leading to conical jet shapes.

Given the high pile-up environment of the LHC, it is important to distinguish jets originating from the hard scattering process and those related to pile-up interactions. For this purpose, each jet is characterized by a so-called *jet vertex fraction* (JVF). The value of the JVF is defined as the p_T -weighted fraction of inner detector tracks pointing to the primary vertex among all tracks associated to the corresponding jet:

$$\text{JVF} = \frac{\sum_{\substack{\text{PV-tracks} \\ \text{tracks}}} p_{T,i}}{\sum_{\text{tracks}} p_{T,i}} \quad (3.3)$$

The jet vertex fraction can only be defined for jets within inner detector coverage of $|\eta| < 2.5$, while the calorimeter jet reconstruction itself is possible up to $|\eta| < 4.5$.

Energy Calibration The ATLAS calorimeters are calibrated using test beam electrons [108]. However, the response of the calorimeters to electromagnetic showers differs from the response to hadronic showers. A dedicated jet energy scale (JES) calibration is therefore performed based on simulation [109]: the jet energy is corrected to correspond, on average, to the simulated energy of the corresponding hadronizing parton. The jet direction is also corrected such to point to the primary vertex instead to the origin of the ATLAS detector coordinate system. A set of corrections is evaluated to take into account for pile-up effects [110, 111]. Simulated jet resolution is also corrected to better describe the data [112]. Finally, several jet energy scale corrections are applied for a better agreement between data and simulation. These corrections are determined with 2011 ATLAS data using several techniques exploiting the transverse momentum balance between a jet and a reference object such as a photon, Z boson or another jet [109, 113]. System-

atic uncertainties on the jet energy scale and resolution due to imperfect Monte Carlo modelling are evaluated to range from 1-6% depending on the jet p_T and pseudorapidity.

3.6 Identification of Jets from b-quarks

The typical decay length of a b-hadron in the ATLAS detector is of the order of few millimetres. Exploiting the high precision of the inner detector tracker it is possible to discriminate between the jets originating from b-quarks and those from other quarks or gluons (also referred to as light-jets). The identification technique used for this purpose is called *b-tagging* and the identified b-tagged jets are referred to as b-jets.

Several b-tagging algorithms have been developed in ATLAS. The relevant algorithms for this thesis are briefly described in what follows, for a more detailed description see [93]. The b-tagging algorithm starts by associating tracks to the jets based on their angular distance ΔR to the jet. The mentioned tracks should satisfy strict selection criteria aimed to ensure a good track quality and to reject tracks likely to come from strange hadron decays or photon conversions. The discrimination between the b-jet and other jets is based on simulated distributions of several discriminating variables. Given the relatively high mass of b-hadrons, the tracks associated to a b-jet will have a relatively wide spread of impact parameter values. This feature is used by the IP3D b-jet tagging algorithm, where a corresponding discriminating variable is defined based on impact parameter significance¹ of all tracks associated to the jet. An alternative approach, used by the *SV1* algorithm, is instead to search for inclusive secondary vertex formed by the decay products of the b-hadron. The search includes also the subsequent charm hadron decays. Another algorithm, called JetFitter [115], relies instead on the direction of the jet to fully reconstruct the decay chain of a b-hadron, under the assumption that the decayed particles will be emitted along the jet axis. The outputs of each of

¹ The significance is defined as the value of the impact parameter divided by the error on its measurement.

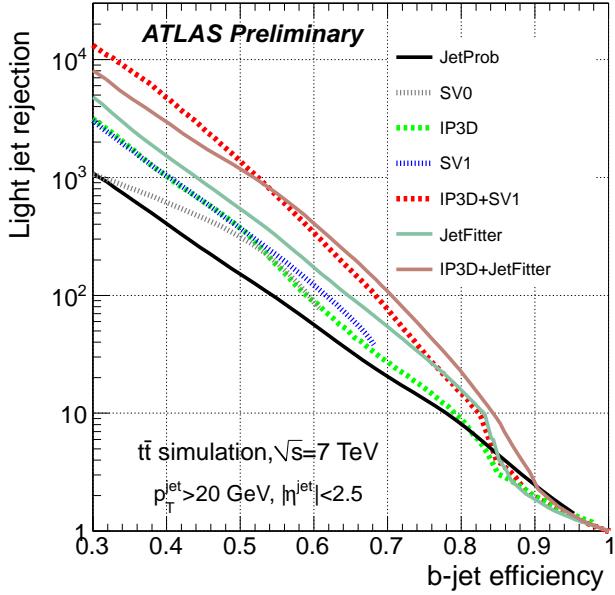


Figure 3.1: Light-jet rejection as a function of the b-jet tagging efficiency for several different tagging algorithms [116], obtained with simulated $t\bar{t}$ events. The rejection is defined as the inverse of mistagging rate of light jets.

these three algorithms gives a measure of the probability that the reconstructed jet originates from a b-quark. Finally, the outputs of the three described algorithms are combined based on an artificial neural network multivariate program [100] to maximise the discriminating power. The output of this neural network is referred to as *MV1* tagger and is used for the Higgs boson search presented in this thesis.

The performance of the mentioned algorithms is evaluated in data selecting $t\bar{t}$ events and compared to simulation [116]. Figure 3.1 shows the b-tagging efficiency as a function of the inverse of the light-jet mistagging rate for different b-tagging algorithm on $t\bar{t}$ simulated events. The tagging efficiency $\epsilon_b^{t\bar{t}}$ obtained from $t\bar{t}$ events is used to define several b-tagging working points. Corrections due to non perfect modelling of the b-tagging performance are evaluated by means of several methods in [117, 118] and used to determine event weights for simulated events. The uncertainties on these corrections range from 5-10% depending on the

p_T and pseudorapidity of the jet.

3.7 Tau-Jet Reconstruction

The reconstruction of jets originating from hadronically decaying τ leptons (in the following τ -jets) is described in detail in [93]. A τ -jet candidate is seeded by reconstructed calorimeter jets with $p_T > 10$ GeV and $|\eta| < 2.5$. Tracks are then associated to the jet and a combination of the tracking and calorimeter information is performed. τ -jets can be distinguished from other jets by their low track multiplicity and a narrower clustering of energy deposit in the electromagnetic and hadronic calorimeters. The τ -jet identification in ATLAS is based on a Boosted Decision Trees (BDT) multivariate procedure [121]. One BDT discriminant has been developed to discriminate τ -jets from quark and gluon initiated jets and a separate one was developed to reject electrons.

3.8 Missing Transverse Energy

The missing transverse energy, E_T^{miss} , is the vectorial sum of the transverse momenta of all the physics objects and calorimeter cells in the event changed of sign. Undetected particles, such as neutrinos, lead to an unbalance of the total transverse momentum, thus, to a non zero value of E_T^{miss} .

Reconstruction and calibration of E_T^{miss} with the ATLAS detector is described in detail in [119]. The missing transverse energy measurement relies on the reconstruction of all physics objects in the event, it includes: muons and their energy deposits in the calorimeter, electrons, jets (weighted by their corresponding JV), inner detector tracks (to take into account low- p_T particles which are not well reconstructed in the calorimeters), photons and τ leptons. The calorimeters cells are calibrated depending on the physics object with which they are associated. The transverse energy of cells not associated to any object is taken into account in the so called “CellOut” contribution. This contribution, together with the one related to jets with $10 < p_T < 20$ GeV are referred to as the *soft term* of the missing

transverse energy. The soft term is found to be very sensitive to pile-up. In order to reduce the impact of pile-up, the soft term is scaled by the corresponding soft-term-vertex-fraction (STVF), which is calculated in the same way as JVF for jets.

A detailed description of the performance of the E_T^{miss} reconstruction and calibration may be found in [120].

3.9 Overlap Removal

Reconstruction of physics objects defined in the previous section may sometimes be ambiguous. For example, a τ -jet is always reconstructed also as a common jet. To avoid double counting of the physics objects originating from the same particle, an overlap removal procedure is performed. A match between physics object of different sort is seeded in a cone of $\Delta R < 0.2$. If the matching occurs, the object with the lowest ranking is removed from the event. Physics object are ranked according to the following order, starting with the highest rank: muon, electron, τ -jet and finally common jets.

3.10 Trigger

The ATLAS trigger system [122] consists of three stages. The Level-1 (L1) trigger is an hardware trigger which reduces the event rate to approximatively 100 kHz and selects the Regions of Interest (RoI) to be further investigated by the High Level Trigger (HLT). The HLT comprises the Level-2 (L2) trigger employing fast reconstruction algorithms and the Event Filter (EF) exploiting the full ATLAS event reconstruction.

In the presented search two triggers are employed: an electron EF trigger, which selects events containing an electron with $p_T > 24$ GeV and a combined muon-electron EF trigger, which requires the presence of a muon with $p_T > 8$ GeV and an electron with $p_T > 12$ GeV in the event. Detailed description of the muon and electron triggers can be found in [123, 124]. Trigger efficiency for both triggers is

evaluated in data selecting Z candidate events and compared with prediction from simulation. Corrections are derived as function of the lepton pseudorapidity and transverse momentum to match the simulated trigger efficiency with the one in data [123, 124].

3.11 Truth Particles

In case of a simulated event, the ATLAS reconstruction software provides the generator-level information regarding simulated particles (also called *truth particles*). Their identity, kinematic properties, decays and interactions are stored in the event based on the conventions defined in [125]. A particle is defined stable if $c\tau > 1$ m, where τ is its mean life time. Particle emerging from interaction with the detector are excluded from this definition. Each particle has an associated “bar-code” which is a unique identifier for that particle in that event. Jets reconstructed with the common ATLAS reconstruction algorithm from stable particles are called *truth jets*.

Chapter 4

Search for neutral MSSM

Higgs Bosons in

$A/h/H \rightarrow \tau^+\tau^- \rightarrow e\mu + 4\nu$ decays

In light of the recent discovery of a Higgs boson with mass of ~ 126 GeV at the LHC [16,17], it remains an open question whether this new particle is the only missing piece of the electroweak symmetry breaking sector or whether it is one of several Higgs bosons predicted in theories that go beyond the SM. The most recent measurements [126–129] of its properties shows this new boson to be, within experimental uncertainties, fully compatible with the SM Higgs boson. Nevertheless, such a new particle can also still be accommodated within several theories beyond the standard model (BSM). Among all of them, Supersymmetry is a theoretically favoured scenario as the most predictive framework beyond the Standard Model.

This chapter presents the search for the neutral MSSM Higgs bosons decaying into pairs of tau leptons in the fully leptonic final state, published in Ref. []¹ as a part of the search for the neutral MSSM Higgs bosons in all final states of the tau leptons decay. The search is based on 20.3 fb^{-1} of data at a centre-of-mass energy of $\sqrt{s} = 8 \text{ TeV}$ recorded by the ATLAS experiment during 2012. This chapter is organised as follows: a brief summary of the

¹to Sandra: I'll remove this sentence if Conf note wont be ready in time

MSSM Higgs sector and an introduction to the analysis strategy is given in Section 4.1, while the event selection and categorization are described in Section 4.2. Section 4.3 describes the estimation of the backgrounds and in Section 4.4 methods to evaluate systematic uncertainties are discussed. Finally, in section 4.5, an overview of the statistical methods employed along with the corresponding result of the search are presented.

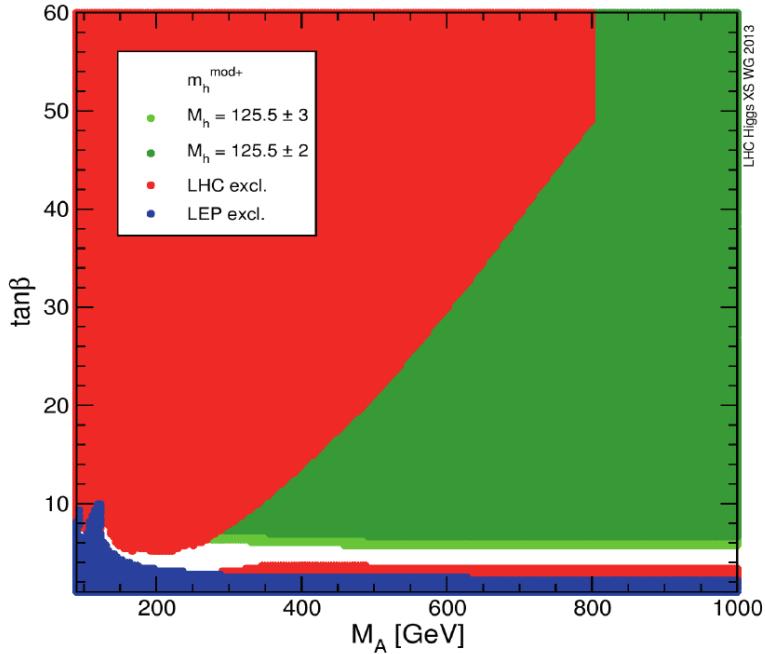


Figure 4.1: Excluded and allowed regions of the $m_A - \tan\beta$ parameter space for the m_h^{mod+} MSSM benchmark scenario. Excluded regions are determined based on direct Higgs boson searches at LEP (blue) and LHC (red). The two green bands correspond to the parameter regions which are compatible with the assumption that the lightest MSSM Higgs boson, h , has a mass respectively of $M_h = 125.5 \pm 2$ (dark green) or 125.5 ± 3 GeV (light green). For more detail see [4].

4.1 Introduction

4.1.1 The Higgs Sector in the MSSM

In the Minimal Supersymmetric extension of the Standard Model (MSSM) [36,37] the Higgs sector is composed of two Higgs doublets of opposite hyper-charge, resulting in five observable Higgs bosons: two of these are neutral and CP -even (h, H), one is neutral and CP -odd (A) and two are charged (H^\pm). At tree level their properties such as masses, widths and branching ratios can be predicted in terms of only two parameters, often chosen to be the mass of the CP -odd Higgs boson

m_A and the ratio of the vacuum expectation values of the two Higgs doublets $\tan \beta$ (for more details see chapter ??). The MSSM predicts the existence of a Higgs boson with properties that resemble those of a SM Higgs boson in large regions of its parameter space. This is usually the case for the lightest Higgs boson, h , while the other two, H and A , tend to be degenerate in mass and decouple from gauge bosons. On the other hand, the couplings of the latter two Higgs bosons with down (up) type fermions are enhanced (suppressed) proportionally to the value of $\tan \beta$, meaning that for large $\tan \beta$ bottom-quark and τ lepton will play an important role for the Higgs bosons production and its decays.

The two most relevant MSSM Higgs bosons production mechanisms at the LHC are gluon fusion, $gg \rightarrow A/H/h$, and the production in association with b -quarks, $pp \rightarrow b(b)A/h/H$, the latter becoming increasingly important for large values of $\tan \beta$. These two are the only production mechanisms considered in this analysis. Assuming there are no decays into supersymmetric particles since these are too heavy, the favoured neutral MSSM Higgs bosons decay mode is the decay into a pair of b -quark and anti-quark, $A/h/H \rightarrow b\bar{b}$. This is followed, for the CP-odd A and CP-even H Higgs bosons, by the decay into pairs of τ leptons. Given that it is very difficult to distinguish the former decay from the large $b\bar{b}$ background, the decay mode $A/h/H \rightarrow \tau^+\tau^-$ provides the highest sensitivity in the search for neutral MSSM Higgs bosons.

Searches for neutral MSSM Higgs bosons have been performed at LEP [64], the Tevatron [65] and the LHC [66, 67]. In the following the search for the neutral MSSM Higgs bosons in the final state $A/h/H \rightarrow \tau^+\tau^- \rightarrow e\mu+4\nu$ is presented. This search is complementary to the searches in other $\tau^+\tau^-$ final states characterised by the presence of one or two hadronically decaying τ leptons. Despite of the fact that the $\tau\tau$ branching ration in $e\mu + 4\nu$ is only 6%, this decay channel provides a sensitivity to the signal comparable to those in other $\tau\tau$ final states, especially for low m_A values. This is mainly due to the high transverse momentum threshold at the trigger level for hadronically decaying τ leptons.

As it is impractical for an experimental search to explore the full parameter space of the MSSM, which has many free parameters, several benchmark scenarios

are introduced by fixing all except m_A and $\tan\beta$ parameters to values typical for most interesting physics cases. With the recent Higgs boson discovery, benchmark scenarios of the MSSM have been updated to accommodate for new experimental constraints. As an example, Figure 4.1 shows the currently excluded and allowed regions of the MSSM parameter space for the m_h^{mod+} updated benchmark scenario. In this scenario a large region of the $m_A - \tan\beta$ parameter space is compatible with the assumption that the observed Higgs boson correspond to the supersymmetric CP-even neutral Higgs boson, h . A large part of this parameter space is still experimentally unexplored, this is a strong motivation to pursue the search for additional neutral MSSM Higgs bosons.

4.1.2 Signal and Background Processes

Signal events in which the neutral MSSM Higgs bosons decay through $A/h/H \rightarrow \tau^+\tau^- \rightarrow e\mu + 4\nu$ process are characterised by the presence of one electron and one muon of opposite charge. These two leptons are isolated and have relatively high transverse momenta. In addition, four neutrinos contribute to the missing transverse energy in the event. Figure 4.2 shows leading order Feynman diagram for the two considered signal production modes, gluon fusion and in association with b -quarks. The presence (absence) of a b -jet in the final state serves as a main characteristic for the event categorization in the latter (former) case, as described later on.

The described signal topology is common to several other known SM background processes which in general have higher cross sections than the sought signal. The dominant background processes are the $Z/\gamma^* \rightarrow \tau^+\tau^-$ production either via Drell-Yan process or in association with jets and the top quark production ($t\bar{t}$ and single top quark production). Additional significant background contributions originate from the dibosons production (WW , WZ , ZZ) and QCD multi-jet events with non-prompt leptons from hadron decay. Vector boson production ($W \rightarrow \ell\nu$ or $Z \rightarrow \ell\ell$, where $\ell \equiv e, \mu$) in association with jets is also considered, but has small impact on the total background contamination. Examples of leading order Feynman diagrams for the dominant background processes are shown in

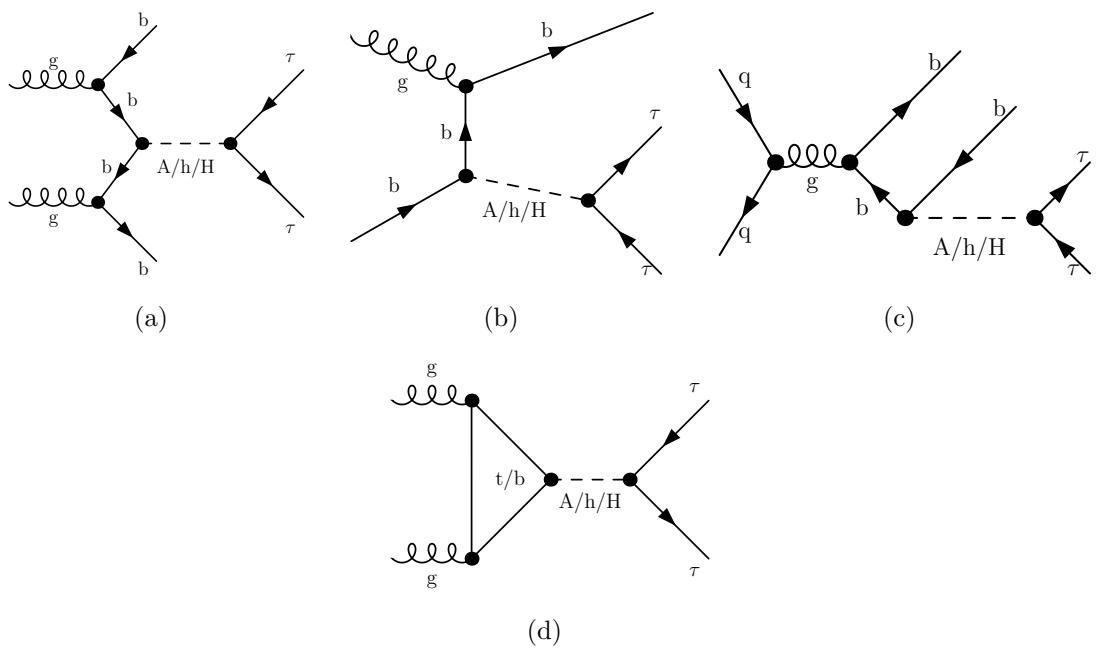


Figure 4.2: Feynman diagram for the production of the neutral MSSM Higgs bosons in association with b -quarks (a,b,c) and via gluon fusion (d) process, subsequent decay in tau lepton pairs is considered.

Process	Cross-section (pb) [\times BR]
Signal ($m_A = 150$ GeV, $\tan \beta = 20$, m_h^{mod} scenario)	
$gg \rightarrow A/h/H \rightarrow \tau\tau \rightarrow e\mu + 4\nu$	0.24/0.20/0.95
$pp \rightarrow b\bar{b}A/h/H \rightarrow \tau\tau \rightarrow e\mu + 4\nu$	0.53/0.05/0.49
Backgrounds	
$W \rightarrow \ell\nu + \text{jets}$	12.22×10^3
$Z/\gamma^* \rightarrow \ell\ell + \text{jets}$	5.5×10^3
$t\bar{t} \rightarrow \ell\ell + X$	137.3
Single top quark ($t-$, $s-$ and $Wt-$ channels) $\rightarrow \ell + X$	28.4, 1.8, 22.4
Dibosons (WW, WZ and ZZ) $\rightarrow \ell\ell + X$	20.6, 6.8, 1.55

Table 4.1: The cross sections multiplied by the relevant branching ratios (BR) for signal and the considered background processes. The symbol ℓ stands for $\ell = (e, \mu, \tau)$. Signal cross sections are calculated for the m_h^{mod} scenario assuming $m_A = 150$ GeV and $\tan \beta = 20$. The masses of the other two neutral MSSM Higgs bosons are in this case $m_H = 151$ GeV and $m_h = 125$ GeV.

Figure 4.3. The production cross sections times the relevant branching fraction for signal and background processes are summarized in Table 4.1.

4.1.3 Analysis Strategy

In this thesis a search for the MSSM $A/h/H \rightarrow \tau^+\tau^- \rightarrow e\mu + 4\nu$ decays is presented. The $ee + 4\nu$ and $\mu\mu + 4\nu$ final states are not considered since a large background contribution is expected from $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ decays, respectively, such that the sensitivity of the search in these final state is significantly reduced.

Candidate events are selected based on the topological properties of the Higgs boson production and decay. The presence of exactly one electron and one muon is required in each event. Electron and muon are required to be isolated and of opposite electrical charge. The events are categorized into two mutually orthogonal event categories. In the so called *b-vetoed* event category, the absence of

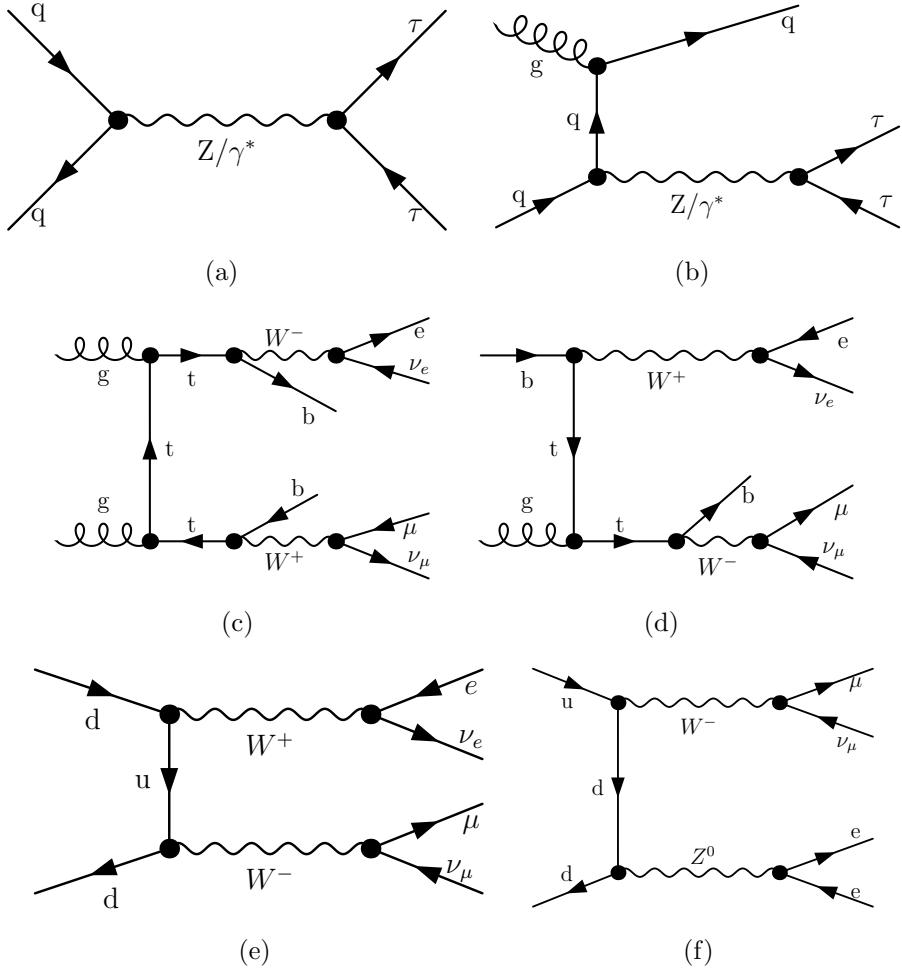


Figure 4.3: Examples of tree level Feynman diagrams for the production and decays of the most relevant backgrounds. The production of $Z/\gamma^* \rightarrow \tau^+\tau^-$ either via Drell-Yan process or in association with jets is shown in (a) and (b), top quark pair and single top quark production in (c) and (d), while examples of WW and WZ production are shown in (e) and (f) respectively.

any b-tagged jets is required, thus searching mainly the signal produced via gluon fusion. The main background process in this category is $Z/\gamma^* \rightarrow \tau\tau$. In contrast, the presence of exactly one b-tagged jet is required in the so called *b-tagged* event category, searching predominantly signal produced in association with b-quarks. The requirement of a b-jet in the final state suppresses the $Z/\gamma^* \rightarrow \tau\tau$ background, consequently, $t\bar{t}$ and single top quark production are the main background processes in this event category. Further selection criteria are introduced in both event categories, these are optimised to enhance the signal produced by the corresponding production mode.

The search is performed within the MSSM m_h^{mod} benchmark scenario scanning the $m_A - \tan\beta$ plane in the ranges $90 \leq m_A \leq 300$ GeV and $5 < \tan\beta < 60$. The prediction of the signal event yields and kinematical distributions are evaluated by simulation. The contribution of the dominant $Z/\gamma^* \rightarrow \tau\tau$ background process is measured in a dedicated signal-depleted control data sample, in order to reduce the systematic uncertainties of the simulation. Similarly, the QCD multi-jet background contribution is also estimated from dedicated data control sample since this background process is hardly modelled by simulation. Contribution of all other background processes is estimated from simulation. The modelling of the background processes is validated using different signal-depleted validation data samples and good agreement is found.

The systematic uncertainties on cross section calculations and the modelling of the detector response are taken into account for simulated signal and background processes. For background processes that are measured with data, the uncertainties of the corresponding measurement methods are evaluated.

The final statistical interpretation of the data is based on the comparison of the observed $\tau\tau$ invariant mass distributions with the prediction of the background-only and signal-plus-background hypothesis. Exclusion limits on the signal production are set by means of a binned profiled likelihood ratio test statistic. The limits are interpreted within the MSSM m_h^{mod} scenario in terms of the constraints on the m_A and $\tan\beta$ values. Furthermore, the results are also expressed in a less model-dependent way in terms of the upper limits on the cross section for the pro-

duction of a generic Higgs boson ϕ with a mass m_ϕ via the production processes $pp \rightarrow b\bar{b}\phi$ and $gg \rightarrow \phi$.

4.1.4 Data and Simulated Event Samples

Data Sample

The presented result are based on proton-proton collision data recorded by the ATLAS experiment during 2012 at a centre-of-mass energy of $\sqrt{s} = 8$ TeV, corresponding to an integrated luminosity of 20.3 fb^{-1} . The events used in this analysis are recorded using a combination of a single electron and combined electron-muon triggers. Only recorded events in which all relevant components of the ATLAS detector were fully operational are considered. Additional data quality requirements are applied to each event according to [114]. These requirements assure the rejection of events with jet activity in known noisy calorimeter regions.

Signal Samples

Signal production via the gluon fusion process, $gg \rightarrow A/H/h$, was simulated with POWHEG [77] and the associated $b\bar{b}A/H/h$ production with SHERPA [78]. The pseudo-scalar Higgs boson samples were generated in the mass range from 90 GeV to 300 GeV assuming $\tan\beta = 20$, all three neutral Higgs bosons ($A/h/H$) are assumed to decay with the same kinematic properties. Appropriate re-weighting of the production cross sections is applied to simulate other $\tan\beta$ values. The m_h^{mod} MSSM benchmark scenario is assumed.

Background Samples

The production of W and Z/γ^* bosons in association with jets was simulated with the ALPGEN [70] generator. The $t\bar{t}$ process was generated using the POWHEG generator. The single top quark production via s-channel and in Wt process were generated using MC@NLO [72], while single top quark production via t-channel was generated with the AcerMC [73] generator. Diboson processes (WW , WZ , ZZ) were generated with the HERWIG [74] generator. For all ALPGEN

and MC@NLO event samples described above, the parton shower and hadronization were simulated with HERWIG and the underlying event activity with the JIMMY [75] programme. Different sets of parton density functions (PDFs) are used depending on the generator: CTEQ6L1 [79] is used by ALPGEN and AcerMC while CT10 [80] is used by SHERPA, POWHEG and MC@NLO.

TAUOLA [82] and PHOTOS [83] are used to model the tau lepton decay and additional photon radiation from charged leptons in the leading-log approximation, respectively.

The ATLAS detector response is simulated for all generated samples using the GEANT4 [84, 85] package, the reconstruction of physics objects, described in chapter 3, is performed with the same software used also for the data. The effects of the simultaneous recording of additional proton collisions from the same or neighbouring bunch crossings (pile-up) are taken into account in the simulation.

4.2 Event Selection and Categorization

4.2.1 The Common Selection Criteria

According to the characteristic properties of signal events, each event in data and simulation should satisfy the selection criteria described in the following. Since these are shared by both the b-tagged and b-vetoed event category, they are referred to as common selection criteria:

- (i) A trigger selection, requiring the presence of a single electron with $p_T > 24$ GeV, or alternatively, an electron with $p_T > 12$ GeV together with a muon with $p_T > 8$ GeV.
- (ii) At least one reconstructed vertex with more than three associated tracks. This selection is aimed to reject background from cosmic muons.
- (iii) Exactly one reconstructed “Tight” electron with $|\eta| < 1.37$ or $1.52 < |\eta| < 2.47$. The electron should have $p_T > 15$ or 25 GeV depending on the trigger that selected the event.

- (iv) Exactly one “Combined” muon with $|\eta| < 2.5$ and $p_T > 10$ GeV.
- (v) The electron should be isolated with $E_T^{cone}/p_T < 0.08$ and $P_T^{cone}/p_T < 0.06$.
- (vi) The muon should be isolated with $E_T^{cone}/p_T < 0.04$ and $P_T^{cone}/p_T < 0.06$.
- (vii) Muon and electron should be of opposite charge.
- (viii) Overlap removal between electron, muon, τ -jets and jets is performed.
- (ix) The event is rejected if at least one hadronic τ lepton decay is found with τ -jet transverse momentum $p_T > 15$ GeV. τ -jets candidate are required to be associated to one or three charged tracks, for the identification a “Medium” BDT working point is chosen, additionally, a BDT-based electron veto is applied.
- (x) To reduce QCD-multijet background contamination, the invariant mass obtained from the sum of the electron and muon four-momenta should be greater than 30 GeV.

Details on the definition of physics objects and the applied quality criteria can be found in chapter 3.

Events accepted by the common selection criteria are categorized into the *b-tagged* and *b-vetoed* event categories by requiring the presence of exactly one b-tagged jet or the absence of any b-tagged jet in the event, respectively. A jet is considered b-tagged if it has $p_T > 20$ GeV, $|\eta| < 2.5$, $JVF > 0.5$ and it passes the selection of the *MV1* b-tagging algorithm at 70% of efficiency for b-quark, $\epsilon_b^{t\bar{t}}$. Further selection criteria are applied to each category and optimized separately, as described in the following.

4.2.2 b-vetoed Event Category

A veto on the presence of b-tagged jets in the final state allows for the selection of signal events which are produced predominantly via gluon fusion. In this event category the $Z/\gamma^* \rightarrow \tau\tau$ process is an irreducible background due to the same

topology of the Higgs and Z boson decay. Other background processes can still be discriminated against the signal due to their kinematic properties. The τ leptons from the Higgs boson decay are highly boosted and so are their decay products, this results in significantly different lepton kinematics with respect to diboson or $t\bar{t}$ background processes. Firstly, the electron and muon from Higgs boson decay will be more likely emitted back-to-back. This is illustrated in Figure 4.4(a), showing the angular distance between the two leptons in the transverse plane $\Delta\phi_{e,\mu} = |\phi_e - \phi_\mu|$ for the signal and relevant background processes. Secondly, the neutrinos from the Higgs boson decay will be more likely collinear with the charged leptons, thus, the angular correlation between the direction of the missing transverse energy and the two leptons, derived as:

$$\hat{E}_T^{miss} \cdot (\hat{P}_T^\mu + \hat{P}_T^e) = \cos(\Delta\phi_{E_T,\mu}) + \cos(\Delta\phi_{E_T,e}) = \sum_\ell \cos(\Delta\phi_{E_T,\ell})$$

is expected to tend to zero, as is shown in Figure 4.4(b). These two features can be used to discriminate the signal from the W boson, top quark and dibosons background processes. No further selection criteria are applied in this event category, as it has been shown that no significant improvement of the analysis sensitivity can be achieved. The exact selection criteria are listed in Table 4.2, while in Table 4.3 the predicted number of background and signal events after each stage of selection are reported.

4.2.3 b-tagged Event Category

The request of exactly one b-tagged jet in the b-tagged event category selects predominantly signal events produced in b-quarks associated production mode. Background processes with b-jet activity, as the top quark and single top quark production become enhanced compared to the $Z/\gamma^* \rightarrow \tau\tau$ background. Also in this category selection requirement on $\Delta\phi_{e,\mu}$ and $\sum \cos \Delta\phi$ are imposed to reduce the top quark and diboson background contributions, as described for the b-vetoed event category. Further selection criteria specific for this category are employed as described below.

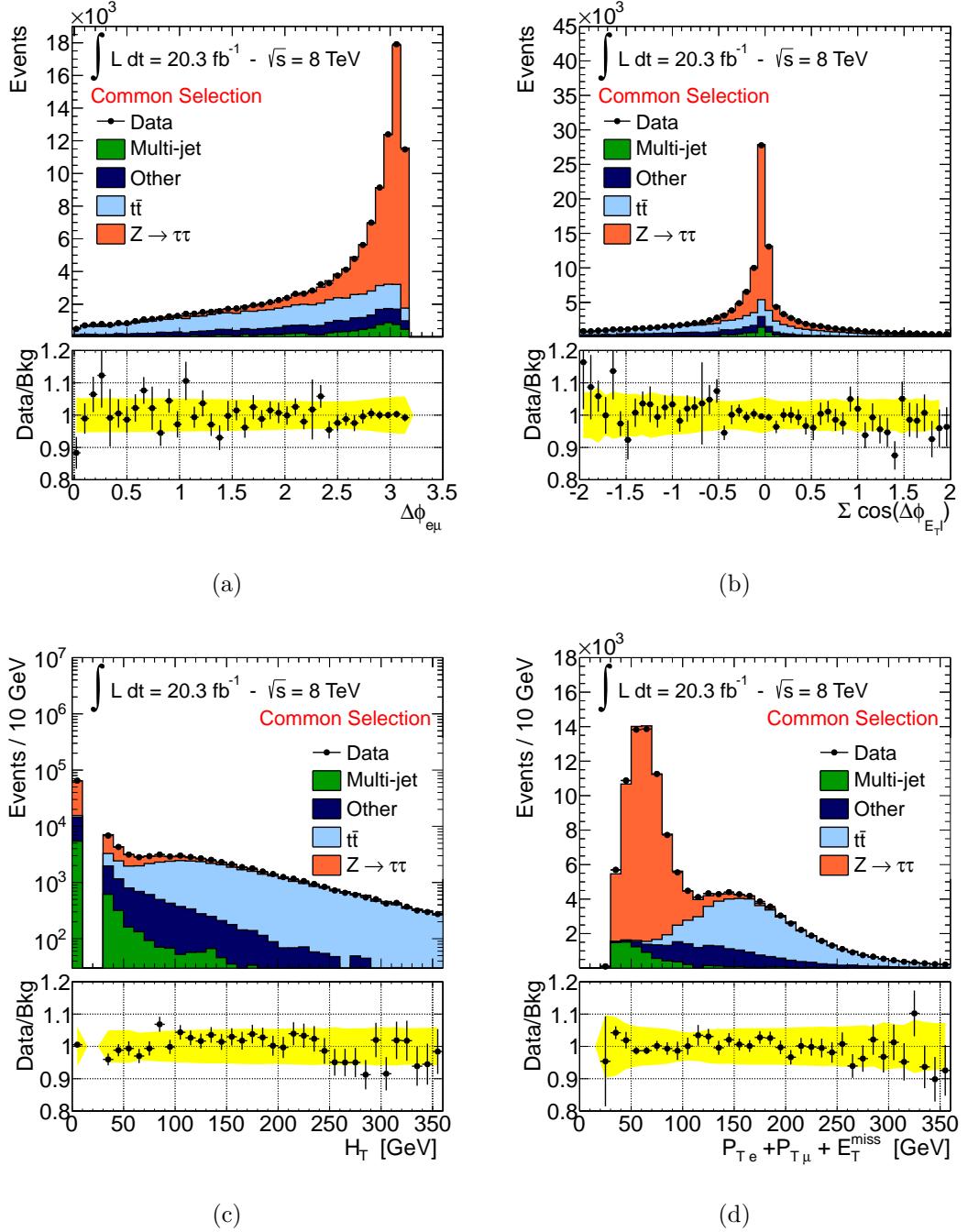


Figure 4.4: Distributions of relevant discriminating variables shown after the common selection has been applied. The prediction of the background model is compared to data. The contribution of the $Z/\gamma^* \rightarrow \tau\tau$ and QCD multi-jet background processes is measured in dedicated signal-depleted control data samples, the prediction for all the other background processes is obtained from simulation. The notation “Other” stands for the electroweak processes $W \rightarrow \ell\nu$, $Z \rightarrow \ell\ell$, diboson and single top quark production. The yellow band represents the total systematic uncertainty for the background model prediction (see Section 4.4).

Category	Selection
b-vetoed	No b-tagged jets $\Delta\phi_{e,\mu} > 1.6$ $\sum \cos \Delta\phi > -0.4$
b-tagged	Exactly one b-tagged jet $\Delta\phi_{e,\mu} > 2$ $\sum \cos \Delta\phi > -0.2$ $H_T < 100 \text{ GeV}$ $P_{T\mu} + P_{Te} + E_T^{miss} < 100 \text{ GeV}$

Table 4.2: Summary of the event selection criteria in the b-tagged and b-vetoed event categories, applied after the common event selection has been performed.

Signal events in this event category can be discriminated from the top quark process given their relatively low jet activity. The top quark process is very likely to have two or more highly energetic jets in the event, unlike the signal b-jet which are relatively low energetic. Weak jet activity is ensured by requesting the sum of the jets transverse momenta H_T in the event to be small. The H_T distribution is shown in Figure 4.4(c). The jets used for the calculation of the H_T value should have $p_T > 30 \text{ GeV}$, $|\eta| < 4.5$ and $\text{JVF} > 0.5$ (if $|\eta| < 2.5$).

Another feature that discriminate top quark pair production from the Higgs boson signal is the higher invariant mass of the former final state, as the highest Higgs mass considered for the presented search is 300 GeV. The sum of electron and muon transverse momenta with E_T^{miss} is used as a corresponding discriminating variable, whose distribution is shown in Figure 4.4(d) .

The summary of the exact optimized selection criteria for the b-tagged event category is shown in Table 4.2. In Table 4.4 the predicted number of background and signal events after each stage of selection in the b-tagged event category is reported.

	Common Selections	n(b-jet)=0	$\Delta\phi(e - \mu) > 1.6$	$\sum \cos \Delta\phi > -0.4$
Data	125886	89155	-	--
Multi-jet	6693 ± 456	6357 ± 461	5322 ± 438	4137 ± 339
$Z \rightarrow \ell\ell$	569 ± 48	564 ± 48	516 ± 47	434 ± 44
$W \rightarrow \ell\nu$	1625 ± 155	1604 ± 155	1145 ± 125	714 ± 101
Dibosons	9338 ± 48	9235 ± 48	7358 ± 43	4002 ± 31
$t\bar{t}$	40632 ± 106	7707 ± 46	5044 ± 37	3416 ± 31
Single Top	4449 ± 44	1664 ± 27	1124 ± 22	682 ± 18
$Z/\gamma^* \rightarrow \tau\tau$	61503 ± 68	60440 ± 67	58078 ± 65	55303 ± 64

Table 4.3: Number of observed and predicted signal and background events, after each selection stage in the b-vetoed event category.

4.2.4 Mass Reconstruction with MMC Technique

Acurate invariant mass reconstruction of a di- τ resonace is a challenging task due to the undetected neutrinos. In the presented analysis a total of four neutrinos are involved in the final state, two for each of the τ lepton decays. The invariant mass depends on eight unknowns given by the two sums of neutrino four-momenta, one for each τ lepton decay. These unknowns can be constrained by four parameters obtained from the measured missing transverse energy and from the τ lepton mass using the following equations:

$$\vec{E}_T^{miss} = \vec{P}_T^{mis_1} + \vec{P}_T^{mis_2} \quad (4.1)$$

$$M_{\tau_i}^2 = m_{mis_i}^2 + m_{vis_i}^2 + 2\mathbf{P}_{vis_i} \cdot \mathbf{P}_{mis_i}$$

where the index i runs over the two τ leptons in the event. $\vec{P}_T^{mis_i}$, m_{mis_i} and \mathbf{P}_{mis_i} are respectively the transverse momentum, the invariant mass and the four momentum of the pair of neutrinos originating from the decay of the i -th τ lepton with mass M_τ . The subscript vis indicates instead quantities related to the charged lepton from the corresponding τ lepton decay. The remaining four degrees of freedom can be further constrained, for example, assuming that the neutrinos are collinear to the electron or muon from the corresponding τ lepton decay. This approximation, however, introduces limitations on the mass resolution.

In this analysis, the so-called "Missing Mass Calculator" (MMC) algorithm is

	n(b-jet)=1	$\Delta\phi$	$\sum \cos \Delta\phi$	$P_{T\mu} + P_{Te} + E_T^{miss}$	H_T
Data	23352	-	-	-	-
Multi-jet	330 ± 40	208 ± 27	135 ± 22	114 ± 17	100 ± 15
$Z \rightarrow \ell\ell$		5.2 ± 1.8	2.3 ± 1.1	2.3 ± 1.1	1.7 ± 1.0
$W \rightarrow \ell\nu$		20 ± 6	15 ± 6	13 ± 6	10 ± 6
Dibosons		99 ± 5	63 ± 4	36.4 ± 3.0	14.8 ± 1.8
$t\bar{t}$		19810 ± 70	9680 ± 50	6450 ± 50	808 ± 15
Single Top		2456 ± 33	1223 ± 23	784 ± 18	122 ± 7
$Z/\gamma^* \rightarrow \tau\tau$		952 ± 9	625 ± 7	540 ± 7	482 ± 6

Table 4.4: Number of observed and predicted signal and background events, after each selection stage in the b-tagged event category.

used to calculate the most likely invariant mass of the di- τ system for a given event topology. The implementation of the MMC method in this search is based on [130]. The MMC algorithm solves the equations 4.1 for a set of points in a grid of a four-dimensional parameter space. The four independent variables are chosen to be $m_{mis_i}^2$ and $\cos\theta_i^*$, the latter defined as the angle between the charged lepton from the τ lepton decay and the boost direction of the τ lepton. The di- τ invariant mass in each event is then calculated for each given point of the parameter space. Each solution is weighted by the probability that a τ lepton decay assumes a given configuration. The probability of a given configuration is predicted by means of simulation using PYTHIA generator supplemented with TAUOLA package. The invariant mass of the di- τ system, $m_{\tau\tau}^{MMC}$, is then defined as the maximum of the weighted invariant mass distribution obtained from all scanned points.

The resolution of the missing transverse energy measurement impacts the resolution of the invariant mass obtained with the $m_{\tau\tau}^{MMC}$ method. To improve the E_T^{miss} resolution, a scan over a six-dimensional parameter space is performed in a similar manner as described above. For this purpose, the value of \vec{E}_T^{miss} is also considered unknown and a scan is performed over all possible values constrained by the measured E_T^{miss} and its corresponding uncertainty.

Figure 4.5 shows the $m_{\tau\tau}^{MMC}$ invariant mass distribution after the common selection and after the requirement of the presence or the absence of a b-tagged

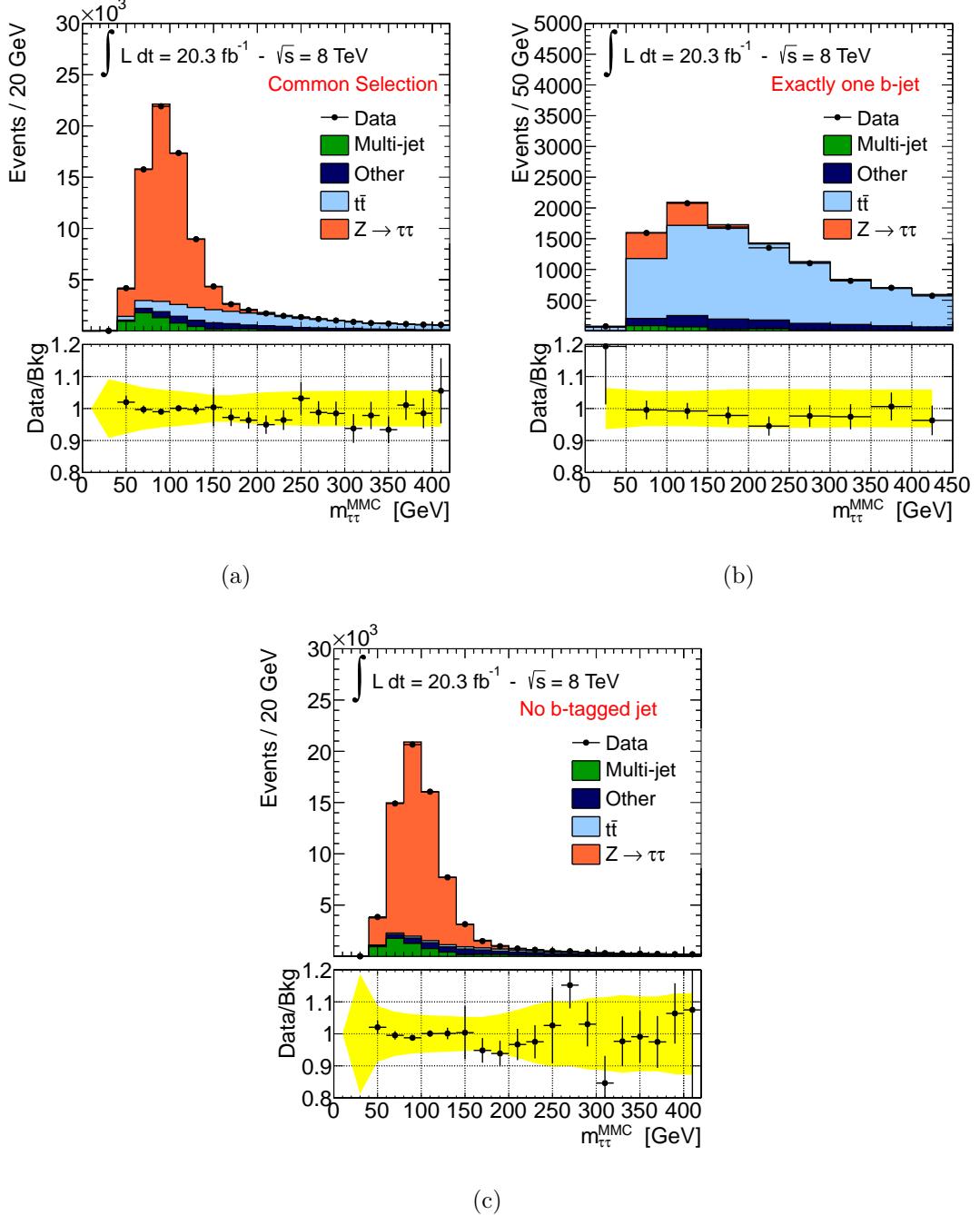


Figure 4.5: Observed and expected distribution of the invariant di- τ mass $m_{\tau\tau}^{MMC}$ for different stages of analysis selection: after requiring the common selection (a), additionally requiring the presence of exactly one b-tagged jet (b) or the absence of a b-tagged jet (c). The prediction of the background model is compared to data. The contribution of the $Z/\gamma^* \rightarrow \tau\tau$ and QCD multi-jet background processes is measured in dedicated signal-depleted control data samples, the prediction for all the other background processes is obtained from simulation. The notation “Other” stands for the electroweak processes $W \rightarrow \ell\nu$, $Z \rightarrow \ell\ell$, diboson and single top quark production. The yellow band represents the total systematic uncertainty for the background model prediction.

jet.

4.3 Background Prediction and Validation

This section describes the strategies for the prediction of the background contributions and validation of these predictions. Monte Carlo simulation is extensively used to model the kinematic properties of the background and signal processes. However, since the simulation of any process is usually prone to systematic uncertainties due to a non-perfect description of the pileup effects, underlying event and detector performance, the background contributions from $Z/\gamma^* \rightarrow \tau\tau$ and QCD multi-jet process are estimated using dedicated signal-free control data samples, as described respectively in section 4.3.3 and 4.3.2. Contributions of other background processes, such as $t\bar{t}$, single top quark, dibosons, $Z \rightarrow ll + \text{jets}$ (where $l = e, \mu$) and $W + \text{jets}$, are estimated from simulation. Given the relatively large $t\bar{t}$ background contribution, a dedicated study to validate this background prediction has been made as described in section 4.3.1.

A good agreement between data and background prediction is found after the common selection, as can be seen in Figure 4.4 and Figure 4.6.

4.3.1 Validation of the $t\bar{t}$ Background Prediction

The background contribution from top quark pair production is estimated using a sample of simulated events generated with POWHEG-PYTHIA Monte Carlo generator. Since this is one of the major background processes for the presented analysis a careful validation of the predicted contribution is needed. For this purpose signal-depleted data validation sample enriched with $t\bar{t}$ events by requiring the presence of exactly two b-tagged jets in all the events passing the common selection is employed. Figures 4.7 and 4.8 show the distributions for a set of kinematic properties and all discriminating variables obtained with this data sample. Good agreement between data and Monte Carlo prediction is found with overall ratio of the observed to the predicted number of $t\bar{t}$ events of $0.998 \pm 0.011(\text{stat.}) \pm 0.110(\text{sys.})$. The total systematic uncertainty on the ratio is dominated by the uncertainty of

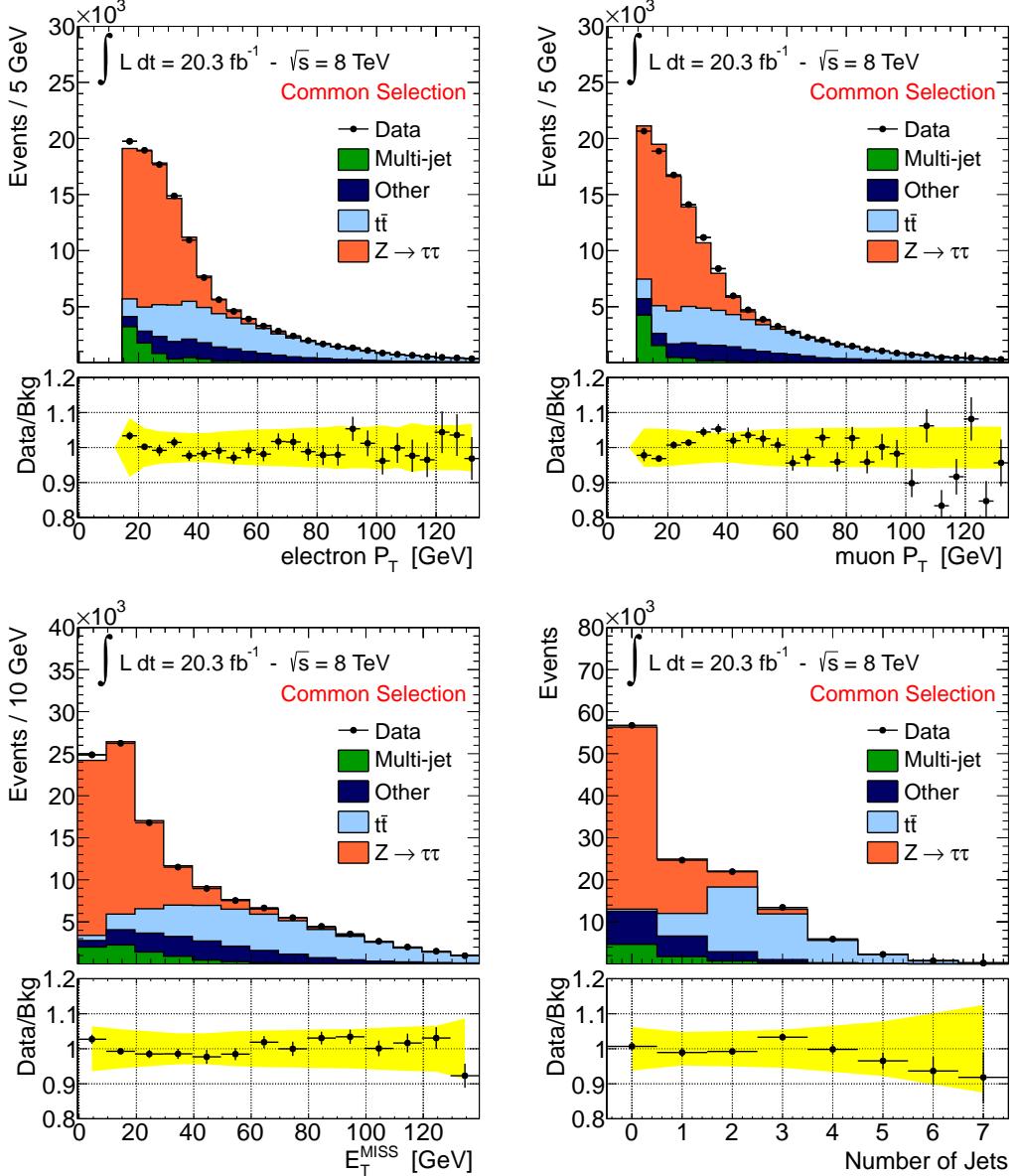


Figure 4.6: Observed and expected distribution of kinematic variables after common selections. The prediction of the background model is compared to data. The contribution of the $Z/\gamma^* \rightarrow \tau\tau$ and QCD multi-jet background processes is measured in dedicated signal-depleted control data samples, the prediction for all the other background processes is obtained from simulation. The notation “Other” stands for the electroweak processes $W \rightarrow \ell\nu$, $Z \rightarrow \ell\ell$, diboson and single top quark production. The yellow band represents the total systematic uncertainty for the background model prediction.

Data Sample	Relative Lepton Charge	Lepton Isolation
A (signal sample)	OS	isolated
B	SS	isolated
C	OS	anti-isolated
D	SS	anti-isolated

Table 4.5: Control data samples for the measurement of the QCD multi-jet background contribution. The samples are defined by the requirements on the relative charge sign of the two leptons (OS,SS) and the isolation criteria applied on them (isolated or anti-isolated). See text.

the b-tagging efficiency.

4.3.2 Multi-jet Background Measurement

The QCD multi-jet process represents an important background, especially in the b-vetoed event category, due to its high cross-section and the relatively low threshold on the lepton p_T used in this analysis. The contribution of this background is evaluated by the so-called ABCD data-driven technique. The ABCD method splits the sample of data events after the common selection into four sub-samples: the signal sample (A), defined by the event selections criteria described in Section 4.2 and three signal-depleted control data sample (B,C,D), which are orthogonal to each other and are enriched in multi-jets events. The three control data samples are defined by inverting the requirements on the relative sign of the electron and muon charge and on the isolation criteria. Both the calorimetric and tracking isolation criteria described in Section 4.2.1 are inverted for each electron and muon with respect to the nominal values, thus defining the so-called anti-isolated leptons. The data are divided into four samples of events with leptons of opposite sign charge (OS) or same sign charge (SS) with respectively isolated or anti-isolated leptons, as summarized in Table 4.5.

The ABCD method assumes that there is no correlation between the relative charge and lepton isolation in QCD multi-jet events, or in other words that the

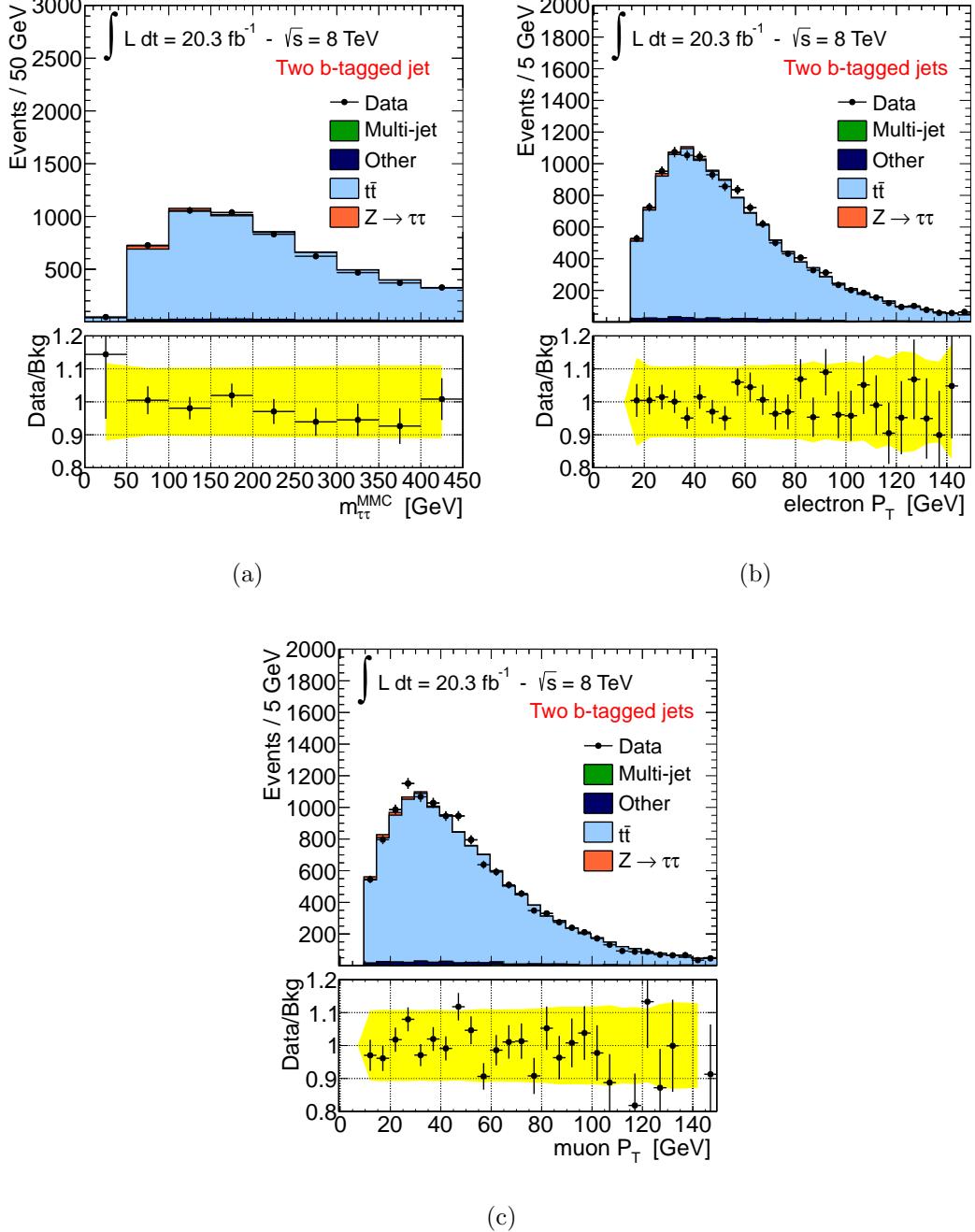


Figure 4.7: Observed and expected distributions of (a) the di- τ invariant mass $m_{\tau\tau}^{MMC}$, (b) the electron transverse momentum and (c) the muon transverse momentum in the $t\bar{t}$ validation sample. The contribution of the $Z/\gamma^* \rightarrow \tau\tau$ and QCD multi-jet background processes is measured in dedicated signal-depleted control data samples, the prediction for all the other background processes is obtained from simulation. The notation ‘‘Other’’ stands for the electroweak processes $W \rightarrow \ell\nu$, $Z \rightarrow \ell\ell$, diboson and single top quark production. The error bars on the observed to the predicted events ratio indicates the statistical uncertainty, whereas the yellow band indicates the total systematic uncertainty of this ratio.

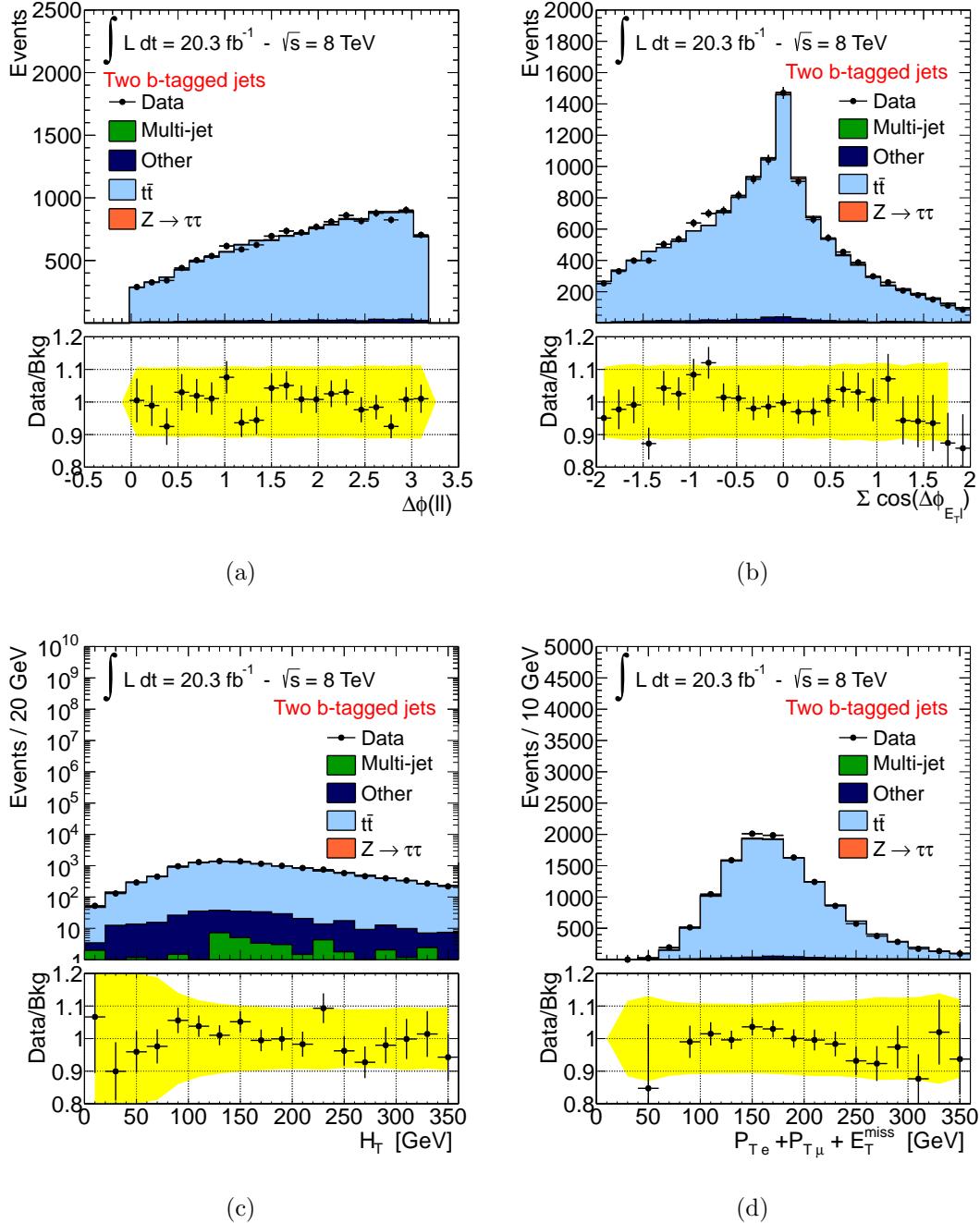


Figure 4.8: Observed and expected distributions of the discriminating variables, (a) $\Delta\phi(e - \mu)$, (b) $\sum \cos \Delta\phi$, (c) $p_{T\mu} + p_{Te} + E_T^{\text{miss}}$ and (d) H_T in the $t\bar{t}$ validation sample. The background prediction is performed as previously stated. The error bars on the observed to the predicted events ratio indicates the statistical uncertainty, whereas the yellow band indicates the total systematic uncertainty of this ratio.

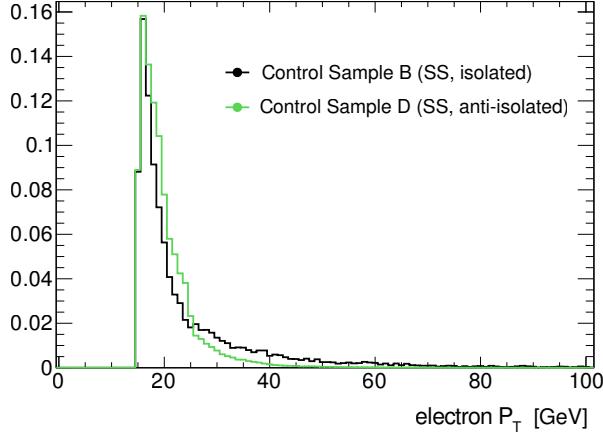


Figure 4.9: Comparison of the electron p_T distribution in control samples B and sample D, showing the bias due to the trigger. The histograms are normalized to the same area.

ratio of OS/SS events is uncorrelated with the lepton isolation criteria. In this case, the number (N_A) of QCD multi-jet events in the signal sample A can be estimated from the yields (N_B , N_C , N_D) of multi-jet events in the control samples B , C and D , using the equation

$$N_A = N_B \times \frac{N_C}{N_D} = N_B \times R_{QCD} \quad (4.2)$$

To obtain the pure QCD multi-jet event yields in the data control samples, the contribution from contaminating electroweak (W+jets, Z+jets and dibosons) and top quark processes ($t\bar{t}$, single top quark production) is subtracted in each control sample based on the prediction from simulation. Tables 4.7 and 4.6 show the observed event yield in each control sample at different stages of the event selection along with the predictions of non-QCD background contributions which are subtracted. Signal contamination has been evaluated in all three control samples for different signal mass points. For the range of m_A and $\tan\beta$ values considered in this analysis, the highest signal contamination is seen in sample B for the mass point $m_A = 300$ GeV and $\tan\beta = 50$, where a contamination of 0.2% is observed².

² This contamination signal originates mainly from the production in association with b-quarks

The modelling of the shapes of kinematic distributions in QCD multi-jet events is given by the data sample B. The events in this sample are expected to have similar kinematic properties as in the signal sample. A drawback of this choice is a rather low number of events and a higher contamination with non-QCD process compared to samples C and D. Sample B is chosen to avoid a shape bias due to isolation requirements at the trigger level, since the single-electron trigger already imposes isolation requirements. Figure 4.9 shows the comparison of the electron p_T distributions in sample B and D. In the latter sample high- p_T electrons are suppressed as they do not pass the trigger selection. Eventually the trigger isolation requirement could also bias the ratio R_{QCD} . This possibility has been carefully studied in a dedicated study as reported in Appendix A. To a good approximation, the mentioned trigger effects cancel out in the ratio R_{QCD} and no additional systematic uncertainty needs to be taken into account.

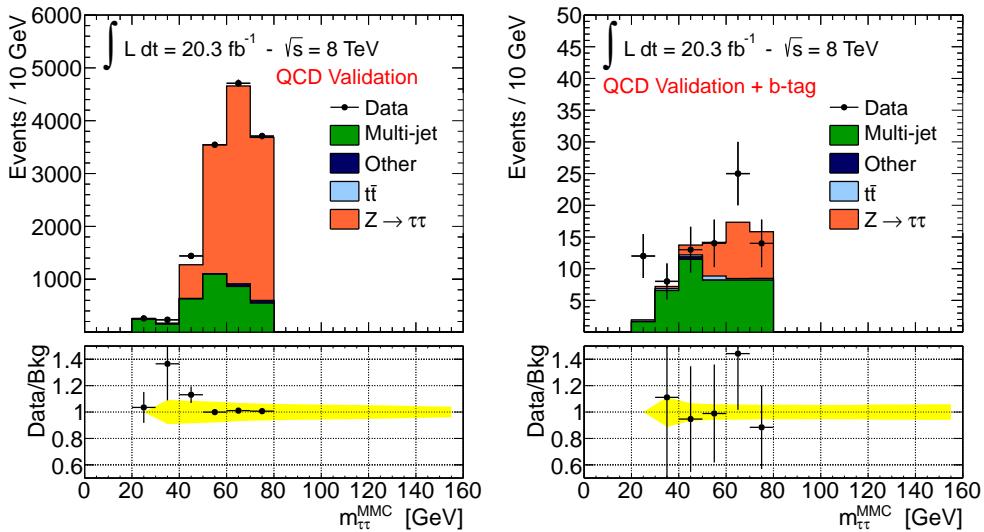


Figure 4.10: $m_{\tau\tau}^{MMC}$ distribution obtained with QCD validation samples defined in Section 4.3.2, without (left) and with an additional requirement of exactly one b-tagged jet in the final state (right). The error bars and the yellow band indicates the statistical and systematic uncertainty respectively.

and, as it scales with the cross section, it will be an order of magnitude smaller for $\tan\beta = 20$.

Event Selection		B	C	D	R_{QCD}
Common Selection	Data	6189	604628	312901	1.929 ± 0.004
	non-QCD	2510 ± 180	1090 ± 30	730 ± 35	
B-veto	Data	5673	558217	284847	1.960 ± 0.004
	non-QCD	2220 ± 180	710 ± 30	415 ± 30	
$\Delta\phi_{e,\mu}$	Data	4610	532583	271404	1.962 ± 0.005
	non-QCD	1700 ± 170	580 ± 30	345 ± 30	
$\sum \cos \Delta\phi$	Data	3417	486747	247712	1.965 ± 0.005
	non-QCD	1120 ± 100	370 ± 20	230 ± 20	
$m_{\tau\tau}^{MMC} > 0.$	Data	3177	479967	244276	1.965 ± 0.005
	non-QCD	1000 ± 100	300 ± 17	190 ± 20	

Table 4.6: Number of observed events and the predicted non-QCD contribution at different stages of the event selections for b-veto category. The error on the R_{QCD} ratio is of statistical nature only.

To test the predictions of the ABCD method an additional validation sample has been defined with the following selection criteria after applying the common selection:

- $E_T^{miss} < 20$ GeV
- $H_T < 70$ GeV and $p_{T\mu} + p_{Te} + E_T^{miss} < 50$ GeV
- $0 < m_{\tau\tau}^{MMC} < 80$ GeV

This validation sample is designed to enhance the multi-jet background contribution with respect to $Z/\gamma^* \rightarrow \tau\tau$ keeping the final state kinematics as similar as possible to the signal sample. Figure 4.10 shows the $m_{\tau\tau}^{MMC}$ distribution for this validation sample with and without the b-tagging requirements. Agreement between data and the background predictions is found within statistical and detector-related systematics uncertainty.

Systematic uncertainties are assigned on the scaling factor R_{QCD} and on the shape of the discriminating variable $m_{\tau\tau}^{MMC}$ to take into account any correlation

Event Selection		B	C	D	R_{QCD}
Common Selection	Data	6189	604628	312901	1.929 ± 0.004
	non-QCD	2510 ± 180	1090 ± 30	730 ± 35	
B-tag	Data	419	44619	27257	1.64 ± 0.01
	non-QCD	215 ± 10	310 ± 12	277 ± 13	
$\Delta\phi_{e,\mu}$	Data	230	38810	23316	1.67 ± 0.01
	non-QCD	104 ± 6	200 ± 10	175 ± 7	
$\sum \cos \Delta\phi$	Data	149	31379	18779	1.67 ± 0.02
	non-QCD	67 ± 5	127 ± 8	114 ± 6	
$\sum H_T$	Data	83	27781	15626	1.78 ± 0.02
	non-QCD	23 ± 4	25 ± 3	22 ± 3	
$p_{T\mu} + p_{Te} + E_T^{miss}$	Data	71	27735	15590	1.78 ± 0.02
	non-QCD	10 ± 3	22 ± 3	18 ± 2	
$m_{\tau\tau}^{MMC} > 0.$	Data	70	27634	15522	1.78 ± 0.02
	non-QCD	9 ± 3	20 ± 3	17 ± 2	

Table 4.7: Number of observed events and the predicted non-QCD contribution at different stages of the event selections for b-tag category. The error on the R_{QCD} ratio is of statistical nature only.

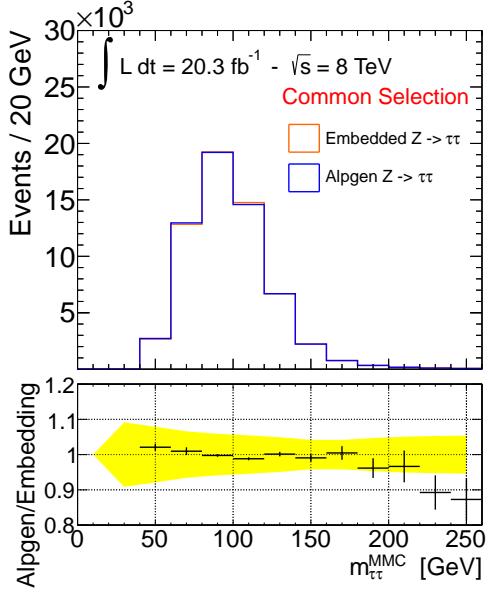


Figure 4.11: Comparison of the $m_{\tau\tau}^{MMC}$ distributions obtained from the ALPGEN $Z/\gamma^* \rightarrow \tau\tau$ simulation and from the embedding technique after the requirements of the common selection has been applied. The yellow band indicates the total systematic uncertainty relative to the ALPGEN simulation sample.

between the isolation and the relative charge of the leptons as detailed in Section 4.4.

4.3.3 $Z \rightarrow \tau\tau + \text{Jets}$ Background Measurement

The $Z/\gamma^* \rightarrow \tau\tau$ decays are the major source of background to the presented analysis, calling for its thorough understanding. Unfortunately, for a light Higgs boson, it is impossible to fully discriminate the $Z/\gamma^* \rightarrow \tau\tau$ decays from the signal and thus a dedicated signal-free data control sample cannot be defined. However, thanks to the small Higgs boson coupling to muons, $Z \rightarrow \mu\mu$ decays provide a good starting point to model $Z/\gamma^* \rightarrow \tau\tau$ events based on data. A hybrid approach relying on data and simulation known as the "embedding" is used for this purpose. The $Z \rightarrow \mu\mu$ event candidates are selected in data. The two muons

from the Z decay are then substituted with the decay products from simulated τ lepton decays. The energy deposit in the calorimeter and the reconstructed tracks in a cone of given size around the muon are subtracted and substituted with the corresponding predictions from the τ lepton decay. These τ leptons have the same kinematic properties as the original muons. Further details on the embedding technique may be found in [86, 87].

As the Trigger is not simulated in the described embedded samples, only the shapes of kinematic distributions are modelled by the embedded sample, while the $Z \rightarrow \tau\tau$ event yield is normalized to ALPGEN $Z/\gamma^* \rightarrow \tau\tau$ prediction at the common selection stage. Furthermore, a set of corrections as described in [88], is applied to unfold from the original $Z \rightarrow \mu\mu$ trigger and reconstruction efficiency. Subsequently, the trigger and reconstruction efficiency of the $e\mu + 4\nu$ final state are emulated by means of event weights.

The embedding technique has been validated in several studies detailed in [86, 88], demonstrating a reliable performance of the embedding technique and a good description of data. Figure 4.11 shows a comparison of the $m_{\tau\tau}^{MMC}$ distribution between embedded and ALPGEN $Z/\gamma^* \rightarrow \tau\tau$ events, no significant differences are seen. Other relevant discriminating variables, such as the E_T^{miss} and the number of b-jets in the final state, are instead slightly better described by the embedded $Z/\gamma^* \rightarrow \tau\tau$ sample, as shown in Figure 4.12, this behaviour is expected due to the imperfect modelling of these variables with simulation.

The embedded sample is based on the selected $Z \rightarrow \mu\mu$ event candidates in data. The $Z \rightarrow \mu\mu$ selection criteria assure a rather pure $Z \rightarrow \mu\mu$ sample. However, further event selection criteria used in the presented analysis, for example the b-tagging requirements, could enhance the contamination of this sample with events from other processes. Dedicated studies have been made to estimate the $t\bar{t}$ and QCD multi-jet contamination in the embedded sample. The $t\bar{t}$ contamination is estimated by evaluating the yield of embedded $Z \rightarrow \tau\tau$ events in a validation sample with two b-tagged jets (as described in Section 4.3.1). These events are assumed to originate solely from the $t\bar{t}$ process and the corresponding yield in the signal sample is extrapolated from simulation. Table 4.8 summarises the evaluated

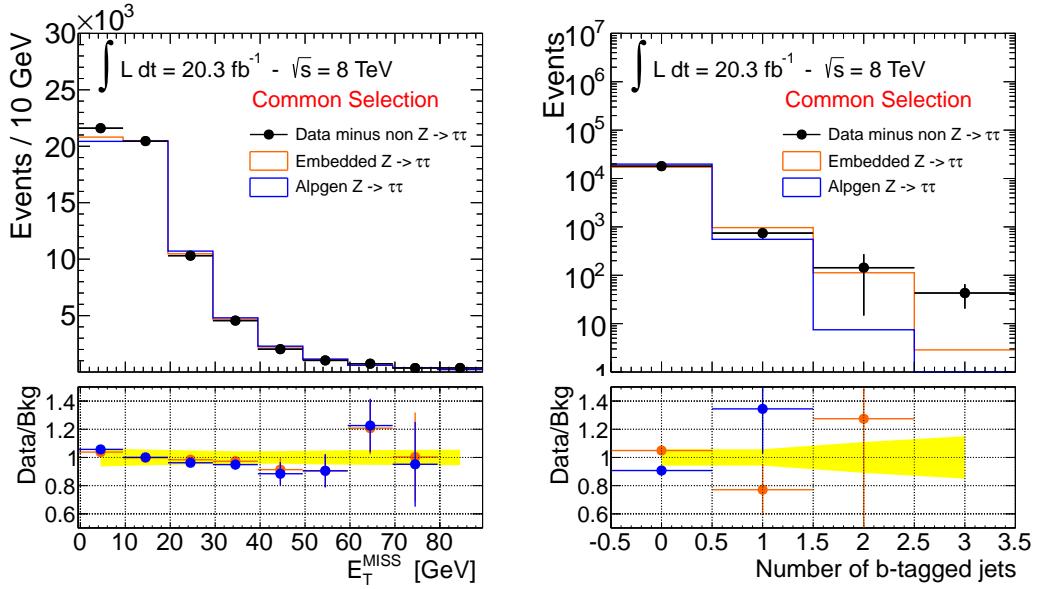


Figure 4.12: Comparison of the (left) E_T^{miss} and (right) b-tagged jet multiplicity distributions in embedded and ALPGEN $Z/\gamma^* \rightarrow \tau\tau$ events after the requirements of the common selection has been applied. Data are superimposed after subtracting the contribution of non- $Z/\gamma^* \rightarrow \tau\tau$ processes. The yellow band indicates the total systematic uncertainty relative to the ALPGEN simulation sample.

top quark contamination in the embedded $Z \rightarrow \tau\tau$ sample, separately for the two event categories. The multi-jet contamination can be estimated starting from the yield of embedded events in the sample C defined by the ABCD method. It is assumed that all events in this validation sample are QCD multi-jet events. The QCD multi-jet contamination of the embedded events in the signal sample can be estimated as:

$$N_A^{QCD-emb} = N_C^{QCD-emb} \times \frac{N_B^{\mu\mu}}{N_D^{\mu\mu}} = N_B \times R_{QCD}^{\mu\mu} \quad (4.3)$$

The transfer factor $R_{QCD}^{\mu\mu}$, is evaluated using a di-muon final state with same kinematic selection criteria as for $Z \rightarrow \mu\mu$ candidates entering the embedding procedure. Table 4.9 shows the estimated contamination of QCD multi-jet in the embedded sample. Contamination effects are considered negligible.

	Embedded event yield in $t\bar{t}$ control sample	Transfer factor	Estimated events in signal sample	Contamination
b-taged	84 ± 9	$(2.6 \pm 0.1) \times 10^{-2}$	2.2 ± 0.2	0.5 %
b-vetoed	84 ± 9	$(1.74 \pm 0.02) \times 10^{-1}$	15 ± 2	0.03 %

Table 4.8: Evaluation of the $t\bar{t}$ contamination in the embedded $Z \rightarrow \tau\tau$ sample using a two b-tag validation sample. The transfer factor is a multiplicative factor obtained from simulation that allows to extrapolate the measurement from the validation sample into the signal sample.

4.4 Systematic Uncertainties

This section describes a number of sources of systematic uncertainties that are relevant for the presented analysis. To account for differences in the observed and simulated detector response a set of corrections is applied at the level of object reconstruction and at event level as described in chapter 3. The uncertainties due to such corrections are referred to as detector-related systematic uncertainties and are addressed in section 4.4.1. For all processes whose contributions are predicted

	Embedded event yield in QCD control sample C	Transfer factor	Estimated events in signal sample	Contamination
B-tag	12 ± 3	$(7 \pm 1) \times 10^{-3}$	$(8.4 \pm 0.3) \times 10^{-2}$	0.03 %
B-veto	390 ± 20	$(2.5 \pm 0.1) \times 10^{-2}$	10.0 ± 0.5	0.02 %

Table 4.9: Evaluation of the QCD multi-jet contamination in the embedded $Z \rightarrow \tau\tau$ sample using the control sample with OS anti-isolated events (sample C). The transfer factor $R_{QCD}^{\mu\mu}$ is the multiplicative factor that allows to extrapolate the measured events in control sample C into the signal sample.

from simulation, also theory-related systematic uncertainties need to be accounted for. These include uncertainties on the cross-section and on the acceptance of events after the given analysis selection criteria and are described in section 4.4.2 . Further systematic uncertainties related to the background measurements with dedicated control data are described in Sections 4.4.3 and 4.4.4 .

Each source of systematic uncertainty can contribute separately to the uncertainty on the final event yield and on the shape of the $m_{\tau\tau}^{MMC}$ distribution which is used as final discriminating variable in the statistical interpretation of data. Systematic uncertainties that affect the shape of the mass distribution are documented in appendix C.3. Uncertainty on the $m_{\tau\tau}^{MMC}$ shape distribution are found to be negligible for all except the embedded sample, for which significant deviation from the nominal distribution are found in the b-vetoed category only. Systematic uncertainties that do not affect the shape of the mass distribution and have an impact on the event yield of less than 0.5% for each process are neglected.

4.4.1 Detector-related Systematics Uncertainties

Systematic uncertainties related to object reconstruction and event-by-event corrections are based on the calibration measurements of the relevant parameter. Each of those parameters correspond to a nuisance parameter in the probability model used for the statistical interpretation as described in Section 4.5. Each parameter is varied independently by one standard deviation according to its measured uncer-

tainty. The corresponding impact on the yield of simulated events is evaluated for each signal and background sample. In the following, detector-related uncertainty are described in more details. Tables 4.11 and 4.10 briefly summarize the impact of these uncertainties on the predicted sample yields.

Luminosity The integrated luminosity of the 8 TeV data recorded with the ATLAS detector during 2012 is measured to be 20.3 fb^{-1} [63] with an uncertainty of 2.8%.

Pileup Simulated events are re-weighted to reproduce the average number of interactions per bunch crossing $\langle \mu \rangle$ as seen in data. Those event weights have an uncertainty which is propagated to each simulated sample.

Trigger Efficiency is corrected in simulation to match (on the average) the one observed in data. Those correction weights are evaluated as a function of p_T and η of the corresponding leptons and have associated uncertainties. Systematic uncertainties on both the single electron and electron-muon trigger efficiency are taken into account independently and range approximately from 1-2%.

In the embedded $Z \rightarrow \tau\tau$ sample, the trigger is emulated by applying weights according to the event topology. Those weights are related to the ones described above and have similar uncertainties. Trigger efficiency uncertainty for the embedded sample are considered uncorrelated with those of other samples.

Electrons Two sources of uncertainty on reconstructed electron objects are considered: the first related to electron identification and reconstruction efficiencies ("Electron ID") which range approximately from 1-2% depending on the transverse momentum of the electron, the second related to electron energy scale and resolution corrections, both ranging from 0.3-3% depending on η . The energy scale uncertainties are described by a set of six different nuisance parameters [102]. However, only a few of them give a non-negligible contribution to the analysis. Two of them are found to affect the shape of the $m_{\tau\tau}^{MMC}$ distribution and are considered independently: uncertainty arising from the electron momentum measurement with

$Z \rightarrow ee$ data ("Electron Zee") and the one related to low momentum electrons ("Electron LOWPT"). All other uncertainties related to energy scale and resolution are summed in quadrature ("Electron E").

Muons The uncertainty on muon identification efficiency depends on the charge and momentum of the muon. Typically these uncertainties are of the order of a fraction of percent, and are referred as "Muon ID". The uncertainties on the muon energy scale and resolution are considered independently for the inner detector and muon spectrometer measurements and are then added in quadrature to estimate the final uncertainty ("Muon E").

Taus-Jets The jets from the hadronically decaying τ leptons are only considered in the analysis by applying a veto on their presence in an event. Uncertainties on both τ -jet energy scale and identification efficiency have been investigated and are found to be negligible for this analysis.

Jets The systematic uncertainties on the Jet Energy Scale (JES) are described by multiple sets of nuisance parameters [113] related to different effects and jet energy components, for example the sensitivity to pileup effects or to the flavour composition of the jet. The overall uncertainty on the JES ranges between 3% and 7%, depending on the p_T and η of the jet. The overall impact of the JES uncertainty on the analysis yields is shown in Tables 4.11 and 4.10 by summing all component in quadrature, while in the statistical interpretation of data those uncertainties are considered uncorrelated. Systematic uncertainty due to the jet resolution ("Jet Resolution") are obtained by smearing the jet energy according to the measured uncertainty which ranges from 10-20% depending on the direction of the jet.

b-Tagging Corrections are applied to simulation to match the b-tagging efficiency observed in data. Uncertainties on the knowledge of the b-tagging efficiencies for the 70% $\epsilon_b^{t\bar{t}}$ working point of the MV1 b-tagger are considered [117, 118]. These uncertainties range from 5-10% depending on the p_T of the jet. The effect of

those uncertainties is evaluated independently for the b-quark, c-quark and light or gluon initiated jets and referred to respectively as "B Eff", "C Eff" and "L Eff". The tagging and mis-tagging efficiency uncertainties are considered to be fully anti-correlated.

Missing Transverse Energy The effect of the energy scale uncertainties for all physics objects is propagated to the E_T^{miss} calculation. In addition, uncertainty on the energy scale and resolution due to the remaining unassociated calorimeter energy deposits, the "soft-terms", is considered and estimated to be of the order of 10% [119]. E_T^{miss} uncertainties are independently propagated through the analysis and are added in quadrature, this final term is referred to as the "MET" uncertainty.

4.4.2 Theoretical Uncertainties

Uncertainties on the cross-sections that have been used to normalize the contribution of simulated samples to the integrated luminosity of analyzed data are reported in Table 4.12. These uncertainties include contributions due to parton distribution functions (PDFs), the choice of the value of the strong coupling constant, the renormalisation and factorisation scales. Furthermore, the uncertainties on the signal cross-section depend on the $\tan\beta$ value, the nature of the Higgs boson ($A/h/H$) and its mass.

The impact of systematic uncertainties due to various Monte Carlo tuning parameters for the description of the underlying event and lepton kinematic properties is considered. Since the distribution of the invariant mass of all visible τ lepton decay products is found to be not affected by these systematic, uncertainties only the variation in the acceptance is considered as a systematic uncertainty. The acceptance uncertainties for the simulated ALPGEN $Z/\gamma^* \rightarrow \tau\tau$ sample, which is used for the normalization of the embedded sample, are estimated at the common selection stage to be 4% [137]. Since additional selection criteria are applied directly to the embedded sample, no further acceptance uncertainties are considered. Acceptance uncertainties on the yield of $t\bar{t}$ simulated events are estimated

Source	b-vetoed event category, Uncertainties on event yields (%)				
	Signal bbH	Signal ggH	$Z/\gamma^* \rightarrow \tau\tau$	Top	Other
Electron ID	2.4	2.3	2.9 (s)	1.4	1.6
Electron E.	0.4	0.5	0.4	0.5	0.9
Electron LOWPT	0.3	0.5	0.4 (s)	0.0	1.2
Electron Zee	0.4	0.4	0.4 (s)	0.1	0.3
Muon ID	0.3	0.3	0.3	0.3	0.3
Muon E.	0.1	0.1	0.1	0.5	0.5
Trigger Single Ele.	0.6	0.6	0.5	0.9	0.9
Trigger Dilep.	1.0	1.0	1.3	0.2	0.3
Embedding MFS	-	-	0.1 (s)	-	-
Embedding Iso.	-	-	0.0 (s)	-	-
JES	0.6	0.7	-	1.0	1.2
JER	0.5	0.3	-	0.6	0.3
B Eff	1.8	0.0	-	12.0	0.8
C Eff	0.0	0.1	-	0.1	0.0
L Eff	0.0	0.1	-	0.2	0.1
Pileup	0.5	0.8	0.4	0.3	0.3
MET	0.2	0.8	0.1	0.2	0.5
Luminosity	2.8	2.8	2.8	2.8	2.8

Table 4.10: Impact of the experimental systematic uncertainties on the event yields in different simulated samples in the b-vetoed event category. Here "Other" refers to the sum of all remaining background samples: $W \rightarrow \ell\nu$, dibosons, $Z \rightarrow \ell\ell$ and single top quark processes. The signal produced in association with b-quarks and via gluon fusion is considered separately assuming $m_A = 150$ GeV and $\tan \beta = 20$. Uncertainty that impacts the $m_{\tau\tau}^{MMC}$ mass distribution are noted with the symbol (s).

b-tagged event category, Uncertainties on event yields (%)					
Source	Signal bbH	Signal ggH	$Z/\gamma^* \rightarrow \tau\tau$	Top	Other
Electron ID	2.3	2.6	2.8	1.8	2.0
Electron E	0.7	1.2	0.5	0.5	0.9
Electron LOWPT	0.4	0.0	0.4	0.1	0.4
Electron Zee	0.3	0.6	0.4	0.6	0.5
Muon ID	0.3	0.3	0.3	0.3	0.3
Muon E	0.5	0.8	0.1	0.1	0.2
Trigger Single Ele.	0.7	0.5	0.5	0.8	0.8
Trigger Dilepton	1.0	1.2	1.4	0.6	0.6
Embedding MFS	-	-	0.0	-	-
Embedding Iso.	-	-	1.3	-	-
JES	2.7	7.3	-	10.0	7.0
JER	1.4	6.3	-	2.9	3.0
B Eff	10.2	3.1	-	2.6	5.0
C Eff	0.2	4.3	-	0.0	1.2
L Eff	0.4	8.0	-	0.1	1.2
Pileup	0.4	0.7	0.4	0.4	0.9
MET	0.7	0.5	0.2	1.0	1.2
Luminosity	2.8	2.8	2.8	2.8	2.8

Table 4.11: Impact of the experimental systematic uncertainties on the event yields in different simulated samples in the b-taged event category. Here "Other" refers to the sum of all remaining background samples: $W \rightarrow \ell\nu$, dibosons, $Z \rightarrow \ell\ell$ and single top quark processes. The signal produced in association with b-quarks and via gluon fusion is considered separately assuming $m_A = 150$ GeV and $\tan \beta = 20$.

Generator	Process	Uncertainty
ALPGEN	$Z \rightarrow \tau\tau/ee/\mu\mu$	$\pm 5\%$
POWHEG	$t\bar{t}$	$\pm 5.5\%$
ALPGEN	$W \rightarrow \tau\nu/e\nu/\mu\nu$	$\pm 5\%$
AcerMC	single top	$\pm 13\%$
HERWIG	dibosons	$\pm 6\%$
SHERPA	$bbA/h/H$ ($m_A \geq 120$ GeV)	$-(< 20)\%, +(< 9) \%$
SHERPA	$bbA/h/H$ ($m_A = 110$ GeV)	$-(< 25)\%, +(< 9) \%$
SHERPA	$bbA/h/H$ ($m_A = 100$ GeV)	$-(< 28)\%, +(< 9) \%$
SHERPA	$bbA/h/H$ ($m_A = 90$ GeV)	$-(< 30)\%, +(< 9) \%$
POWHEG	$ggA/h/H$ ($m_A \leq 300$ GeV)	$< 15\%$

Table 4.12: Cross-section uncertainties for background and signal processes, $\tan \beta = 20$ is assumed for all signal samples.

to be 2% [138]. The acceptance uncertainties on dibosons and single top quark production are estimated to be 2% [137].

Uncertainties on the signal acceptance have been estimated with signal samples produced varying generator parameters. The impact on the selection of leptons, τ -jet and jets is evaluated at the particle level, prior to simulation of the detector response. This truth-level study is implemented within the Rivet framework [140], where the b-tagging is performed by identifying the b-quarks and applying weights according to the measured ATLAS b-tagging efficiencies [117]. The variation of the acceptance with respect to the nominal Monte Carlo tune has been considered as a source of systematic uncertainty. For signal a total acceptance uncertainty varies from 4% to 30% depending on M_A , production process and on the analysis category.

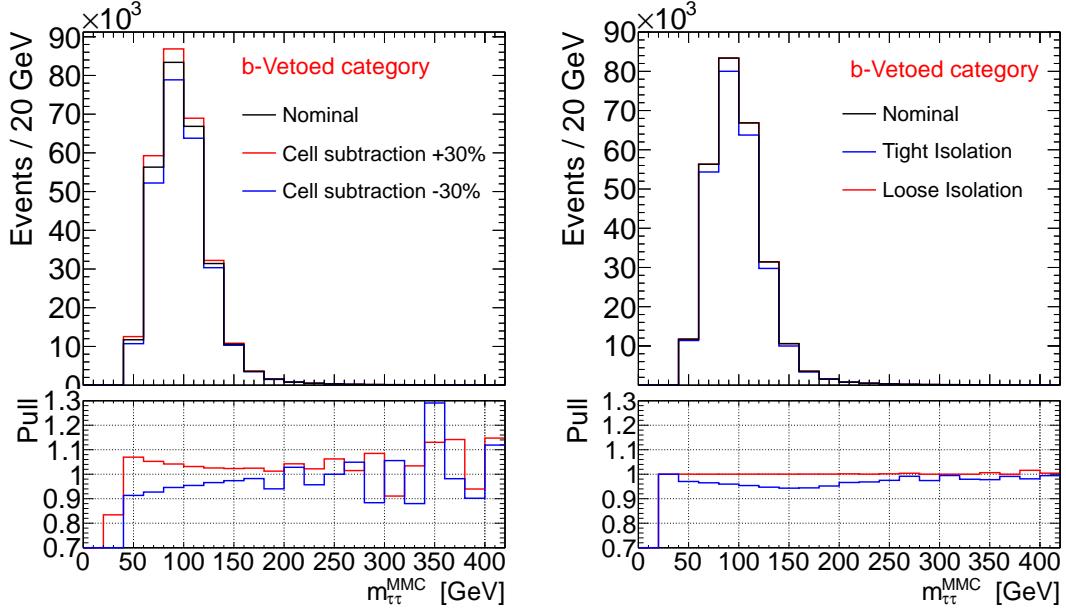


Figure 4.13: Impact of EMB_MFS (left) and EMB_ISO (right) systematic uncertainties on the $m_{\tau\tau}^{MMC}$ mass distribution for in the embedded sample. Significant impact is observed only in the b-vetoed event category.

4.4.3 Systematic Uncertainties of the $Z/\gamma^* \rightarrow \tau\tau$ Embedded Sample

An important element of the embedding method is the subtraction of the calorimeter cells associated with the muons in the original $Z \rightarrow \mu\mu$ event and their substitution with those from the simulated τ lepton decays. To make a conservative estimate of the systematic uncertainty on this procedure, the energy of the subtracted cells is scaled up or down by 30%. The analysis is repeated with those modified samples and the relative uncertainty is referred as "EMB_MFS". This uncertainty affects mainly the shape of the $m_{\tau\tau}^{MMC}$ mass distribution, shown in Figure 4.13.

In the sample of $Z \rightarrow \mu\mu$ candidates used for the embedding, only a loose requirement on the muon track isolation is applied. A different muon isolation requirement may affect the selected sample by modifying the topology of the event,

Selection	R_{QCD}	R_{QCD}^{AB}	R_{QCD}^{iso}
common selection	1.929 ± 0.004	2.12 ± 0.17	2.22 ± 0.16
No b-tagged jets	1.965 ± 0.005	2.10 ± 0.16	2.22 ± 0.16
Exactly one b-tagged jet	1.78 ± 0.02	1.9 ± 0.9	2.0 ± 0.8

Table 4.13: Comparison between R_{QCD} , R_{QCD}^{AB} and R_{QCD}^{iso} after the common selection stage and after requiring or vetoing the presence of b-tagged jets. Only the statistical uncertainty is reported here. The value of R_{QCD}^{iso} is reported for the lepton isolation threshold which is twice the nominal value, while for R_{QCD}^{AB} and R_{QCD} the nominal values are reported.

changing the contamination with other processes or the activity in the calorimeter. To estimate the importance of these effects in the embedded sample, the muon isolation criteria used in the original $Z \rightarrow \mu\mu$ sample are tightened, while an even looser selection would have a rather small impact due to isolation requirements at the trigger level. The resulting uncertainty, referred to as "EMB_ISO", affects both the event yield and the shape of the $m_{\tau\tau}^{MMC}$ mass distribution in the embedded sample, as shown in Figure 4.13.

Finally, because the normalization of the embedded sample is determined from the simulated ALPGEN sample, the uncertainties related to the cross section and to luminosity are assigned. In addition all detector-related systematic uncertainties relevant to the decay products of the simulated τ lepton decay are propagated to the embedded sample.

4.4.4 QCD Multi-Jet Systematic Uncertainties

The QCD multi-jet background is estimated via the ABCD method, as described in Section 4.3.2. This technique relies strongly on the assumption that the lepton isolation variables are uncorrelated to the relative charge of the two leptons. Systematic uncertainties are assigned to take into account possible deviations from this assumption. First the correlation between the ratio R_{QCD} and the lepton isolation

criteria is considered, then the result is compared with an auxiliary measurement.

Figure 4.14 shows the R_{QCD} factor, i.e. the ratio between the QCD yields in data samples C and D, as a function of a sliding lepton isolation threshold relative to the nominal analysis selection (red points). As described previously, the expected contamination of non-QCD background processes is subtracted from the data in samples C and D. To estimate the uncertainty on the value of R_{QCD} an additional transfer factor is defined: $R_{QCD}^{iso} = N_{\hat{A}}/N_{\hat{B}}$, where \hat{A} and \hat{B} are semi-isolated OS and SS samples defined by requiring the lepton isolation to be larger than the nominal one, but smaller than a given sliding threshold value (defined by the X -axis of the plot). Also here the non-QCD contributions are subtracted from the data yields. The semi-isolated samples \hat{A} and \hat{B} are chosen given the high contamination with non-QCD background processes and with possible signal in data samples A and B. Figure 4.14 shows R_{QCD}^{iso} as a function of the relative lepton isolation threshold (black points). The difference between R_{QCD} and R_{QCD}^{iso} in the vicinity of the nominal isolation threshold is then assigned as a systematic uncertainty on R_{QCD} . For the lepton isolation threshold which is twice the nominal value, a systematic uncertainty of 15% is found. The result shown in Figure 4.14 are obtained after common selection. Similar results after the full analysis selection in the two event categories are shown in Appendix B.

As a test of the result described above an additional measurement is performed. The R_{QCD}^{AB} is calculated as the ratio between the estimated QCD multi-jet contributions in samples A and B instead of C and D. The non-QCD contributions are subtracted from data. Due to the large contribution of this non-QCD background, along with small numbers of observed events and possible signal contamination, this measurement is only used for cross check. Table 4.13 shows a comparison between R_{QCD} , R_{QCD}^{iso} and R_{QCD}^{AB} after the common selection stage and after requiring or vetoing the presence of b-tagged jets, at these selection stages the signal contamination is negligible. Good agreement is seen between the results of these methods.

The shape of the $m_{\tau\tau}^{MMC}$ mass distribution differs between the OS and SS in anti-isolated samples (C and D) as shown in Figure 4.15. The size of this

effect is within the above R_{QCD} uncertainty for the relevant mass range of the QCD multi-jet background (QCD multi-jet background contribution is negligible for $m_{\tau\tau}^{MMC} > 150$ GeV) and hence no correction factor is applied to the mass shape of sample B. It is assumed, however, that there could be the same shape difference in the isolated samples and thus a shape uncertainty is assigned to the mass distribution in sample B to account for this deviation. Further shape uncertainties due to non-QCD background subtraction are found to be negligible. The uncertainty due to the use of an isolation requirement at the trigger level is discussed in Appendix A and is found to be negligible.

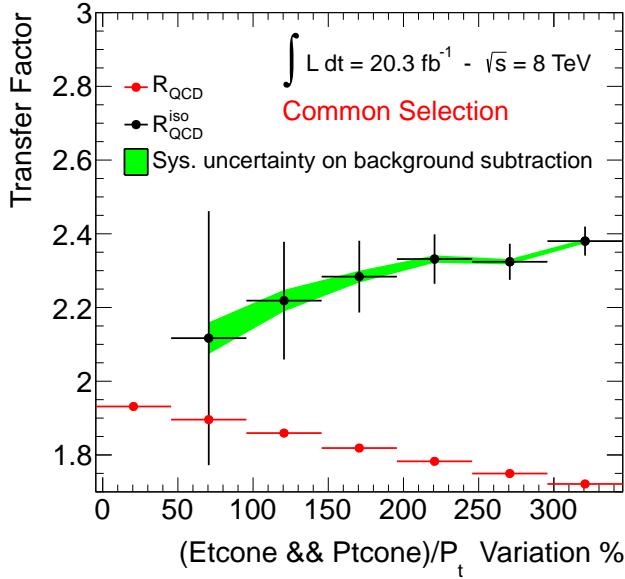


Figure 4.14: Transfer factors R_{QCD} and R_{QCD}^{iso} (see text) as a function of the sliding lepton isolation thresholds. The thresholds are varied in percentages relative to the nominal lepton isolation threshold (value of zero on the plot). The common selection are applied.

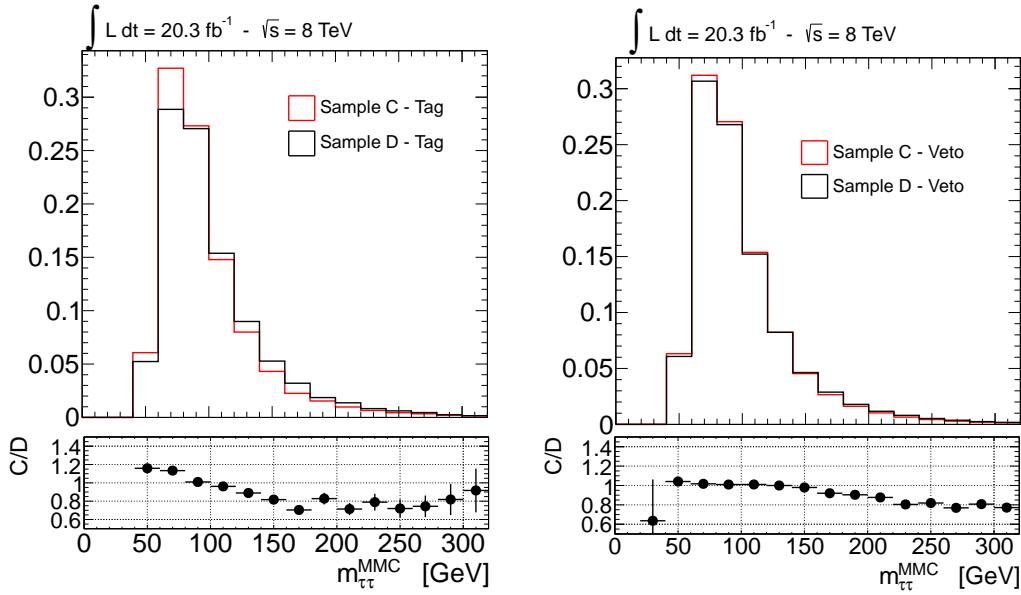


Figure 4.15: Differences in the shape of the invariant $m_{\tau\tau}^{MMC}$ mass distribution in data samples C and D shown separately for the b-taged and b-vetoed event categories. The data samples C and D are normalised to the same number of events.

4.5 Results

4.5.1 Statistical Procedure

The statistical interpretation of data in the presented search is based on profiled likelihood ratio test statistic used for the Higgs boson searches [141]. The statistical procedures described in the following are implemented in the software packages described in [142–144].

The Likelihood Function

The statistical interpretation of data is performed by means of testing the compatibility of the *background only* and the *signal-plus-background* hypotheses with the observed data. The main ingredient of the hypothesis *test statistic*, defined

later on in this section, is a binned likelihood function for the data set, \mathcal{D} :

$$\mathcal{L}(\mathcal{D} \mid \mu, \boldsymbol{\theta}) = \text{Pois}(\mathcal{D} \mid \mu \cdot s(\boldsymbol{\theta}) + b(\boldsymbol{\theta})) \cdot \prod_{j,i} \mathcal{G}(\theta_j^{sys} \mid 0, 1) \cdot \Gamma(\theta_i^{stat} \mid \beta_i) \quad (4.4)$$

describing how likely is a certain hypothesis given the observation of the data set \mathcal{D} . The signal strength modifier μ allows for reproducing a continuous set of signal hypotheses with different cross-section. The value of $\mu = 0$ corresponds to the background-only hypothesis. The vector $\boldsymbol{\theta}$ represents the set of *nuisance* parameters related to the systematic (θ_j^{sys}) and to the statistical (θ_i^{stat}) uncertainties of the background and signal predictions. The functions $s(\boldsymbol{\theta})$ and $b(\boldsymbol{\theta})$ represent the expected signal and background distribution respectively, these are binned histograms of the invariant $m_{\tau\tau}^{MMC}$ mass distribution. The function $\mathcal{G}(\theta_j^{sys} \mid 0, 1)$ is the Gaussian³ probability density function (p.d.f.) for the nuisance parameter θ_j^{sys} , which is assumed to be distributed with mean = 0 and $\sigma = 1$. The impact of the corresponding systematic uncertainty on the signal and backgrounds yields and on the shape of the $m_{\tau\tau}^{MMC}$ invariant mass distribution is evaluated separately as described in Section 4.4. The function $\Gamma(\theta_i^{stat} \mid \beta_i)$ is an extended gamma function⁴ describing the p.d.f. for the nuisance parameter θ_i^{stat} related to the statistical uncertainty β_i , for the bin i of the considered histogram. Each value of the nuisance parameter set $\boldsymbol{\theta}$ is associated with a variation of the predicted signal and background event yields with respect to the nominal prediction. The Poisson distribution in equation (4.4) is a binned p.d.f. and stands for a product of Poisson probabilities to observe n_i events in the bin i of the $m_{\tau\tau}^{MMC}$ mass distribution histogram:

$$\text{Pois}(\mathcal{D} \mid \mu \cdot s(\boldsymbol{\theta}) + b(\boldsymbol{\theta})) = \prod_i \frac{(\mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta}))^{n_i}}{n_i!} e^{-\mu s_i(\boldsymbol{\theta}) - b_i(\boldsymbol{\theta})}$$

The $m_{\tau\tau}^{MMC}$ mass distributions in the b-tagged and b-vetoed category are analysed separately. The actual implementation in the likelihood function of the ABCD

³Evaluation of systematic uncertainties is obtained from auxiliary measurements. From Bayes theorem, assuming a flat prior and a Gaussian distribution for the measured parameter a Gaussian posterior is obtained.

⁴The posterior of a Poisson distribution assuming a flat prior is a gamma function

method follows that suggested in [135] and it is described in more detail in Appendix C.

Statistical Combination of Results From Two Event Categories

Complementary event categories in one search channel, like the b-tagged and b-vetoed event categories, can be combined in order to increase the sensitivity to the signal. If there are no events entering both categories, as is the case for the presented analysis, the statistical combination is the product of the likelihood functions for the individual categories. For the combination, the convention described in [141] is used to take into account the correlation between different sources of uncertainties. Uncertainties are considered either as fully correlated, which means that the same nuisance parameter is describing the given systematic effect in both categories, or fully uncorrelated, in which case different nuisance parameter are employed in the two categories. Partially correlated uncertainties are either split into component which are fully uncorrelated or they are defined as to either fully correlated or uncorrelated, depending on which assumption is the most conservative.

The Test Statistic and the Exclusion Limits

To compute the compatibility of the data with a given hypothesis a test statistic is defined based on the *profiled likelihood ratio* [145]:

$$\tilde{q}_\mu = -2\ln \frac{\mathcal{L}(\mathcal{D} | \mu, \hat{\boldsymbol{\theta}}_\mu)}{\mathcal{L}(\mathcal{D} | \hat{\mu}, \hat{\boldsymbol{\theta}})} \quad \text{with the constraint } 0 \leq \hat{\mu} \leq \mu. \quad (4.5)$$

where \mathcal{L} is the likelihood function defined in equation (4.4), $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ are the global maximum likelihood estimators for μ and $\boldsymbol{\theta}$ given the data, whereas $\hat{\boldsymbol{\theta}}_\mu$ is the maximum likelihood estimator of $\boldsymbol{\theta}$ given the data but with a signal strength fixed to the value of μ . \tilde{q}_μ is increasing with increasing disagreement between data and the given μ hypothesis. Based on this, the procedure for setting upper exclusion limits on the signal cross-section is defined as follows:

1. The probability density function of \tilde{q}_μ is determined under the background-only (H_0) and the signal-plus-background (H_μ) hypotheses for any given value of μ . Since the determination of these distributions by means of pseudo-data demands large computing resources, asymptotic approximation formulas described in [145] are employed.
2. Once the p.d.f. for the background-only and signal-plus-background hypothesis are obtained, it is possible to define for a given dataset two probability values (p-values) for any given value of μ . These are the probabilities to obtain data less compatible with the considered hypothesis than the actual observation:

$$p_{s+b} = P(\tilde{q}_\mu > \tilde{q}_\mu^{\text{observed}} \mid H_\mu)$$

$$p_b = P(\tilde{q}_\mu > \tilde{q}_\mu^{\text{observed}} \mid H_0)$$

The ratio of this two probabilities defines the quantity $CL_s = p_{s+b}/p_b$ [146, 147].

3. If for a given μ the CL_s value of $CL_s \leq \alpha$ is obtained, the signal-plus-background hypothesis (with the corresponding μ value) is considered to be excluded at a $(1 - \alpha)$ CL_s confidence level. The 95% confidence level upper limit on μ , denoted as $\mu^{95\%}$, is defined as the smallest value of μ for which the CL_s is no longer greater than 0.05 .

By construction, rejecting all values of $\mu > \mu^{95\%}$, the signal-plus-background hypothesis will be rejected when it is true at most 5% of the time. The CL_s prescription is a conservative approach protecting the signal exclusion upper limit from downward fluctuation of the data. The expected median exclusion upper-limit and its error are evaluated with the procedure described above under the background-only hypothesis. The obtained results have been cross-checked using generated pseudo-data instead of the asymptotic approximation for the determination of the \tilde{q}_μ probability density functions.

4.5.2 Exclusion Limits on the Signal Production

The statistical procedure described in section 4.5.1 is the general one used for the SM Higgs boson where only the Higgs boson mass determines the signal properties. For the MSSM further complication arises: there are three neutral Higgs bosons contributing to the total signal yield, in a particular scenario the masses and cross section are defined by two parameter $\tan\beta$ and m_A . Thus, the previously described procedure has to be repeated for each point in the $\tan\beta - m_A$ plane. For the m_h^{mod} scenario exclusion limits at 95% CLs confidence level are derived on the cross section for the neutral MSSM Higgs bosons ($A/h/H$) production via gluon fusion and in association with b-quarks, the considered Higgs bosons decay is $A/H/h \rightarrow \tau\tau \rightarrow \mu e + 4\nu$. A scan has been performed over for 15 $\tan\beta$ values ranging⁵ from $\tan\beta = 5$ to $\tan\beta = 60$. A point in the $\tan\beta - m_A$ plane is excluded if $\mu^{95} \leq 1$ for that point. A linear interpolation is used to determine the excluded $\tan\beta$ for a given m_A . The procedure is repeated for a set of 12 CP-odd Higgs boson masses values m_A ranging from 90 GeV to 300 GeV⁶. The expected and observed event yields are compared in bins of the $m_{\tau\tau}^{MMC}$ mass distribution. The bin sizes are chosen such that the number of events per bin is high enough to justify the use of the asymptotic approximation. Table 4.14 compares the expected and the observed event yields in the two event categories at the final stage of the event selection. Additionally, Figure 4.16 shows the corresponding $m_{\tau\tau}^{MMC}$ mass distributions.

The resulting exclusion limit on the MSSM $m_A - \tan\beta$ parameter space are interpreted within the m_h^{mod} benchmark scenario and shown in Figure 4.17. The expected and observed exclusion limits at 95% CL_s confidence-level are shown as dashed and solid black lines respectively. The green and yellow bands correspond to the 1σ and 2σ error bands on the expected exclusion limit. The analysis is sensitive to the MSSM Higgs boson production for $\tan\beta \geq 13$ and the mass range $90 < m_A < 200$ GeV. The observed limit is

⁵ The set of $\tan\beta$ values used is 5, 8, 10, 13, 16, 20, 23, 26, 30, 35, 40, 45, 50, 55, 60

⁶ The set of m_A values used is: 90, 100, 110, 120, 125, 130, 140, 150, 170, 200, 250 and 300 GeV

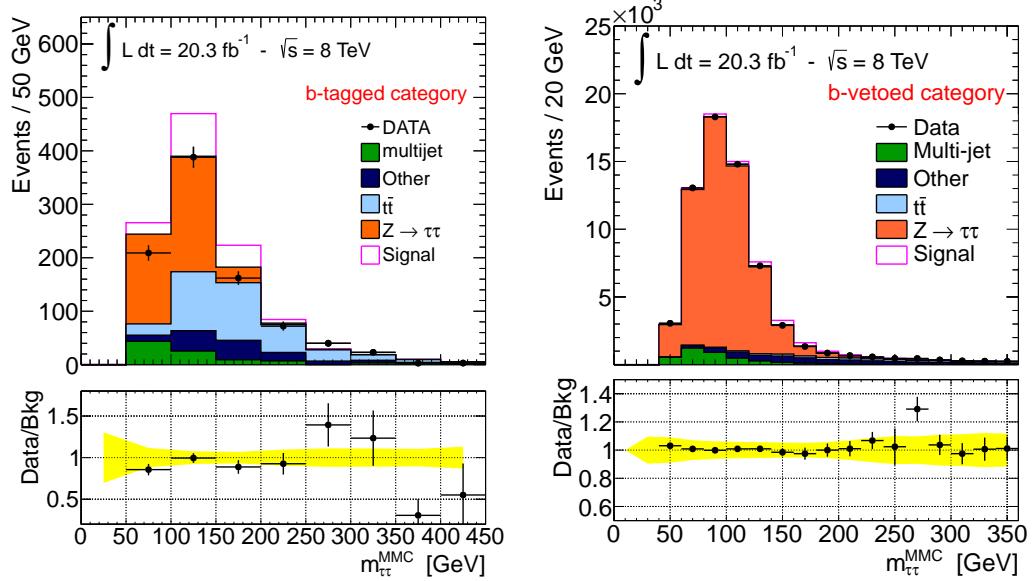


Figure 4.16: Observed and expected distribution of the $m_{\tau\tau}^{MMC}$ mass for (left) the b-tagged event category and (right) the b-vetoed event category after the full event selections. The prediction of the background model is compared to data. The contribution of the $Z/\gamma^* \rightarrow \tau\tau$ and QCD multi-jet background processes is measured in dedicated signal-depleted control data samples, the prediction for all the other background processes is obtained from simulation. The notation “Other” stands for the electroweak processes $W \rightarrow \ell\nu$, $Z \rightarrow \ell\ell$, diboson and single top quark production. The prediction for the signal is evaluated considering the production of the three neutral MSSM Higgs bosons in association with b-quarks and via the gluon fusion processes in the m_H^{mod} scenario for values of $m_A = 150$ GeV and $\tan\beta = 20$. The yellow band represents the total systematic uncertainty for the background model prediction.

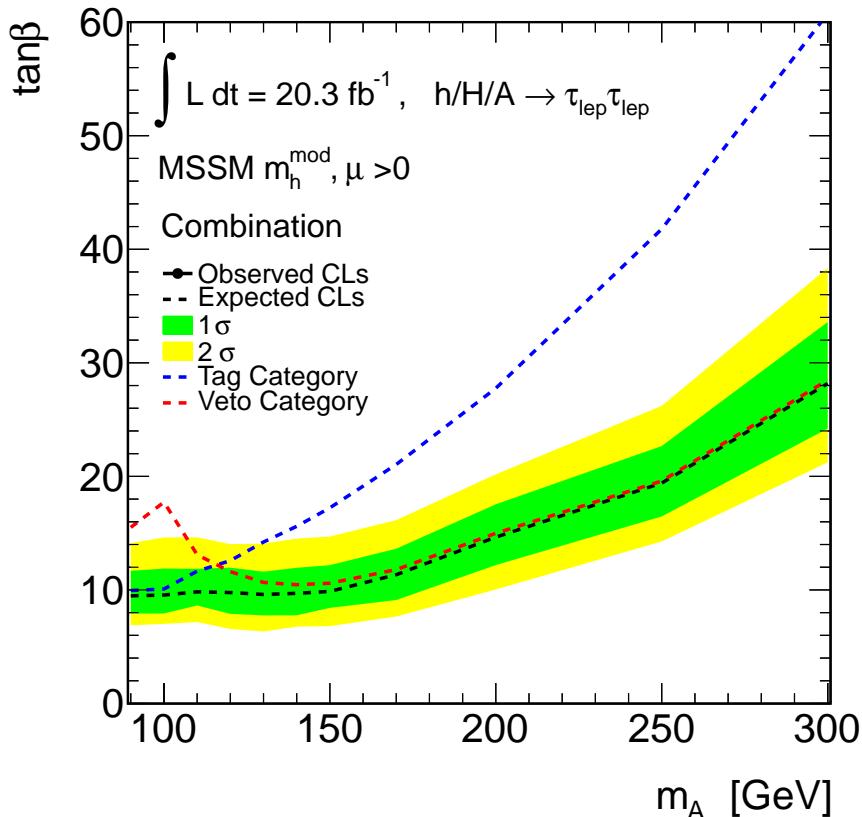


Figure 4.17: Expected and observed exclusion limits at the 95% CL_s confidence-level for MSSM Higgs bosons production interpreted in the $m_A - \tan\beta$ parameter space of the m_h^{mod} scenario. Combined result of the b-tagged and b-vetoed category is shown.

Sample	b-tag category			b-veto category		
	N(event)	Stat.	Syst.	N(event)	Stat.	Syst.
$Z/\gamma^* \rightarrow \tau\tau$	418	± 6	$^{+27}_{-27}$	54680	± 60	$^{+3500}_{-3500}$
$t\bar{t}$	330	± 10	$^{+37}_{-35}$	2159	± 25	$^{+280}_{-300}$
Multijet	101	± 15	$^{+15}_{-15}$	3930	± 330	$^{+590}_{-590}$
Other	114	± 9	$^{+12}_{-12}$	4450	± 110	$^{+250}_{-250}$
Total	963	± 21	$^{+50}_{-50}$	65220	± 360	$^{+3600}_{-3600}$
Signal	144	± 7	$^{+24}_{-33}$	2028	± 27	$^{+150}_{-100}$
Data	-	-	-	-	-	-

Table 4.14: Expected and observed number event yields in the b-tagged and b-vetoed event category after the full event selections. The various background and signal expected event yields are normalized to the integrated luminosity of the data sample (20.3 fb^{-1}). The notation ‘‘Other’’ stands for the electroweak processes $W \rightarrow \ell\nu$, $Z \rightarrow \ell\ell$, diboson and single top quark production. $Z/\gamma^* \rightarrow \tau\tau$ has been estimated using the embedding technique. The uncertainties quoted include the statistical uncertainty (first number) and the systematic uncertainties (second number). For the prediction of the MSSM Higgs bosons signal yield, $m_A = 150 \text{ GeV}$ and $\tan \beta = 20$ is assumed in the m_h^{mod} scenario.

The outcome of the search is also interpreted in a model-independent way, by setting the limits on the production cross-section of a scalar boson produced in the $pp \rightarrow gg \rightarrow \phi$ or $pp \rightarrow bb\phi$ mode and decaying in a di-tau pair. The corresponding expected and observed 95% CL_s confidence-level limits are shown in Figure 4.18. The limits in the production in association with b-quarks and via gluon fusion are shown separately. More information on the limit setting procedure and its validation can be found in Appendix C.

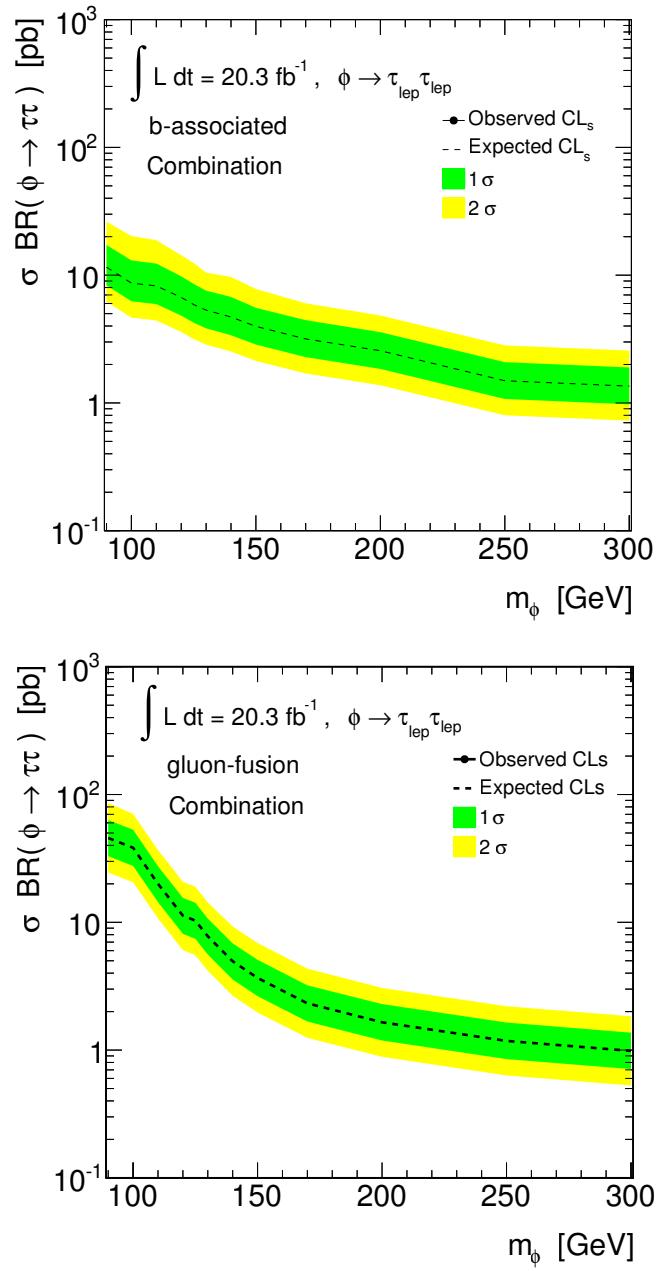


Figure 4.18: Limits on the production of a scalar particle decaying to a di-tau pair and produced in association with b quarks (top) or via gluon-gluon fusion (bottom).

Chapter 5

Prospects for Improved Higgs Search with b-tagging of Track-Jet

The search for the neutral MSSM Higgs bosons described in the previous chapter suffers strongly from poor b-tagging performance due to the relatively low energy of the b-jet produced in association with the signal. Improvements on the b-tagging performance would result in a major improvement of the search sensitivity. This chapter investigates an alternative b-jet identification procedure in which the b-tagging algorithm is applied on track-based jets instead on the commonly used calorimeter jets. While the calorimeter jets are reconstructed from the energy clusters in the calorimeter, the track-based jets consist of inner detector tracks. The performance of the b-tagging on track-based jets is investigated for the first time. The prospects for improvements of the neutral MSSM Higgs bosons search due to the employment of track-based jets for b-tagging are reported. In Section 5.1 the challenges of the MSSM Higgs boson search due to b-tagging are presented along with a description of track-based jet reconstruction. Section 5.2 describes the performance of the b-tagging algorithms with track-based jets in comparison to the nominal b-tagging with calorimeter jets. Preliminary evaluation of the corresponding impact on the analysis is also described

here. Finally, in Section 5.3 a description of the related track-based jets systematic uncertainty is given.

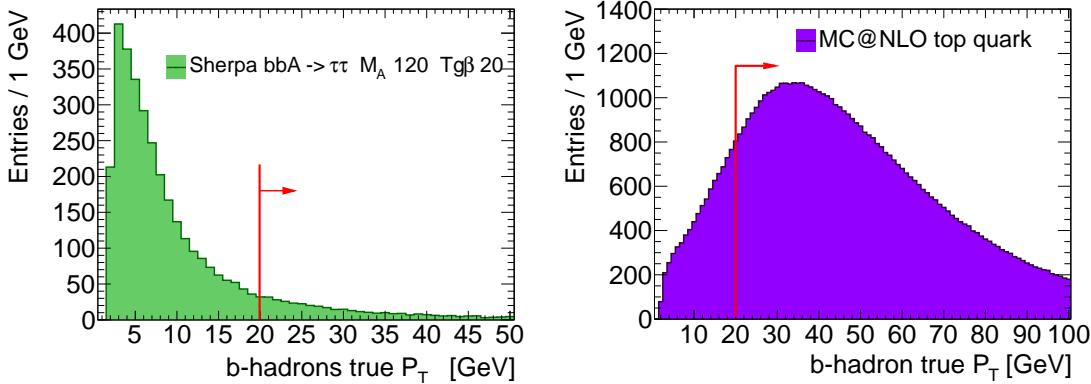


Figure 5.1: Simulated p_T distribution of b-hadron for the neutral MSSM Higgs boson produced in association with b-quarks (left) and for $t\bar{t}$ events (right). The red line in the figure shows the acceptance region in which the energy of the calorimeter jets can be calibrated.

5.1 Track-based Jets

In the search for the neutral MSSM Higgs bosons, as described in chapter 4, the selected events are categorized in two categories based on the presence or the absence of a b-tagged jet in the event. The b-tagged event category is optimized for the selection of Higgs bosons produced in association with b-quarks. Figure 5.1 shows a comparison between the p_T spectra of simulated b-hadron (hadrons containing a b-quark) in $pp \rightarrow b(b) A/h/H$ process and in $t\bar{t}$ events. The signal contains b-hadrons with relatively low transverse momenta, imposing a major challenge for the analysis of the b-tagged event category. Due to the high amount of pileup and ambient energy density in the events, the energy of the calorimeter jets with $p_T < 20$ GeV is not calibrated (see chapter 3), the jet reconstruction performance and related systematic uncertainties are not evaluated for these jets. The reconstruction of calorimeter jets in the $pp \rightarrow b(b) A/h/H$ signal production process is therefore not optimal and represent the major source of the sensitivity losses in the b-tagged event category. An additional challenge of the presented search is the worsening of the b-tagging performance for jets with low transverse momenta. The

efficiency obtained with the *MV1* tagger decreases rapidly with jet p_T , reaching a minimum of 50% at 20 GeV for the tagging point with $\epsilon_b^{t\bar{t}} = 70\%$ [117, 118].

The performance of the jet reconstruction for low transverse momenta may be improved by introducing *track-based* jets (in the following track-jets) in replacement to the commonly used calorimeter jets. Track-jets are reconstructed with the anti- kt algorithm by clustering inner detector tracks. Conversely to calorimeter clusters, tracks have associated impact parameter informations and track-jets can be reconstructed with tracks originating from the same interaction point, making their reconstruction very robust against the impact of pile-up.

Track-jets are built by the *TrackZTool* reconstruction software which runs the anti- kt clustering algorithm on a subset of user-defined tracks. For the purposes of this thesis track-jets are reconstructed from tracks satisfying the following quality selection criteria:

- The track should be associated to the primary vertex (PV), $|z_{track} - z_{PV}| < 2$ mm, where z_{track} and z_{PV} are the absolute z coordinate of the track and of the primary vertex, respectively.
- The track is required to point to the PV in the plane containing the beam axis by $|z_{PV} \cdot \sin(\theta)| < 1.5$ mm, where θ is the angle between the track and the beam axis.
- The distance of minimum approach of the track to the primary vertex in the plane orthogonal to the beam axis is required to be $d_{PV} < 1.5$ mm.
- At least one pixel hit and at least 6 SCT hits (including SCT holes) should be detected for each track.
- A b-layer hit should be present if the b-layer module passed by the track was active.
- The pseudorapidity of the track is required to be $|\eta| < 2.5$, corresponding to the coverage of the inner detector.

Process	MC Generator	Purpose
Minimum bias	Pythia	Systematics study
$b\bar{b}$	Alpgen	Performance for low p_T b-tagging
$Z/\gamma^* \rightarrow \tau\tau$	Pythia	Impact on the MSSM Higgs search
$t\bar{t}$	MC@NLO	Impact on the MSSM Higgs search
MSSM $bb/A/h/H$	Sherpa	Impact on the MSSM Higgs search

Table 5.1: Monte Carlo simulation samples.

- The track transverse momentum should be $p_T > 300$ MeV to ensure a low track fake rate.

A track-jet is seeded by a cluster of at least two tracks which satisfy the above selection criteria, the sum of the transverse momenta of all associated tracks is required to be $\sum_i p_{T,i} > 2$ GeV. It has been shown that the above selection criteria, together with a maximum cone size for clustering of $\Delta R = 0.6$, give the best compromise between the power of rejecting fake tracks and the b-hadron reconstruction efficiency. Several simulation samples have been produced to study the performance of the track-jets reconstruction and of the b-tagging procedure applied to these jets. Table 5.1 gives a list of the produced samples along with the type of studies performed with them.

B-tagging has never been tested before on track-jets, in section 5.2 the first study of b-tagging over track-jets performances is reported.

5.2 Performance of the Track-based Jets Reconstruction and b-tagging

5.2.1 Track-based Jets Reconstruction

Many analysis could profit from an enhanced b-jet reconstruction efficiency at low values of p_T . The studies presented in this section are aimed at comparing the performance of the b-jet reconstruction efficiency and the common b-tagging

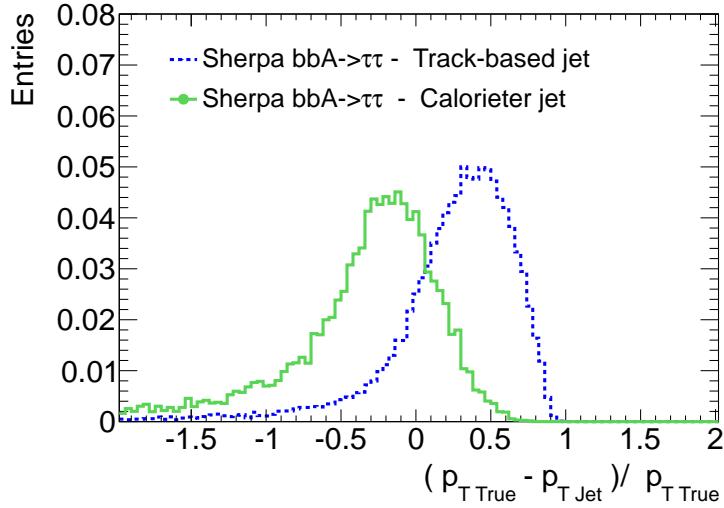


Figure 5.2: Distribution of transverse momentum residuals relative to the true value $P_{T\,true}$, shown separately for the calorimeter and track-jets.

algorithm for the calorimeter and track-based jets, focusing in particular on jets with low transverse momenta.

Even though the track-jets are more robust against pile-up effects, which is the main reason for not use calorimeter jet at low transverse momenta, they contain only the charged fraction of the jet, while the neutral jet component is lost. According to isospin invariance the expected charged fraction in a jet amounts to roughly 2/3 of the total energy. The track-jet momentum is therefore shifted accordingly and there is a larger uncertainty on its measured direction. Figure 5.2 shows the distribution of transverse momentum residual $p_{T\,true} - p_{T\,jet}$ relative to the truth value $p_{T\,true}$ from *truth-jet* (see chapter 3) separately for calorimeter and track-jets. Here truth-jets are matched with reconstructed jets within a cone of size $\Delta R = 0.4$ ¹. As expected the track-jets energy is shifted from zero. This effect can be critical for most of the b-tagging algorithms since the likelihood functions used for decision making are derived separately for different region of jets p_T and pseudorapidity. A dedicated track-jets calibration of b-tagging algorithm is auspucable for future application of b-tagging on these jets.

¹jet splitting effects are resolved by matching with the nearest jet

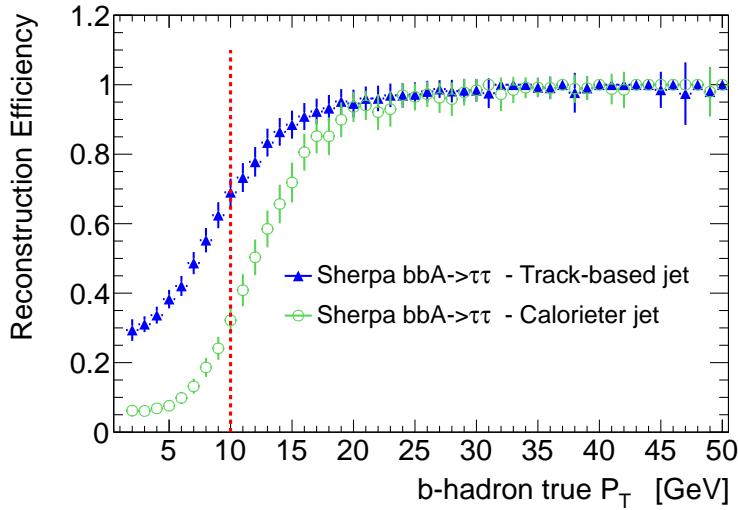


Figure 5.3: b-hadron reconstruction efficiency for track-jet and calorimeter jets as a function of the true b-hadron p_T . Note that calorimeter and track-based jet are required to have at reconstruction level $p_T > 7$ and 2 GeV, respectively. A fair comparison in this plot is only possible for $p_T > 10$ GeV.

To compare the performance of calorimeter and track-jet reconstruction, an anti-kt algorithm with cone size of $\Delta R = 0.4$ is chosen. If the angular distance between the reconstructed jets and a simulated b-hadron in the event is $\Delta R < 0.3$, this jet is said to *match* with a b-hadron. b-hadron *Reconstruction efficiency* is then defined as the ratio between the number of matched b-hadrons and the total number of b-hadrons within inner detector acceptance. Figure 5.3 shows the b-hadron reconstruction efficiency for the calorimeter and track-jets. The latter exhibit a higher reconstruction efficiency at low transverse momenta due to their robustness against pile-up effects.

5.2.2 B-tagging with Track-Based Jets

Performance of the b-tagging algorithms is usually described in terms of the b-tagging efficiency and rejection power against misidentified jets. The *b-tagging efficiency* is the fraction of jets matched to a true b-hadron which pass a given tagging selection criteria, i.e. which are *b-tagged*. The *rejection power* is the in-

verse of the misidentification rate, i.e. the inverse of the fraction of jets which are not matched with a b-hadron or c-hadron, but are b-tagged. Figure 5.4 shows the rejection power as a function of the b-tagging efficiency for various b-tagging algorithms applied on track-jets and calorimeter jets separately. Figure 5.5 shows the rejection power as a function of jet p_T for the b-tagging working point which gives 50% b-tagging efficiency, calorimeter and track-based jet are shown separately. Mistagging rate is rapidly increasing for low transverse momentum jets due to increasing particle multiple scattering and secondary interactions in the material, revealing the necessity of a dedicated b-tagging algorithm for low p_T jets.

The described rejection power and b-tagging efficiency cannot serve for a fair comparison of the track-based and calorimeter jets. The latter can be reconstructed even if there are no associated tracks to them, in which case any b-tagging algorithm will most likely fail. The distribution of the rejection power is therefore altered by such jets. It is instead convenient to introduce the: *effective rejection power*, which is the inverse of the number of mistagged jets per event. Figure 5.6 shows the effective rejection power as a function of the b-hadron reconstruction efficiency, for calorimeter and track-based jets. For a given b-hadron reconstruction efficiency, a higher effective rejection of mistagged jets can be achieved by track-jets. For a fair comparison with calojets the track-jets in Figure 5.6 are selected in the transverse momentum range between 4 and 33 GeV, while the transverse momentum of calorimeter jets ranges from 8 to 50 GeV. The introduced p_T -thresholds corresponds in average to the same p_T range, Figure 5.2 is only valid for low p_T jets and the fraction of undetected momentum from neutral jet component approaches 1/3 for higher p_T track-jets. In conclusion, for jets with low transverse momentum the track-jets provide a higher b-hadron reconstruction efficiency than calorimeter jets and are more suitable for low p_T b-tagging.

5.2.3 Employment of Track-jet for the MSSM Higgs Boson Search

The impact of the track-jets selection on the search for the neutral MSSM Higgs bosons is tested in a preliminary study and reported in the following. Common

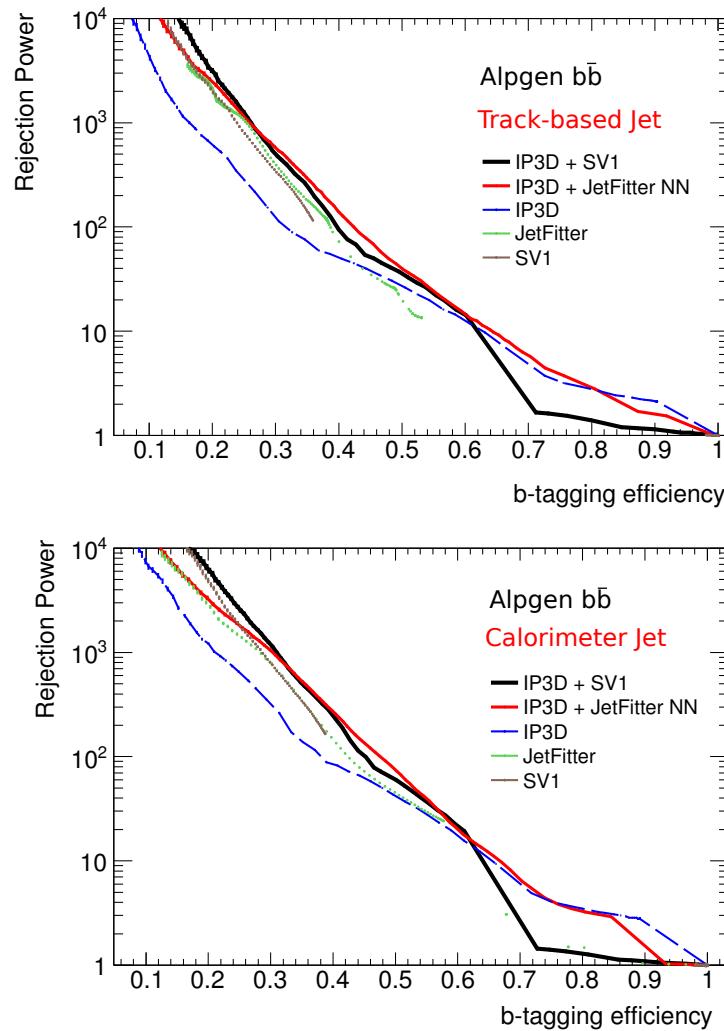


Figure 5.4: Rejection power as a function of the b-tagging efficiency for different b-tagging algorithm applied on track-jets (top) and on calorimeter jet (bottom) for simulated $b\bar{b}$ events.

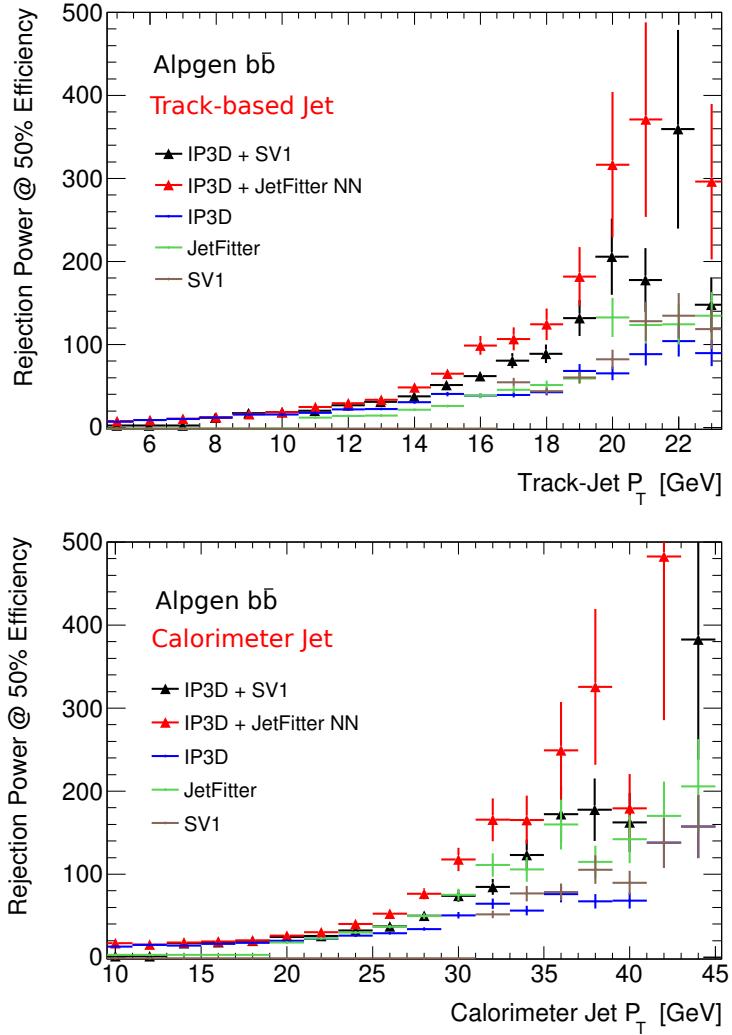


Figure 5.5: Rejection power as a function of the jet transverse momentum for the b-tagging working point with a 50% b-tagging efficiency at that p_T value, track-jet (top) and calorimeter jet (bottom) are shown separately. Results are obtained using simulated $b\bar{b}$ events and shown for several b-tagging algorithms.

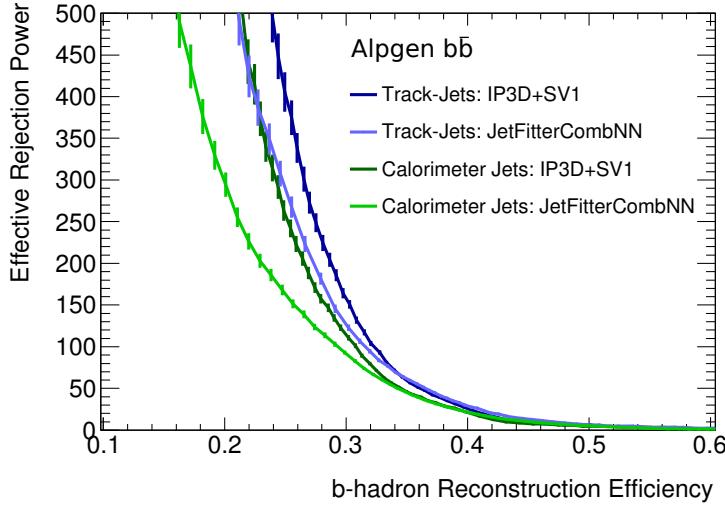


Figure 5.6: Effective rejection as a function of b-hadron reconstruction efficiency, track-jet and calo jets are compared for two different ATLAS tagging algorithms. Track-jets are selected in the transverse momentum range between 4 and 33 GeV, while calojets between 8 and 50 GeV.

selection criteria,², as defined in in Section 4.2.1, are applied to simulated signal and background samples with the following modifications concerning the jets on which b-tagging procedure is applied (taggable jets):

- Calorimeter jets should have $|\eta| < 2.5$ and $20 < p_T < 50$ GeV.
- Track-jets should have $|\eta| < 2.5$ and $5 < p_T < 33$ GeV.
- The b-tagging algorithm applied on the above jets is "IP3D+SV1" at a working point with a $\epsilon_b^{t\bar{t}} = 70\%$ tagging efficiency.

The minimum transverse momentum for which calorimeter jet are calibrated is 20 GeV, track-jets instead can safely access transverse momentum up to 5 GeV. The event yields expected for the $pp \rightarrow b(b)A/h/H$ signal process, $Z/\gamma^* \rightarrow \tau\tau$ and $t\bar{t}$ background processes (the two most important background contributions in the

²This study has not been updated with the latest version of the object reconstruction selections and corrections, a difference of the order of 10% is expected with respect the numbers in table 4.4.

Selection	Signal $bbA/H/h$		$Z/\gamma^* \rightarrow \tau\tau$		$t\bar{t}$	
Common Selection	127.2 ± 2.2		3017 ± 8		2066 ± 5	
	Calo. jet	Track-jet	Calo. jet	Track-jet	Calo. jet	Track-jet
At least one taggable jet	47.3 ± 0.8	106.9 ± 1.8	1146 ± 3	2513 ± 7	1804 ± 4	2014 ± 5
Exactly one jet matched to a b-hadron	18.4 ± 0.3	46.7 ± 0.8	4.5 ± 0.3	18.2 ± 0.5	1054 ± 3	959.1 ± 2.3
Exactly one b-tagged jet	10.2 ± 0.1	21.0 ± 0.6	37.3 ± 0.5	107 ± 1	777 ± 4	630 ± 4

Table 5.2: Event yield for the signal and dominant background processes after different event selection requirements. The yields are shown separately for the calorimeter and the track-based jet reconstruction. The Higgs boson produced in association with b-quarks is simulated for $\tan\beta = 20$ and $m_A = 150$ GeV. The yields are normalized to an integrated luminosity of 1 fb^{-1} .

b-taged event category) are reported in Table 5.2. The yields are normalized to an integrated luminosity of 1 fb^{-1} . In addition to the event yields after the common selection stage other b-tagging related selection requirements are applied to study the impact of the replacement of calorimeter with track-based jets. As expected, after requiring exactly one b-tagged jet, the track-jets reconstruction results in a higher signal efficiency with a relatively similar rejection (10% higher) of top quark background compared to calorimeter-based jet reconstruction. However, lower transverse momentum threshold for the track-jets implies higher mistagging rates, as can be seen from an increase of $Z/\gamma^* \rightarrow \tau\tau$ background. This may also lead to a strong contamination of the QCD multi-jet background, even tough this is a minor background contribution in b-taged event category.

In conclusion, the neutral MSSM Higgs boson search presented in the previous chapter may be improved if track-jet reconstruction is applied instead of calorimeter based one. The sensitivity³ to the signal can be improved in this event category by about a factor two. However, to exploit the full power of this technique a dedicated calibration of the b-tagging algorithms is needed for the track-jets. Additional improvements of the b-tagging algorithms for low p_T b-jets are also desirable. Furthermore, systematic uncertainties of track-jet reconstruction need to be evaluated. A preliminary study, addressing some of the most important of such systematic uncertainties is reported in section 5.3.

5.3 Systematic Uncertainties of Track-jet Reconstruction

There are several sources of systematic uncertainties of the track-jet reconstruction that may contribute to the mismodeling of physics observables. These effects are briefly summarized in the following with an emphasis on uncertainties of the energy scale and reconstruction efficiency.

Uncertainty on the properties of simulated track-jets can arise from the Monte Carlo generator configurations, depending on the particular choice of PDF and fragmentation functions, or details of the parton shower and underlying event modeling, which have a significant impact on physics objects with low transverse momentum. These uncertainties can be evaluated by means of a dedicated analysis with the Rivet package [140]. They depend on the particular use of track-jets and need to be evaluated case by case.

Energy scale and resolution of single tracks is found to be very well modeled by simulation for tracks above 500 MeV [150]. Thus, uncertainty on the track-jet energy scale and resolution that arise from the mismodeling of the pattern

³Note that the sensitivity is estimate according to the s/\sqrt{b} ratio (where s and b are the signal and background yield respectively), considering a counting experiment without systematic uncertainties and with only two background processes, this corresponds to the maximal possible sensitivity achievable with the current b-tagging performance.

recognition procedure are considered to be negligible.

In a dense track environment different tracks may share same hits, leading to a degradation of the track momentum resolution, fake tracks and losses of track reconstruction efficiency. Mismodeling of the hit sharing among several tracks may in general affect the track-jet energy scale, resolution and reconstruction efficiency. Such effects has been studied in [152], where energy scale uncertainties for calorimeter jets is measured based on associated tracks. It has been shown that effects due to the merging of track hits are negligible for jets with $p_T < 300$ GeV.

Mismodeling of the inner detector material budget leads to the mismodeling of the to track reconstruction efficiency, which strongly affects also the track-jets reconstruction. A methodology to estimate the uncertainty of the energy scale and reconstruction efficiency for track-jets due to the mismodelling of the material budget is studied for the first time and presented in section 5.3.1.

5.3.1 Material Budget Uncertainty on Track-Based Jets Reconstruction

An obvious but rather inconvenient way to estimate the uncertainty due to the mismodeling of the inner detector material budget is to simulate the Monte Carlo samples relevant for a given analysis using several different ID material budgets. It can be shown that the mismodeling of material budget primarily influence the track reconstruction efficiency (see section 5.3.2). An alternative approach is therefore to modify the track reconstruction efficiency in a given sample according to the corresponding uncertainty [151, 153] and build track-jets from such new collection of tracks. A tool has been developed which randomly removes tracks according to the uncertainty on reconstruction efficiency. The track-jets which are build from this subset of tracks are called in the following *inefficient* track-jets.

The standard and inefficient track-jets are compared in a simulated sample of minimum bias processes. A set of "isolated" track-jet with cone size $\Delta R = 0.4$ are selected, where the isolation means that no other track-jet should be reconstructed within an angular distance of $\Delta R = 1$. Inefficient track-jets are then matched with the original track-jet in the same event, the matching fails if no inefficient track-jet

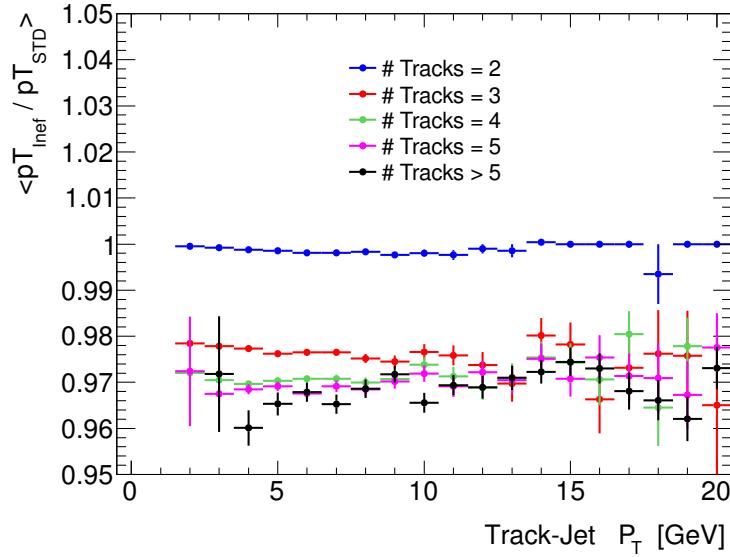


Figure 5.7: Inefficient track-jets are matched with standard track-jets. The average transverse momenta of the inefficient track-jet $p_{T\text{inef}}$ relative to the corresponding standard track-jet $p_{T\text{std}}$ is shown as a function of the p_T and number of associated tracks of the standard track-jet.

is found within a cone of size $\Delta R = 0.8$ around the original track-jet. The impact of tracking inefficiency on track-jet energy scale and reconstruction efficiency is presented respectively in Figure 5.7 and 5.8. These results are based on the current knowledge of the inner detector material budget [151]. Since track-jets are required to have at least two tracks at reconstruction level, if a track is lost that jet cannot be reconstructed anylonger, therefore for track-jet with two tracks the only effect is a loss of reconstruction efficiency. For track-jets with low transverse momentum, uncertainty on the material budget translates into an energy scale shift of 2-4% and in a reduction of the mean number of tracks.

This method can only simulate excess of material (reduced track efficiency) but not a lack of material (increased track efficiency). However, for the latter case a similar, symmetric impact is expected.

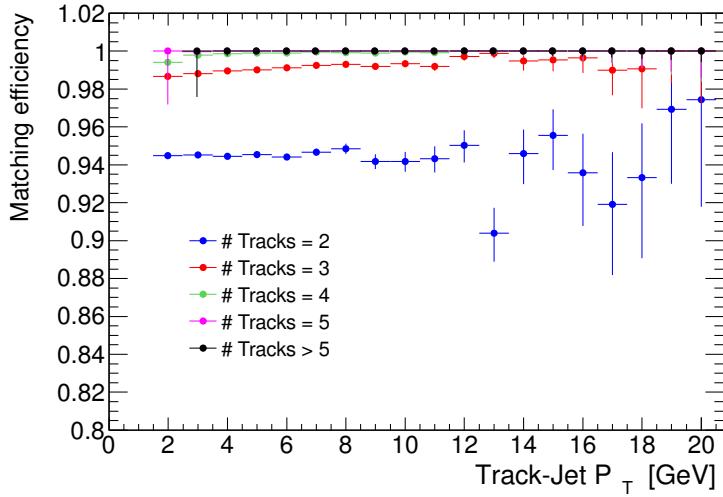


Figure 5.8: Inefficient track-jets are matched with standard track-jets. The matching efficiency is shown as a function of the p_T and number of tracks associated with the standard track-jet.

5.3.2 Validation of the Track Subtraction Method

The method described in section 5.3.1 depends strongly on the assumption that hadronic secondary interactions within the inner detector material lead mainly to the loss of some tracks and only in a marginal way to a decrease of the track quality. As a consequence, the misodeling of the material budget is expected to influence mainly the track reconstruction efficiency. In this section, the impact of the material budget uncertainty on the track momentum resolution and fake rate is evaluated using a simulated sample of minimum bias events, in which additional material is added to the ID increasing uniformly of 10% the interaction length.

For this study, the track selection is performed as in Section 5.1. Furthermore a track should be matched within a cone $\Delta R < 0.1$ with a stable⁴ simulated particle which alone should be causing at least 80% of all the track hits. Tracks that do not fulfill these requirements are considered as fake tracks. Fake tracks originate from a random combination of hits generated by different particles. The track fake rate

⁴This refers to a generator level stable and interacting particle, i.e. a charged particle with decay length greater than 1m. Also stable particles from secondary interactions are considered.

defined as the ratio between the number of fake tracks and the total number of selected tracks is about 1-3% and is shown in Figure 5.9. The additional material leads to a total increase of the track fake rate by about one permille. The track energy resolution as shown in Figure 5.10 is about 1% for a large range of track p_T values, the total degradation of the resolution in the presence of additional ID material is also of the order about one permille. The deterioration of the track energy resolution and fake rate due to the additional material budget is therefore negligible compared to the impact on the track reconstruction efficiency of about 1-2%. Decrease in the track reconstruction efficiency has a strong impact on the track-jet energy scale. Figure 5.11 shows the ratio of the track reconstruction efficiency for the primary particles assuming the nominal and additional material budget.

Reconstruction of inefficient track-jets in a sample with nominal material budget is also directly compared to the track-jets reconstruction in a sample with added additional ID material. Track-jet are matched to truth-jet (as described in section 5.2) in order to determine the track-jet reconstruction efficiency and energy scale, shown respectively in Figure 5.12 and 5.13. Inefficient track-jets reproduce correctly the impact of additional material, giving in most of the cases a conservative estimate of the corresponding systematic uncertainties.

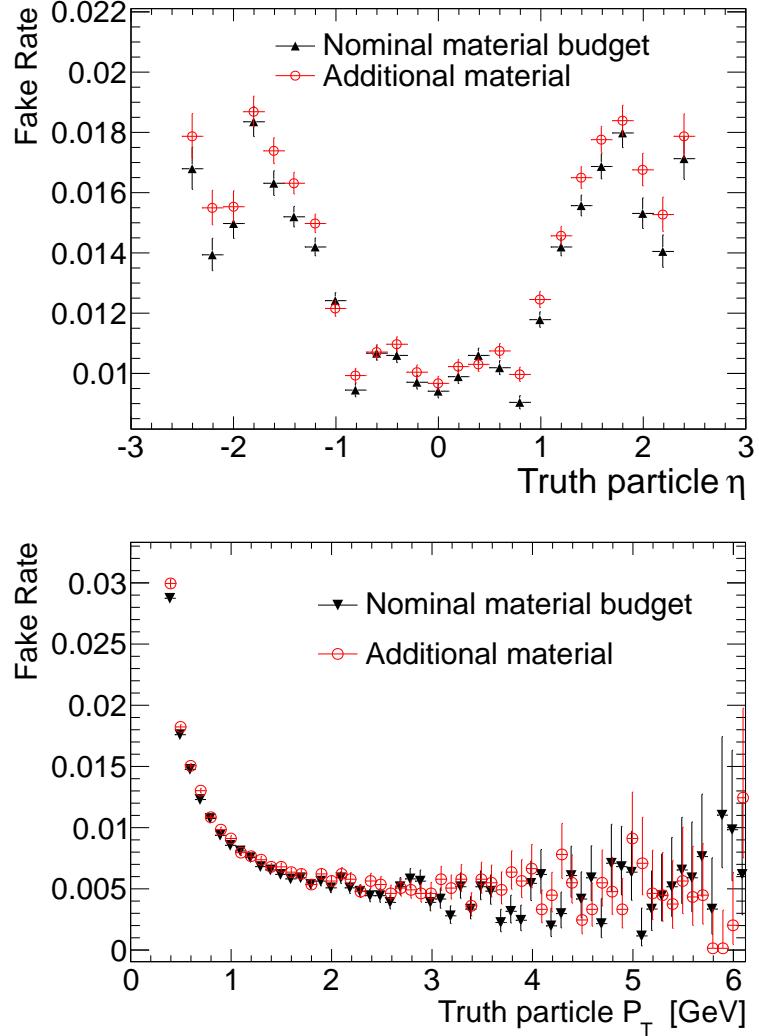


Figure 5.9: Track fake rate as a function of the associated truth particle η (left) and p_T (right). The results are shown for the two simulated samples of minimum bias processes, one with a nominal material budget and one with 10% additional inner detector material.

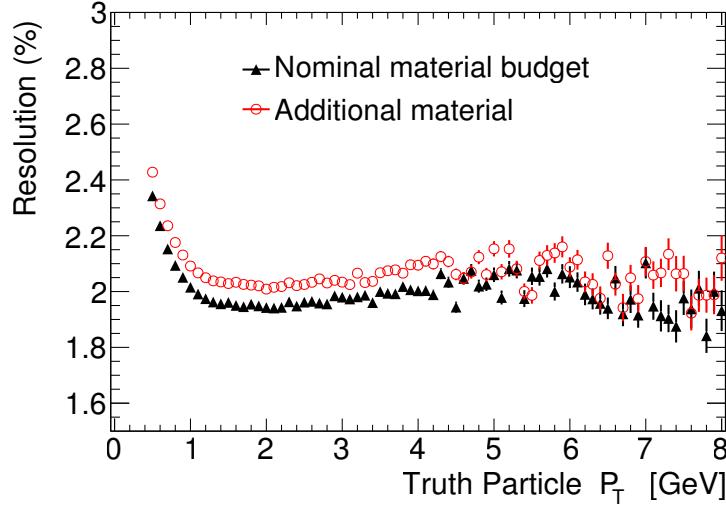


Figure 5.10: Track momentum resolution relative to the matched truth particle as a function of truth particle p_T . The results are shown for the two simulated samples of minimum bias processes, one with a nominal material budget and one with 10% additional inner detector material.

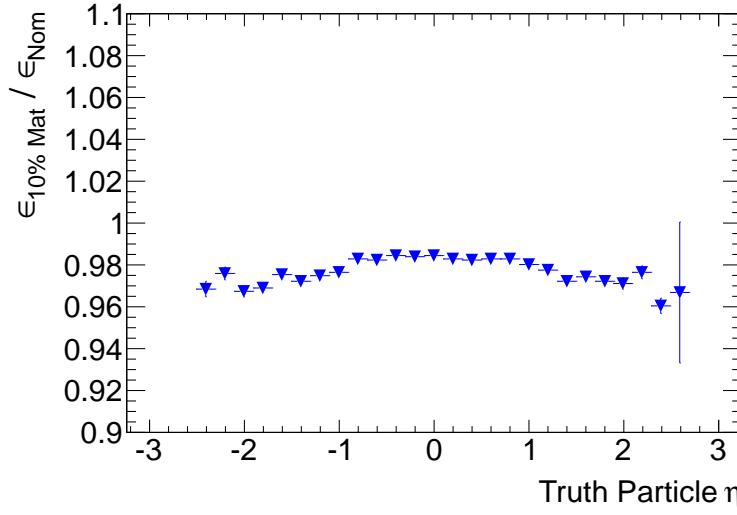


Figure 5.11: Ratio of track reconstruction efficiencies relative to primary truth particles as a function of truth particle η . The ratio is shown for efficiency measured in minimum bias samples with the nominal material budget (ϵ_{Nom}) and in the sample with 10% of additional material ($\epsilon_{10\%Mat}$).

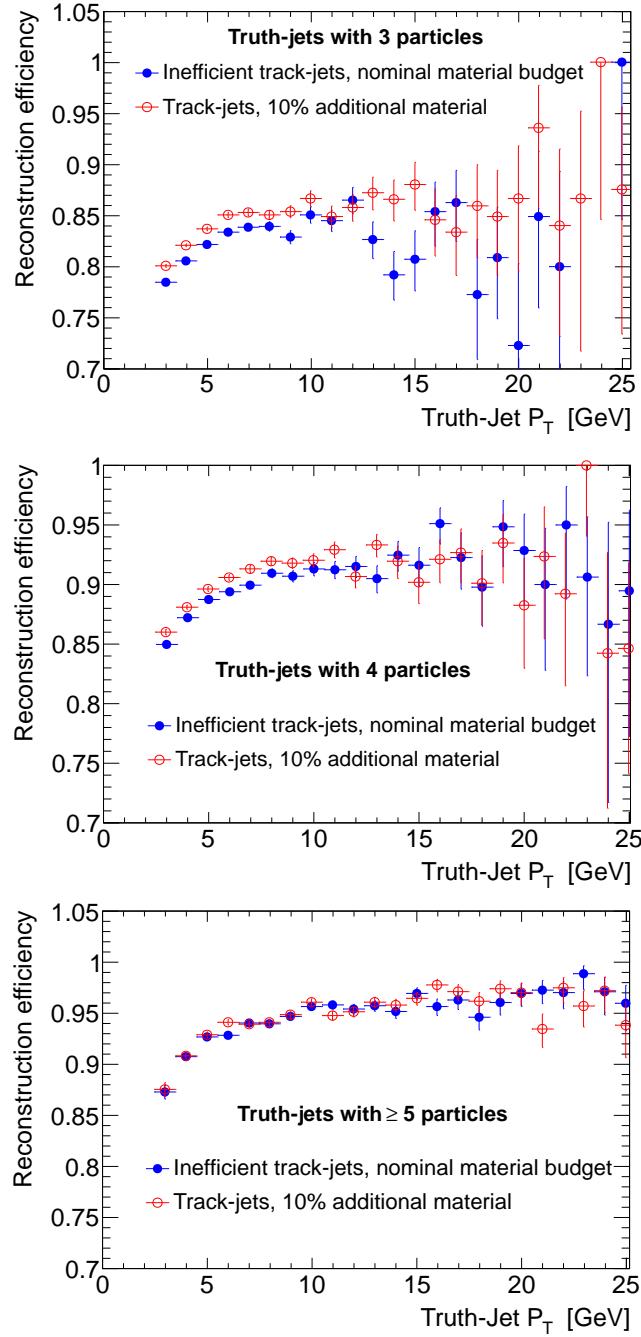


Figure 5.12: Jet reconstruction efficiency relative to truth-jet for inefficient track-jets reconstructed in a minimum bias sample with nominal material budget and for the nominal track-jet reconstruction in a sample with 10% additional material. Results are shown separately for truth-jet consisting of 3, 4 and ≥ 5 truth particle.

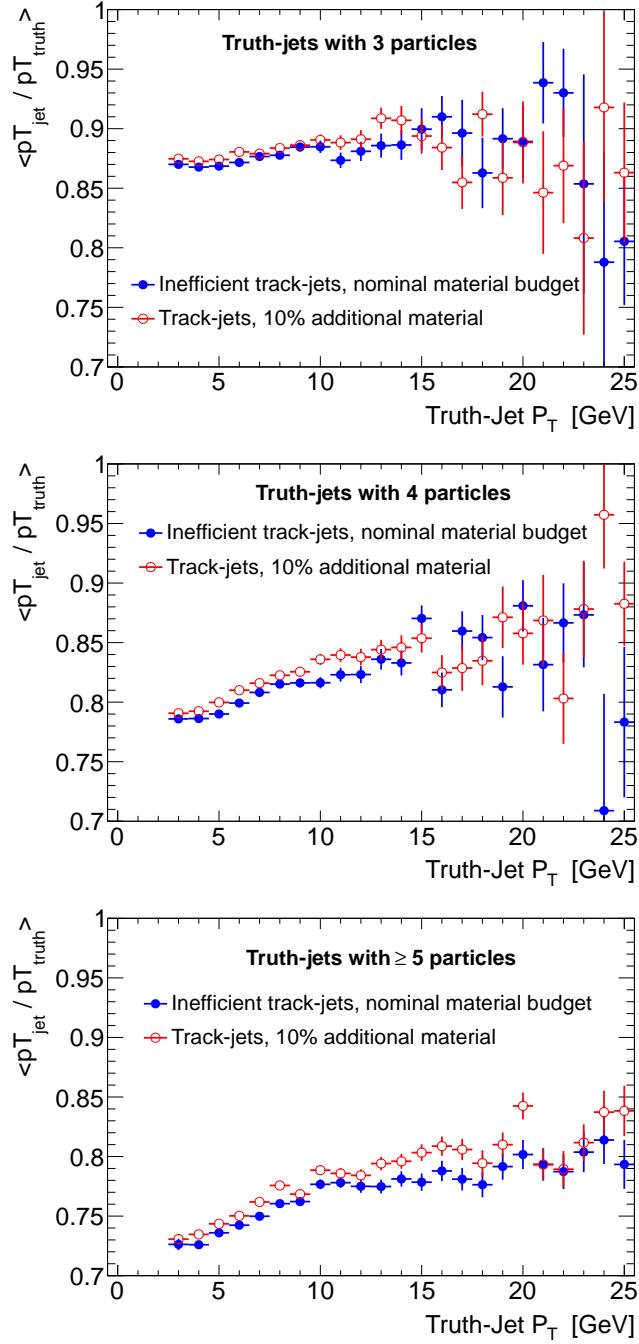


Figure 5.13: Fraction of the jet transverse momentum relative to truth-jet transverse momentum shown for inefficient track-jets reconstructed in a minimum bias sample with nominal material budget and for the nominal track-jet reconstruction in a sample with 10% additional material. Result are shown separately for truth-jet consisting of 3,4 and ≥ 5 truth particle.

Bibliography

- [1] G. Altarelli, “*Collider Physics within the Standard Model: a Primer,*” arXiv:1303.2842 .
- [2] S. P. Martin, “*A Supersymmetry primer,*” In *Kane, G.L. (ed.): Perspectives on supersymmetry II* 1-153 [hep-ph/9709356].
- [3] A. Djouadi, *The Anatomy of Electroweak Symmetry Breaking Tome II: The Higgs Bosons in the Minimal Supersymmetric Model*, Phys. Rep. 459 (2008) 1.
- [4] S. Heinemeyer *et al.* [LHC Higgs Cross Section Working Group Collaboration], “Handbook of LHC Higgs Cross Sections: 3. Higgs Properties,” arXiv:1307.1347
- [5] S. Dittmaier *et al.* [LHC Higgs Cross Section Working Group Collaboration], “*Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables,*” arXiv:1101.0593
- [6] Michael E. Peskin, Dan V. Schroeder, *An Introduction To Quantum Field Theory*, Westview Press, 1995.
- [7] S. L. Glashow, Partial Symmetries of Weak Interactions, Nuc. Phys. 22 (1961) 579.
- [8] A. Salam, Weak and Electromagnetic Interactions, Conf. Proc. C 680519 (1968) 367. Originally printed in Svartholm: Elementary Particle Theory, proceedings of the Nobel Symposium held 1968 at Lerum, Sweden.
- [9] S. Weinberg, A Model of Leptons, Phys. Rev. Lett. 19 (1967) 1264.

- [10] H. Fritsch, M. Gell-Mann, and H. Leutwyler, Advantages of the Color Octet Gluon Picture, *Phys. Lett. B* **47** (1973) 365.
- [11] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321.
- [12] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, *Phys. Lett.* **12** (1964) 132.
- [13] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508.
- [14] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, *Phys. Rev.* **145** (1966) 1156.
- [15] T. W. B. Kibble, Symmetry Breaking in NonAbelian Gauge Theories, *Phys. Rev.* **155** (1967) 1554.
- [16] The ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Physics Letters B* **716** (2012) 1–29.
- [17] The CMS Collatoration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Physics Letters B* **716** (2012) 30–61.
- [18] W. Hollik, “*Electroweak theory*,” hep-ph/9602380.
- [19] The Review of Particle Physics, J. Beringer et al. (Particle Data Group), *Phys. Rev. D86*, 010001 (2012) and 2013 partial update for the 2014 edition.
- [20] The ALEPH, CDF, D0, DELPHI, L3, OPAL and SLD Collaborations, the LEP Electroweak Working Group, Tevatron Working Group and SLD Electroweak and Heavy Flavour Working Groups, *Precision Electroweak Measurements and Constraints on the Standard Model*, arXiv:1012.2367. Updated for 2012 winter conferences, March 2012.

- [21] The Gfitter Group, M. Baak, et al., Updated Status of the Global Electroweak Fit And constraints on New Physics, *Eur. Phys. J. C* 72 (2012) 2003. Updated for 2012 winter conferences, March 2012, <http://gfitter.desy.de>.
- [22] G.W. Bennett et al., *Phys. Rev. Lett.* 89, 101804 (2002); Erratum *ibid. Phys. Rev. Lett.* 89, 129903 (2002); G.W. Bennett et al., *Phys. Rev. Lett.* 92, 161802 (2004); G.W. Bennett et al., *Phys. Rev. D*73, 072003 (2006).
- [23] B. Bhattacherjee, S. S. Biswal and D. Ghosh, *Top quark forward-backward asymmetry at Tevatron and its implications at the LHC*, *Phys. Rev. D* **83** (2011) 091501 [[arXiv:1102.0545 \[hep-ph\]](https://arxiv.org/abs/1102.0545)].
- [24] M. Veltman, *Acta. Phys. Pol. B*8 (1977) 475.
- [25] S. Weinberg, Gauge Hierarchies, *Phys. Lett. B* 82 (1979) 387.
- [26] M. Veltman, The InfraredUltraviolet Connection, *Acta Phys. Polon. B* 12 (1981) 437.
- [27] C. Smith and G. Ross, The Real Gauge Hierarchy Problem, *Phys. Lett. B* 105 (1981) 38.
- [28] F. Zwicky, Spectral Displacement of Extra Galactic Nebulae, *Helv. Phys. Acta* 6 (1933) 110.
- [29] E. W. Kolb and M. S. Turner, The Early Universe, *Front. Phys.* 69 (1990) 1.
- [30] The WMAP Collaboration, D. N. Spergel et al., First Year Wilkinson Microwave Anisotropy Probe (wmap) Observations: Determination of Cosmological Parameters, *Astrophys. J. Suppl.* 148 (2003) 175.
- [31] H. M. Georgi and S. L. Glashow, Unity of All ElementaryParticle Forces, *Phys. Rev. Lett.* 32 (1974) 438.
- [32] J. C. Pati and A. Salam, Lepton Number as the Fourth Color, *Phys. Rev. D* 10 (1974) 275.

- [33] P. Fayet, Phys. Lett. B 64, 159 (1976).
- [34] P. Fayet, Phys. Lett. B 69, 489 (1977), Phys. Lett. B 84, 416 (1979).
- [35] G.R. Farrar and P. Fayet, Phys. Lett. B 76, 575 (1978).
- [36] S. Martin, in Perspectives on Supersymmetry, Ed. G.L. Kane, World Scientific, Singapore, 1998, hep-ph/9709356.
- [37] For reviews on the MSSM, see: P. Fayet and S. Ferrara, Phys. Rep. 32 (1977) 249; H.P. Nilles, Phys. Rep. 110 (1984) 1; R. Barbieri, Riv. Nuovo Cim. 11N4 (1988) 1; R. Arnowitt and Pran Nath, Report CTP-TAMU-52-93; J. Bagger, Lectures at TASI-95, hep-ph/9604232.
- [38] H. E. Haber and G. Kane, Phys. Rep. 117 (1985) 75.
- [39] M. Drees and S. Martin, CLTP Report (1995) and hep-ph/9504324.
- [40] D.J.H. Chung, L.L. Everett, G.L. Kane, S.F. King, J. Lykken and L.T. Wang, Phys. Rept. 407 (2005) 1.
- [41] M. Drees, R.M. Godbole and P. Roy, Theory and Phenomenology of Sparticles, World Scientific, Spring 2004.
- [42] L. Girardello and M.T. Grisaru, Nucl. Phys. B194 (1982) 65.
- [43] Y. Okada, M. Yamaguchi and T. Yanagida, Prog. Theor. Phys. 85 (1991) 1; ibid. Phys. Lett. B262 (1991) 54; J.R. Ellis, G. Ridolfi and F. Zwirner, Phys. Lett. B257 (1991) 83; ibid. Phys. Lett. B262 (1991) 477; H.E. Haber and R. Hempfling, Phys. Rev. Lett. 66 (1991) 1815.
- [44] A. Djouadi and S. RosiersLees (conv.) et al., Summary Report of the MSSM Working Group for the GDRSupersymtrie, hep-ph/9901246.
- [45] K. Inoue, A. Komatsu and S. Takeshita, Prog. Theor. Phys 68 (1982) 927; (E) ibid. 70 (1983) 330.

- [46] E. Witten, Nucl. Phys. B188 (1981) 513; ibid Nucl. Phys. B202 (1982) 253; N. Sakai, Z. Phys. C11 (1981) 153; S. Dimopoulos and H. Georgi, Nucl. Phys. B193 (1981) 150; R.K. Kaul and P. Majumdar, Nucl. Phys. B199 (1982) 36.
- [47] J.F. Donoghue and L.F. Li, Phys. Rev. 19 (1979) 945.
- [48] J.F. Gunion and H.E. Haber, Nucl. Phys. B278 (1986) 449.
- [49] J.F. Gunion and H.E. Haber, Nucl. Phys. B272 (1986) 1; (E) hep-ph/9301205.
- [50] M. Frank et al., The Higgs Boson Masses and Mixings of the Complex MSSM in the FeynmanDiagrammatic Approach, JHEP 0702 (2007) 47.
- [51] M. Carena, S. Heinemeyer, C. E. M. Wagner, and G. Weiglein, *Suggestions for benchmark scenarios for MSSM Higgs boson searches at hadron colliders*, Eur. Phys. J. **C26** (2003) 601–607, hep-ph/0202167.
- [52] S. Heinemeyer, O. Stal, and G. Weiglein, *Interpreting the LHC Higgs Search Results in the MSSM*, Phys. Lett. B 710 (2012) 201206, arXiv:1112.3026.
- [53] The ATLAS Collaboration , *Constraints on New Phenomena via Higgs Boson Coupling Measurements with the ATLAS Detector*. ATLAS-CONF-2014-010.
- [54] L. Maiani, A. Polosa, and V. Riquer, Bounds to the Higgs Sector Masses in Minimal Supersymmetry from LHC Data, arXiv:1305.2172.
- [55] A. Djouadi, L. Maiani, G. Moreau, A. Polosa, J. Quevillon, et al., The post-Higgs MSSM scenario: Habemus MSSM?, arXiv:1307.5205.
- [56] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.
- [57] The ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.
- [58] J. Haffner, *The CERN accelerator complex*, OPEN-PHO-ACCEL-2013-056.
- [59] CMS Collaboration, CMS technical proposal, CERN-LHCC-94-38.

- [60] LHCb Collaboration, LHCb technical proposal, CERN-LHCC-98-004.
- [61] ALICE Collaboration, ALICE: Technical proposal for a large ion collider experiment at the CERN LHC, CERN-LHCC-95-71, CERN, 1995,
- [62] G. S. Guralnik, C.R. Hagen and T. W. B. Kibble Phys.Rev.Lett. **13** (1964) 585.
- [63] The ATLAS Collaboration, *Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV using the ATLAS Detector in 2011*, ATLAS-CONF-2011-116.
- [64] ALEPH, DELPHI, L3 and OPAL Collaboration, *Search for neutral MSSM Higgs bosons at LEP*, Eur. Phys. J. **C47** (2006) 547.
- [65] *Combined CDF and D0 upper limits on MSSM Higgs boson production in tau-tau final states with up to 2.2 fb^{-1} of data*, arXiv:1003.3363 [hep-ex].
- [66] The CMS Collaboration, S. Chatrchyan et al., arXiv:1104.1619 [hep-ex] [hep-ex].
- [67] The ATLAS Collaboration , *Search for the neutral Higgs bosons of the Minimal Supersymmetric Standard Model in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, arXiv:1211.6956 [hep-ex].
- [68] S. Heinemeyer, O. Stål and G. Weiglein, *Interpreting the LHC Higgs search results in the MSSM*, Phys.Lett. **B710** (2012) 201–206, arXiv:1112.3026 [hep-ph].
- [69] A. Arbey, M. Battaglia, A. Djouadi and F. Mahmoudi, *The Higgs sector of the phenomenological MSSM in the light of the Higgs boson discovery*, JHEP **1209** (2012) 107, arXiv:1207.1348 [hep-ph].
- [70] M. L. Mangano et al., *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003) 001.

- [71] J. Alwall et al., *Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions*, Eur. Phys. J. **C53** (2008) 473, [arXiv:0706.2569](https://arxiv.org/abs/0706.2569).
- [72] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **06** (2002) 029, [hep-ph/0204244](https://arxiv.org/abs/hep-ph/0204244).
- [73] B. P. Kersevan and E. Richter-Was, *The Monte Carlo Event Generator AcerMC 2.0 with Interfaces to PYTHIA 6.2 and HERWIG 6.5*, [arXiv:0405247v1 \[hep-ph\]](https://arxiv.org/abs/0405247v1).
- [74] G. Corcella et al., *HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **01** (2001) 010.
- [75] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, *Multiparton Interactions in Photoproduction at HERA*, Z. Phys. **C72** (1996) 637.
- [76] T. Binoth, M. Ciccolini, N. Kauer, and M. Kramer, *Gluon-induced W -boson pair production at the LHC*, JHEP **12** (2006) 046.
- [77] A. S. et al., *Higgs boson production in gluon fusion*, JHEP **02** (2009) 029.
- [78] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007.
- [79] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. M. Nadolsky and W. K. Tung, “New generation of parton distributions with uncertainties from global QCD analysis,” JHEP **0207** (2002) 012 [[hep-ph/0201195](https://arxiv.org/abs/hep-ph/0201195)].
- [80] H. -L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin and C. -P. Yuan, “New parton distributions for collider physics,” Phys. Rev. D **82** (2010) 074024 [[arXiv:1007.2241 \[hep-ph\]](https://arxiv.org/abs/1007.2241)].
- [81] The ATLAS Collaboration, *ATLAS Monte Carlo Tunes for MC09*, ATL-PHYS-PUB-2010-002.

- [82] S. Jadach, J. H. Kuhn and Z. Was, *TAUOLA - a library of Monte Carlo programs to simulate decays of polarized τ leptons*, Comput. Phys. Commun. **64** (1990) 275.
- [83] E. Barberio, B. V. Eijk and Z. Was, *Photos - a universal Monte Carlo for QED radiative corrections in decays*, Comput. Phys. Commun. **66** (1991) 115.
- [84] The GEANT4 Collaboration, S. Agostinelli et al., *GEANT4 - a simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.
- [85] The ATLAS Collaboration, G. Aad et al., *The ATLAS Simulation Infrastructure*, ATLAS-SOFT-2010-01-004, submitted to Eur. Phys. J. C., [arXiv:1005.4568](https://arxiv.org/abs/1005.4568).
- [86] The ATLAS Collaboration, *Estimation of $Z/\gamma^* \rightarrow \tau\tau$ Background in VBF $H \rightarrow \tau\tau$ Searches from $Z \rightarrow \mu\mu$ Data using an Embedding Technique*, ATL-PHYS-INT-2009-109.
- [87] The ATLAS Collaboration, *Search for the Standard Model Higgs boson in the $H \rightarrow \tau\tau$ decay mode with 4.7 $f\ b1$ of ATLAS detector*, Tech. Rep. ATLAS-CONF-2012-014, CERN, Geneva, Mar, 2012.
- [88] The ATLAS Collaboration, *Search for the Standard Model Higgs boson $H \rightarrow \tau\tau$ decays with the ATLAS detector*, ATL-COM-PHYS-2013-722.
- [89] T. S. et al., *Z physics at LEP 1*, CERN 89-08 **3** (1989) 143.
- [90] The ATLAS Collaboration, Inner Detector: Technical Design Report, CERN/LHCC/97-016/017 (1997).
- [91] The ATLAS Collaboration, G. Aad et al., The ATLAS Experiment at the CERN Large Hadron Collider, 2008 JINST **3** S08003.
- [92] A. Bazan, T. Bouedo, P. Ghez, M. Marino and C. Tull, “The Athena data dictionary and description language,” eConf C **0303241** (2003) MOJT010 [[cs/0305049 \[cs-se\]](https://arxiv.org/abs/cs/0305049)].

- [93] The ATLAS Collaboration, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, CERN-OPEN-2008-020, arXiv:0901.0512.
- [94] T. Cornelissen et al., Concepts, Design and Implementation of the ATLAS New Tracking, ATLAS Note ATL-SOFT-PUB-2007-007 (2007).
- [95] Kalman, R. E. (1960). "A New Approach to Linear Filtering and Prediction Problems". Journal of Basic Engineering 82 (1): 3545. doi:10.1115/1.3662552
- [96] The ATLAS Collaboration, Performance of primary vertex reconstruction in proton-proton collisions at $s = \text{sqrt}(7)$ TeV in the ATLAS experiment. ATLAS-CONF-2010-069.
- [97] R. Fruhwirth, W. Waltenberger, P. Vanlaer, *Adaptive vertex fitting*, J. Phys. G34 (2007).
- [98] The ATLAS Collaboration, *Characterization of Interaction-Point Beam Parameters Using the pp Event-Vertex Distribution Reconstructed in the ATLAS Detector at the LHC*, ATL-CONF-2010-027.
- [99] The ATLAS collaboration, *Expected electron performance in the ATLAS experiment*, ATL-PHYS-PUB-2011-006
- [100] A. Hocker, P. Speckmayer, J. Stelzer, F. Tegenfeldt, and H. Voss, *TMVA, toolkit for multivariate data analysis with ROOT*, CERN-2008-001 .
- [101] The ATLAS Collaboration, *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data*, CERN-PH-EP-2014-040, arXiv:1404.2240
- [102] The ATLAS Collaboration, G. Aad et al., *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, Eur.Phys.J. C72 (2012) 1909.
- [103] S. Hassini, et al., *A muon identification and combined reconstruction procedure for the ATLAS detector at the LHC using the (MUONBOY, STACO, MuTag) reconstruction packages*, NIM A572 (2007) 7779.

- [104] The ATLAS Collaboration, G. Aad et al., *Preliminary results on the muon reconstruction efficiency, momentum resolution, and momentum scale in ATLAS 2012 pp collision data*, ATLAS-CONF-2013-088, CERN, 2013,
- [105] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet user manual*, Eur.Phys.J. C72 (2012) 1896.
- [106] W. Lampl et al., *Calorimeter Clustering Algorithms : Description and Performance*, ATL-LARG-PUB-2008-002.
- [107] M. Cacciari, G. P. Salam, and G. Soyez, *The anti- kt jet clustering algorithm*, JHEP 04 (2008) 63.
- [108] E. Abat, J. Abdallah, T. Addy, P. Adragna, et al., *Combined performance studies for electrons at the 2004 ATLAS combined test-beam*, JINST 5 (2010) P11006.
- [109] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV*, Submitted to EPJ (2011) , arXiv:1112.6426
- [110] The ATLAS Collaboration, *Pile-up corrections for jets from proton-proton collisions at ATLAS in 2011*, ATLAS-CONF-2012-064, July, 2012.
- [111] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, Phys.Lett. B659 (2008) 119.
- [112] The ATLAS Collaboration, G. Aad et al., *Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector*, Eur.Phys.J. C73 (2013) 2306
- [113] The ATLAS collaboration, *Jet energy scale and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with ATLAS 2011 data*, ATLAS-CONF-2013-004

- [114] The ATLAS Collaboration, *Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of $\sqrt{s} = 7$ TeV*, ATLAS-CONF-2010-038.
- [115] G. Piacquadio, C. Weiser, *A new inclusive secondary vertex algorithm for b-jet tagging in ATLAS*, JPCS 119 (2008) 032032
- [116] The ATLAS Collaboration, G. Aad et al., *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*, ATLAS-CONF-2011-102, CERN, 2011, ATLAS-CONF-2011-102.
- [117] The ATLAS Collaboration, *Measuring the b-tag efficiency in a ttbar sample with 4.7 fb^{-1} of data from the ATLAS detector* ATLAS-CONF-2012-097.
- [118] The ATLAS Collaboration, *Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment* ATLAS-CONF-2014-004.
- [119] The ATLAS Collaboration, *Reconstruction and Calibration of Missing Transverse Energy and Performance in Z and W events in ATLAS Proton-Proton Collisions at $\sqrt{s}=7$ TeV*, ATLAS-CONF-2011-080.
- [120] ATLAS Collaboration, G. Aad et al., Performance of missing transverse momentum reconstruction in proton-proton collisions at 7 TeV with ATLAS, Eur.Phys.J. C72 (2012) 1844.
- [121] TheATLAS Collaboration, *Performance of the Reconstruction and Identification of Hadronic tau Decays in ATLAS with 2011 Data*, ATLAS-CONF-2012-142.
- [122] The ATLAS Collaboration, G. Aad et al., *Performance of the ATLAS trigger system in 2010*, Eur.Phys.J. C72 (2012) 1849.
- [123] The ATLAS Collaboration, G. Aad et al., *Performance of the ATLAS muon trigger in 2011*, ATLAS-CONF-2012-099, CERN, 2012.

- [124] The ATLAS Collaboration, G. Aad et al., *Performance of the ATLAS electron and photon trigger in p-p collisions at $\sqrt{s} = 7$ TeV in 2011*, ATLAS-CONF-2012-048, CERN, 2012.
- [125] M. Dobbs and J.B. Hansen, *The HepMC C++ Monte Carlo Event Record for High Energy Physics*, Computer Physics Communications, ATL-SOFT-2000-001.
- [126] The ATLAS Collaboration, *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, Phys. Lett. B 726 (2013), pp. 120-144.
- [127] The ATLAS Collaboration, *Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC*, Phys. Lett. B 726 (2013), pp. 88-119.
- [128] The CMS Collaboration, *Evidence for the direct decay of the 125 GeV Higgs boson to fermions*, arXiv:1401.6527 [hep-ex].
- [129] The CMS Collaboration, *Higgs boson width from on- vs. off-shell production and decay to Z-boson pairs*, arXiv:1405.3455.
- [130] A. Elagin, P. Murat, A. Pranko, and A. Safonov, *A New Mass Reconstruction Technique for Resonances Decaying to di-tau*, arXiv:1012.4686 [hep-ex]. * Temporary entry *.
- [131] T. A. Collaboration, *Search for neutral MSSM Higgs bosons decaying to $\tau\tau$ pairs in proton-proton collisions at with the ATLAS detector*, Physics Letters B 705 (2011) no. 3, 174 – 192.
- [132] The ATLAS Collaboration, *Data-driven estimation of the background to charged Higgs boson searches using hadronically-decaying tau final states in ATLAS*, ATLAS-CONF-2011-051.
- [133] The ATLAS Collaboration, *Measurement of the $Z \rightarrow \tau\tau$ cross section with the ATLAS detector*, Phys. Rev. D 84 (2011) 112006.

- [134] T. A. Collaboration, *Search for the neutral Higgs bosons of the Minimal Supersymmetric Standard Model in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP , [arXiv:1211.6956](https://arxiv.org/abs/1211.6956).
- [135] Atlas statistics forum, *ABCD method in searches*, link
- [136] The ATLAS Collaboration, *Search for Neutral MSSM Higgs Bosons H to $\tau\tau$ to $l\tau_h$ with the ATLAS Detector in 7 TeV Collisions*, ATL-COM-PHYS-2012-094.
- [137] The ATLAS Collaboration, *Search for neutral Higgs Bosons in the decay mode $H \rightarrow \tau\tau \rightarrow ll + 4\nu$ in proton proton collision at $\sqrt{7}$ TeV with the ATLAS Detector*, ATL-COM-PHYS-2011-758.
- [138] The Atlas Collaboration, *Measurement of the $t\bar{t}$ production cross-section in pp collisions at $\sqrt{s} = 8$ TeV using e-mu events with b-tagged jets* . ATLAS-CONF-2013-097.
- [139] T. Sjostrand, S. Mrenna and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026.
- [140] A. B. et al., *Rivet user manual*, [arXiv:1003.0694 \[hep-ph\]](https://arxiv.org/abs/1003.0694) .
- [141] The ATLAS and CMS Collaborations, *Procedure for the LHC Higgs boson search combination in summer 2011*, ATL-PHYS-PUB-2011-011, ATL-COM-PHYS-2011-818, CMS-NOTE-2011-005.
- [142] W. Verkerke and D. P. Kirkby, *The RooFit Toolkit for Data Modeling*, eConf C0303241 (2003) MOLT007, arXiv:physics/0306116
- [143] K. S. Cranmer et al., *The Roostats Project*, PoS ACAT2010 (2010) 57.
- [144] K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke, *Histfactory: A Tool for Creating Statistical Models for use with RooFit and Roostats*, CERNOPEN2012016 (2012)

- [145] E. G. G. Cowan, K. Cranmer and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, arXiv:1007.1727 [hep-ex].
- [146] A. L. Read. *Presentation of search results: the CLs technique*. J. Phys. G: Nucl. Part. Phys., 28, 2002.
- [147] A. L. Read. *Modified frequentist analysis of search results (the CLs method)*. In Proceedings of the First Workshop on Confidence Limits, CERN, Geneva, Switzerland, 2000.
- [148] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, R. Tanaka (Eds.), et al., *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*, arXiv:1101.0593 [hep-ph].
- [149] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.), *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*, CERN-2012-002 (CERN, Geneva, 2012) , arXiv:1201.3084 [hep-ph].
- [150] ATLAS collaboration *Performance of the ATLAS Silicon Pattern Recognition Algorithm in Data and Simulation at $\sqrt{s} = 7 \text{ TeV}$* , ATLAS-CONF-2010-072
- [151] The ATLAS Collaboration, *A measurement of the material in the ATLAS inner detector using secondary hadronic interactions*, arXiv:1110.6191, JINST 7 (2012) P01013
- [152] The ATLAS Collaboration, *Validation of the ATLAS jet energy scale uncertainties using tracks in proton-proton collision $\sqrt{s} = 7 \text{ TeV}$* , ATLAS-CONF-2011-067
- [153] The ATLAS Collaboration, *Track Reconstruction Efficiency in $\sqrt{s} = 7 \text{ TeV}$ Data for Tracks with $p_T > 100 \text{ MeV}$* , ATL-PHYS-INT-2010-112
- [154] D. de Florian, G. Ferrera, M. Grazzini and D. Tommasini, *Transverse-momentum resummation: Higgs boson production at the Tevatron and the LHC*, JHEP 1111 (2011) , arXiv:1109.2109 [hep-ph].

- [155] Statistical twiki, NuissanceCheck. <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/Nuisance>

Appendix A

QCD Trigger bias

The single electron trigger (`EF_e24vhi_medium1`) used in this analysis includes the following isolation cut: $p_T(\text{cone})20/p_T < 0.1$. This means that the kinematical distributions in the anti-isolated ABCD regions will be biased due to a reduced efficiency for high p_T electrons. This unwanted feature may potentially effect the R_{QCD} factor, as the OS/SS ratio may differ due to different p_T spectrum. To check the effect on R_{QCD} the ABCD method has been repeated using the `EF_e24vh_medium1` trigger, which doesn't include isolation and hence is prescaled in 2012 8 TeV data. The prescale of a factor around 100 has been taken in consideration using trigger information stored in D3PD. Figure A.1 shows $p_T(\text{cone})$ distribution for the standard and test triggers. The comparable event yields in the region $p_T(\text{cone})20/p_T < 0.1$ show that the prescale normalisation for the test trigger has been correctly accounted for.

Figure A.2 shows the behaviour of R_{QCD} factor as a function of isolation variable for the two triggers under test. As the deviations are within statistical uncertainty, we conclude that the isolation requirement used at trigger level does not influence the OS/SS ratio. Hence no further systematic uncertainty is assigned.

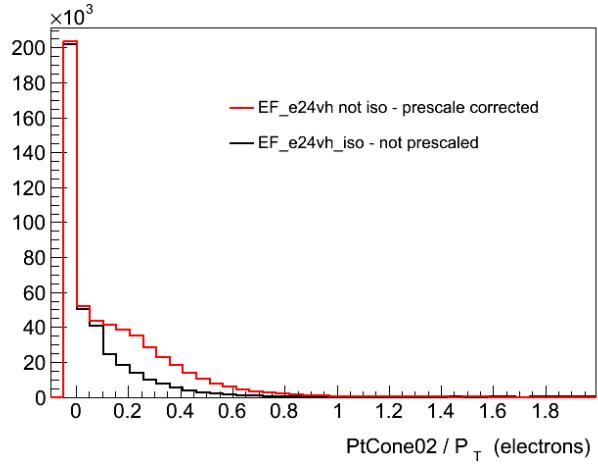


Figure A.1: $p_T(\text{cone}) / p_T$ distribution for the analysis standard trigger and its corrispective without isolation requirement, this second trigger is rescaled according to prescales information.

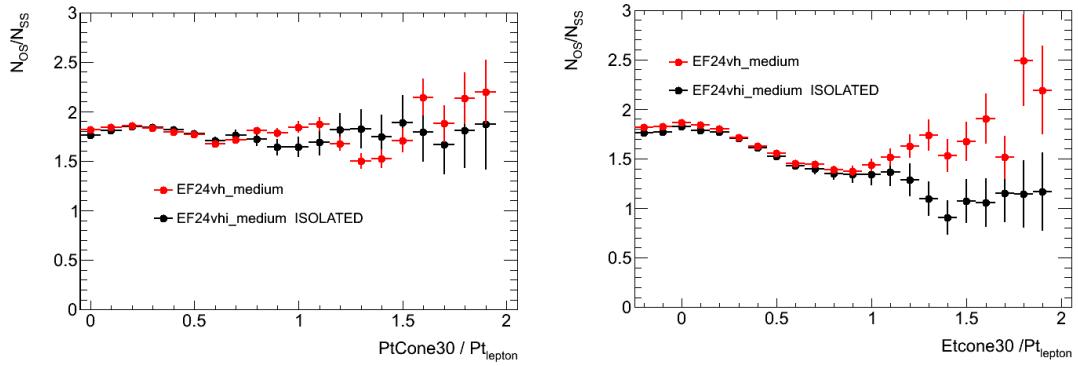


Figure A.2: R_{QCD} as a function of (a) $p_T(\text{cone}) / p_T$ and (b) $E_T(\text{cone}) / \text{PT}$ for the electron triggers with and without isolation requirement.

Appendix B

QCD Additional Plots

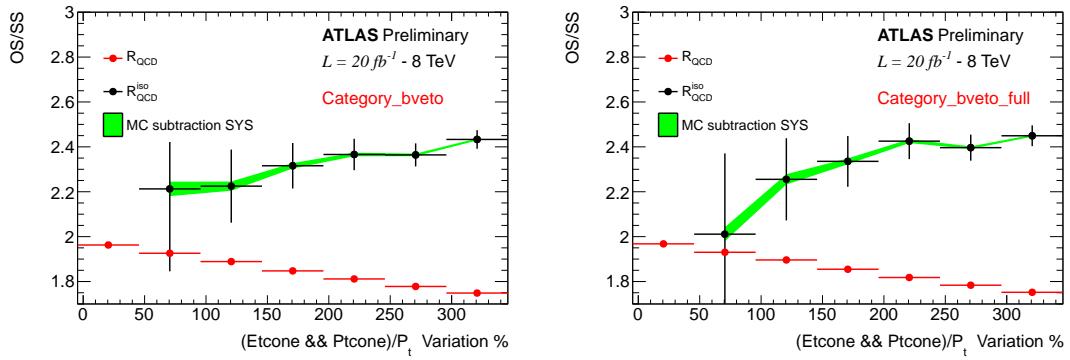


Figure B.1: OS/SS ratio as a function of lepton isolation variable selections (a) after the requirement of zero b-jets and (b) for the full b-veto selection. The isolation selections are varied as a percentage relative to the standard lepton isolation cut values (0 in the plot). The red points show the anti-isolated scale factor R_{QCD} , i.e. the ratio between regions C and D. The black points show the isolated SF, which is defined as the ratio between region \hat{A} and \hat{B} , where the leptons have isolation values larger than the nominal value but smaller than the sliding cut.

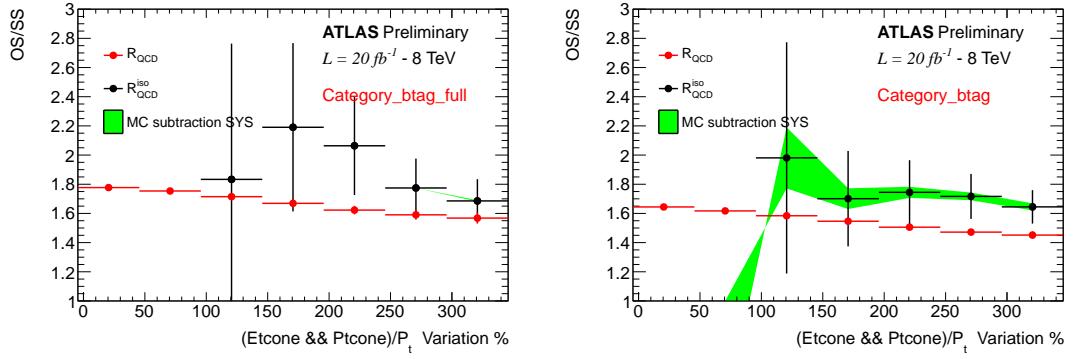


Figure B.2: OS/SS ratio as a function of lepton isolation variable selections (a) after the requirement of one b-jet and (b) for the full b-tag selection. The isolation selections are varied as a percentage relative to the standard lepton isolation cut values (0 in the plot). The red points show the anti-isolated scale factor R_{QCD} , i.e. the ratio between regions C and D. The black points show the isolated SF, which is defined as the ratio between region \hat{A} and \hat{B} , where the leptons have isolation values larger than the nominal value but smaller than the sliding cut.

Appendix C

Limits cross checks and additional plots

C.1 The ABCD Method

The actual implementation in the limit framework of the ABCD method follows that suggested in [135]. The control data samples B,C and D are considered as additional channels to be statistically combined to two the signal event category. Three free parameters are fitted in B,C and D channels which are: the number of multi-jet events in the data sample B, N_B^{QCD} , the factor R_{QCD} and the factor that extrapolates from isolated to anti-isolated data control samples R_{BD} . Neglecting signal contributions, the following equations can be written for the event yield of the B,C and D control data samples:

$$N_B = N_B^{BKG} + N_B^{QCD}$$

$$N_C = N_C^{BKG} + N_B^{QCD} \times R_{QCD} \times R_{BD}$$

$$N_D = N_D^{BKG} + N_B^{QCD} \times R_{BD}$$

where N^{BKG} represent the prediction of non-QCD background in the relative data samples. The estimate of multi-jet event yield in the signal sample will be then $N_B^{QCD} \times R_{QCD}$. This method is particularly powerful because in the best fit of

R_{QCD} the statistical and systematics uncertainty for non-QCD backgrounds and data are considered.

C.2 Additional Limit Checks

During the limit derivation, the systematic uncertainties (translated in term of nuisance parameter) are fitted to the data, several checks have been performed to ensure the quality of our statistical model. If some of the nuisance parameters are significantly different from their nominal value (ie before fit), it can be symptomatic of an important mis-modelling and must be carefully scrutinised. Also the correlation between the nuisance parameter and the signal strength (which reflects the degeneracy of the fit) is an important element to keep under control, in fact it reflects how well the data can constraint the nuisance parameters. Finally, to have a feeling of the behaviour of the likelihood at its minimum one can check the negative log likelihood profile in each nuisance parameter direction. We performed all this checks using the package NuisanceCheck-00-00-05 described in [155].

The signal and background model with the signal normalisation free (unconditional fit) is fitted to the data, in the following example plots the signal is assumed for the mass point $m_A = 120$ GeV, $\tan\beta = 20$, The difference between the post fit and pre-fit value of the nuisance parameter along with their errors is shown in figure C.1-C.3, respectively for the b-veto, b-tag channel and the combination between them. Figure C.4-C.6 shows the correlation matrix between the nuisance parameters respectively for the veto category, tag category and the combination between the two channels. Figure C.7-C.9 shows the behaviour of the likelihood at its minimum for each of the nuisance parameters (while a nuisance parameter is investigated the other are kept constant) for the combination between the channels.

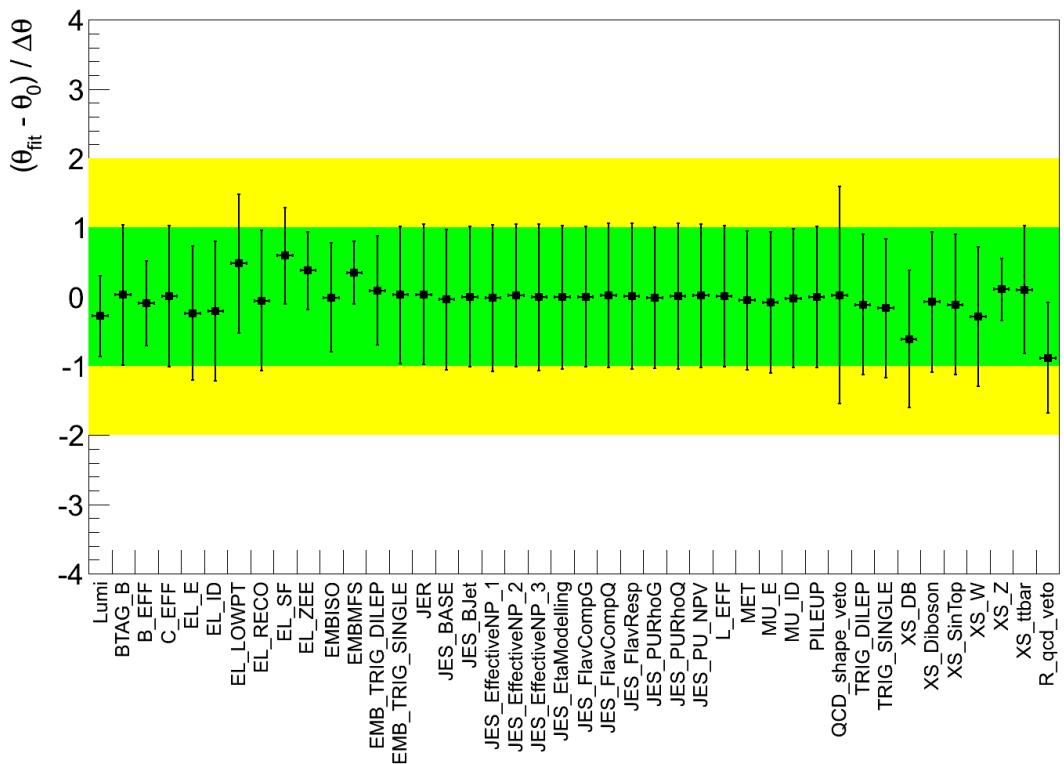


Figure C.1: Pulls for nuisance parameter considered in the fit, mA = 120 GeV, $\tan\beta = 20$, for the veto channel.

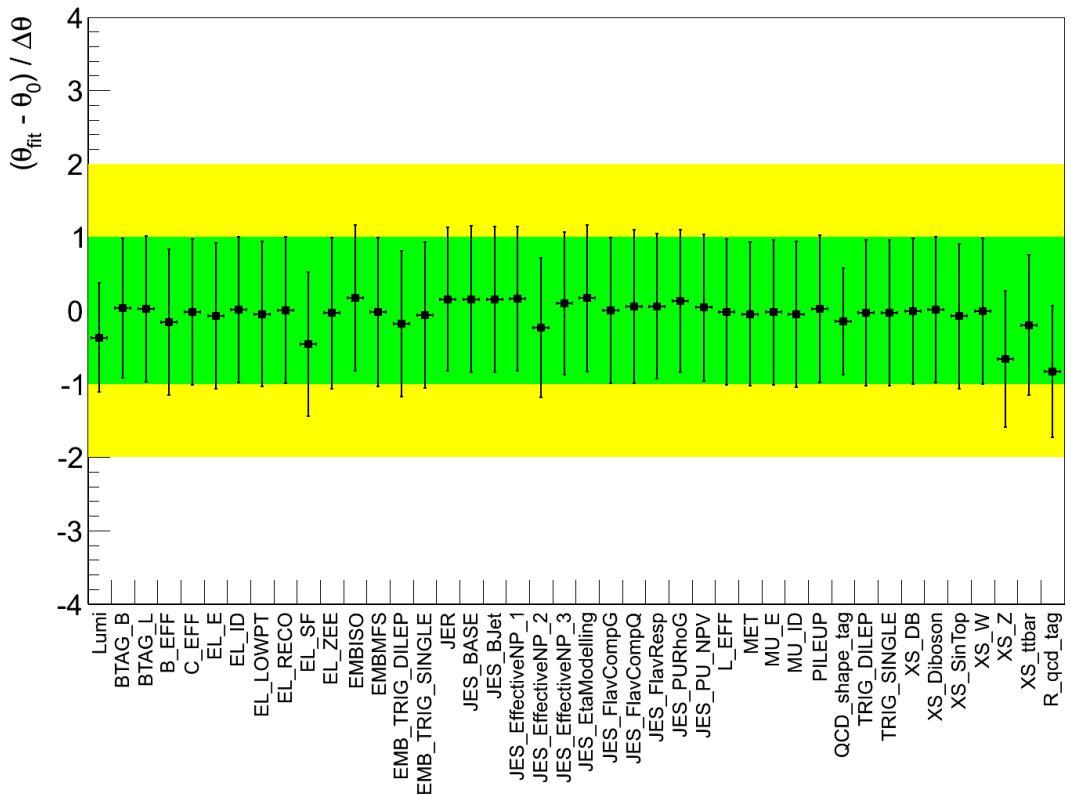


Figure C.2: Pulls for nuisance parameter considered in the fit, mA = 120 GeV, $\tan\beta = 20$, for the tag channel.

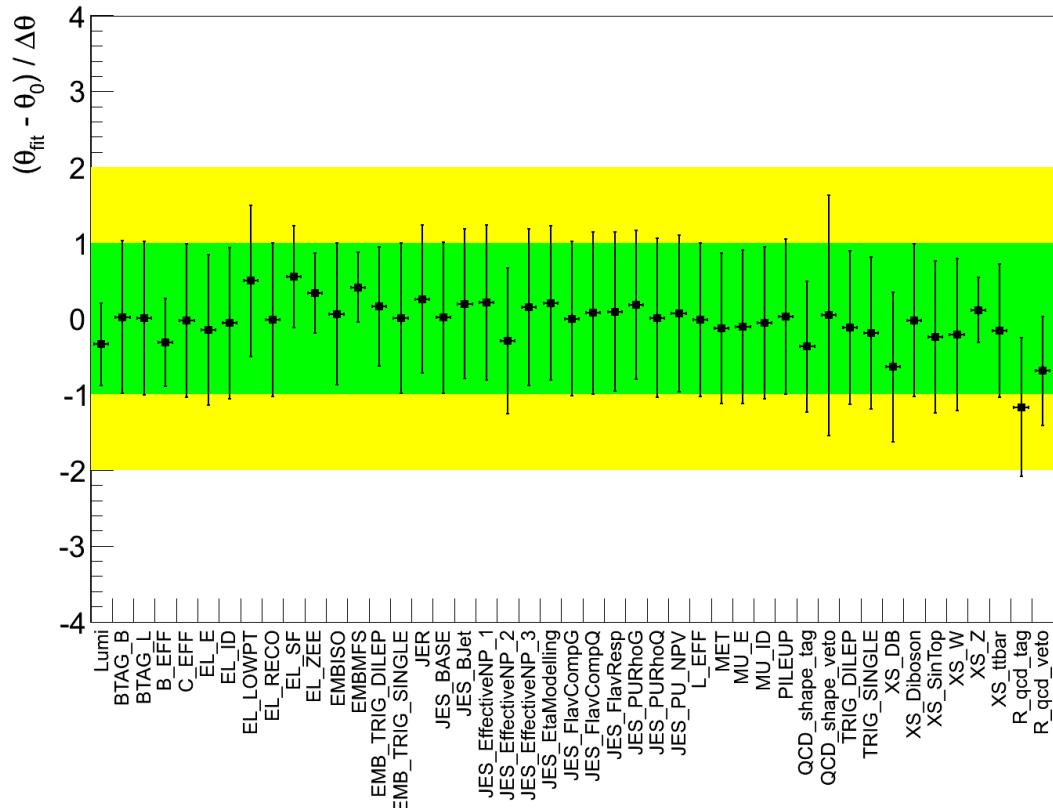


Figure C.3: Pulls for nuisance parameter considered in the fit, mA = 120 GeV, $\tan\beta = 20$, combination between the two channel.

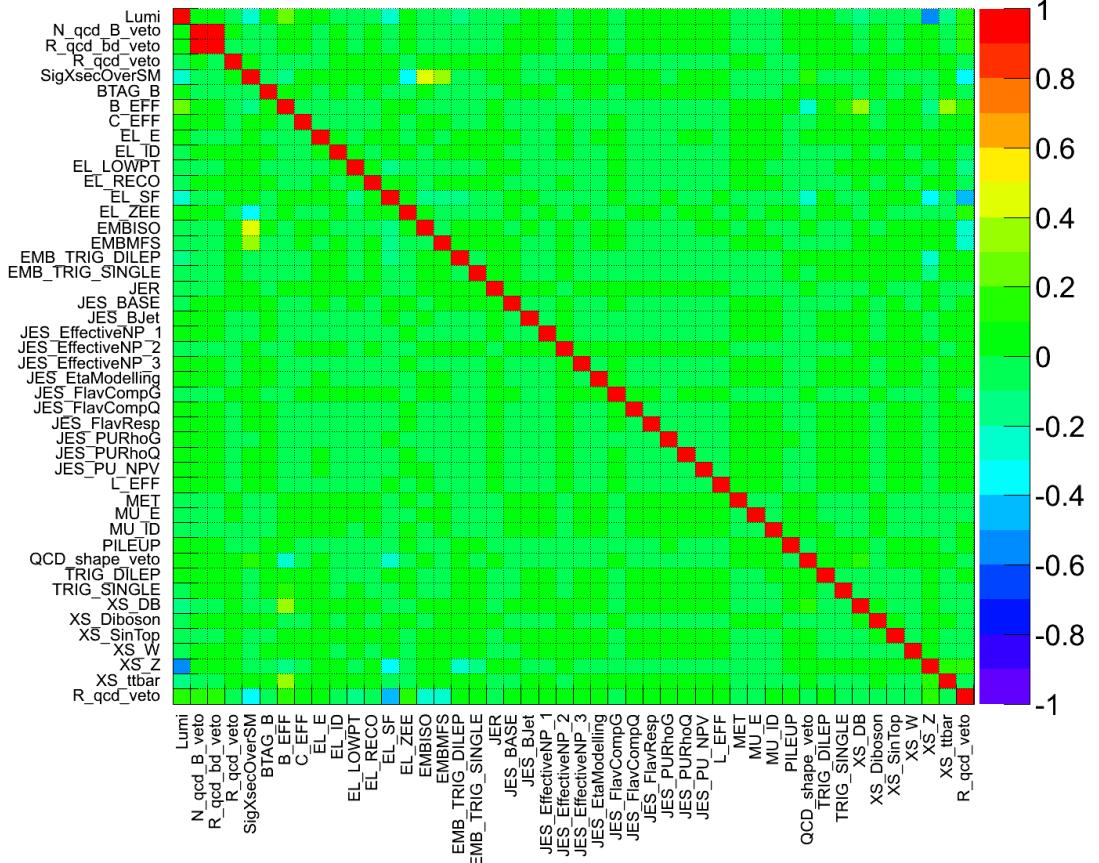


Figure C.4: Correlation matrix for nuisance parameters considered in the fit. The point $m_A = 120$ GeV and $\tan\beta = 20$ is considered for the tag channel.

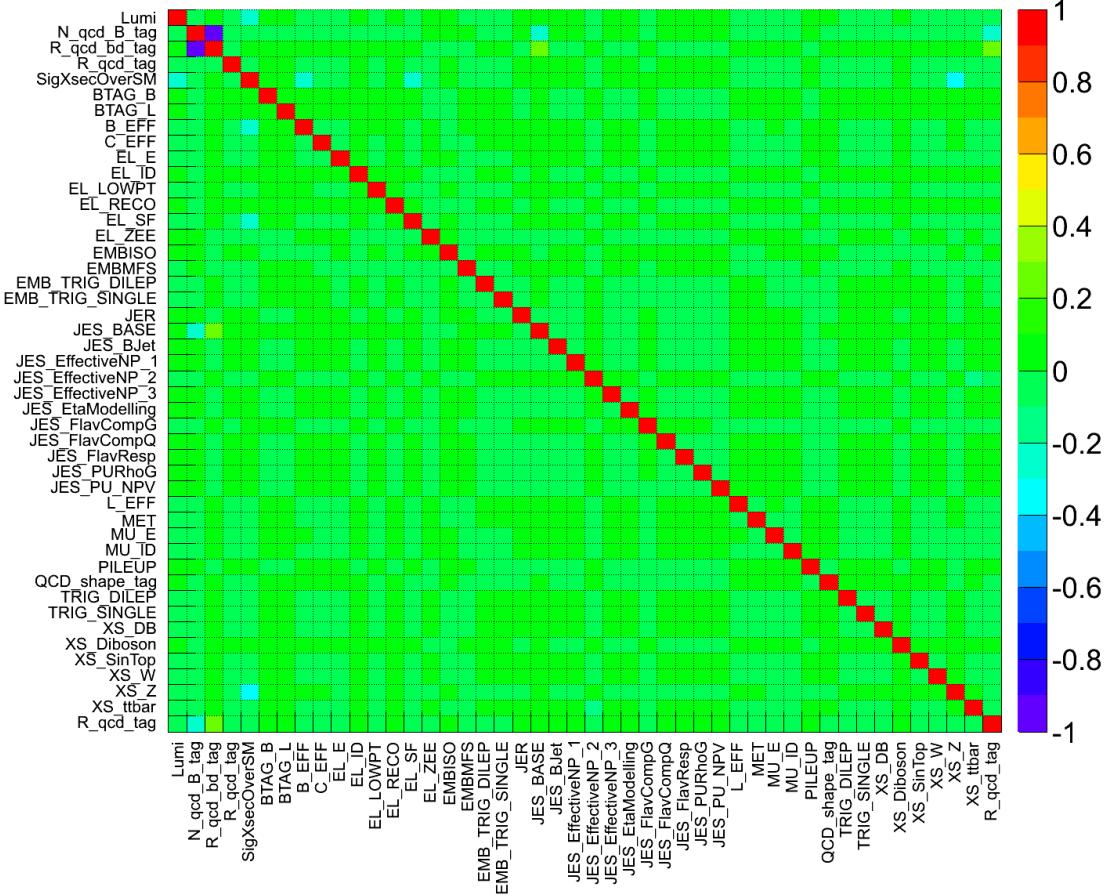


Figure C.5: Correlation matrix for nuisance parameters considered in the fit. The point $m_A = 120$ GeV and $\tan\beta = 20$ is considered for the tag channel.

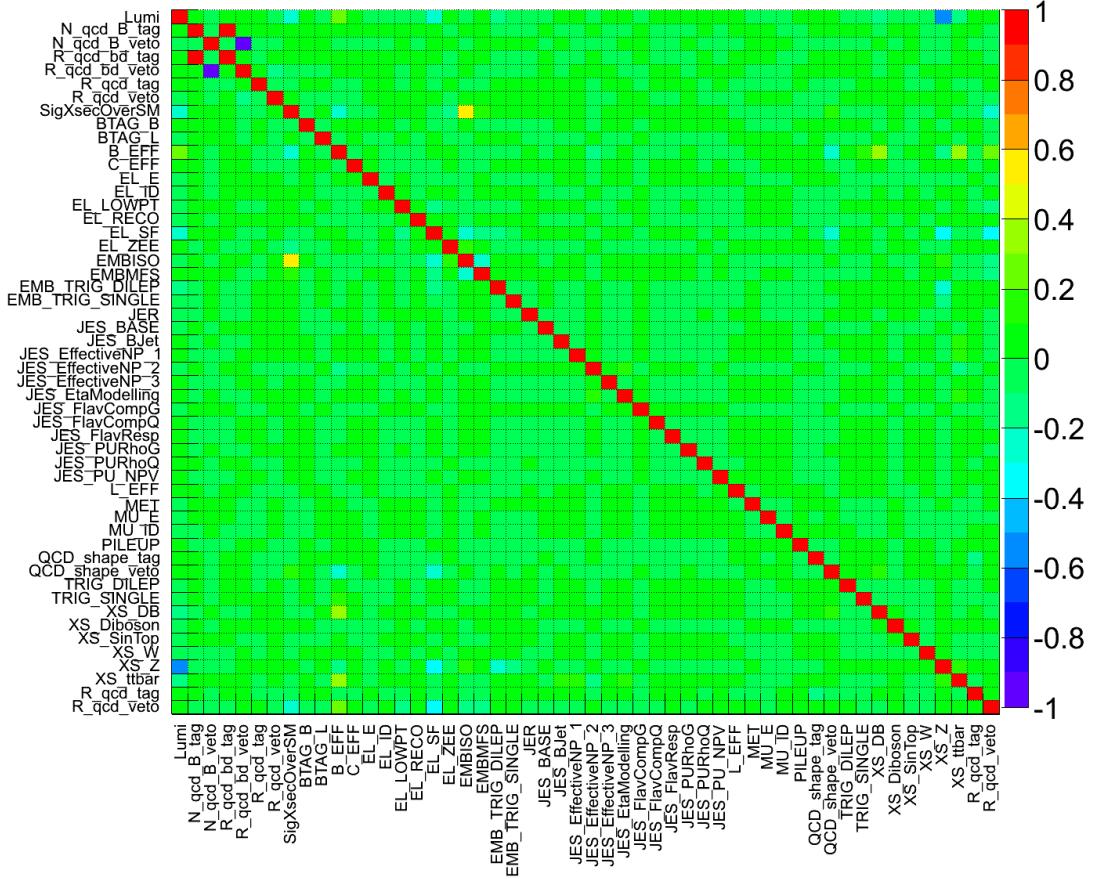


Figure C.6: Correlation matrix for nuisance parameters considered in the fit. The point $m_A = 120$ GeV and $\tan\beta = 20$ is considered for the combination of the b-tag and b-veto channels.

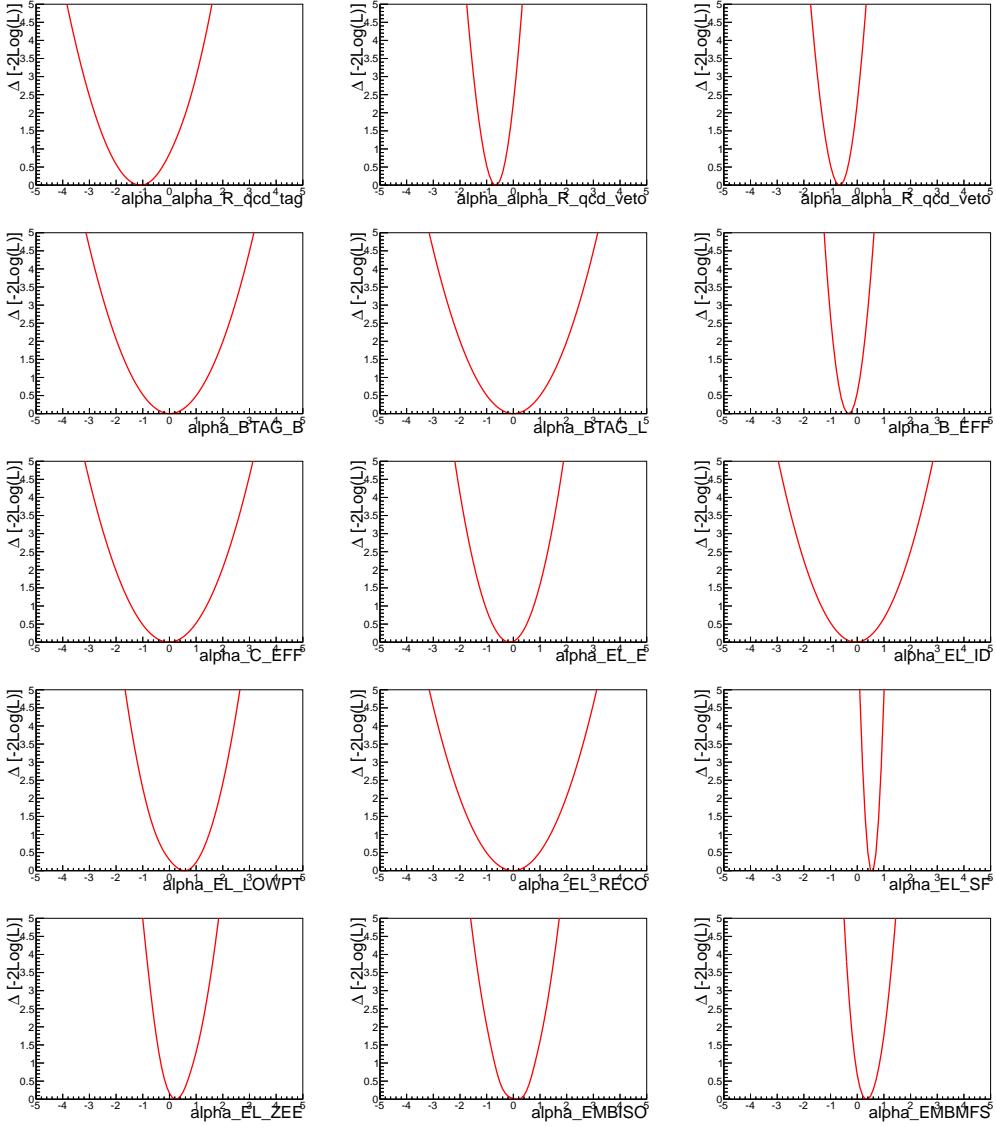


Figure C.7: Likelihood scans for nuisance parameter considered in the fit, mA = 120 GeV, $\tan\beta = 20$, combination between the two channel.

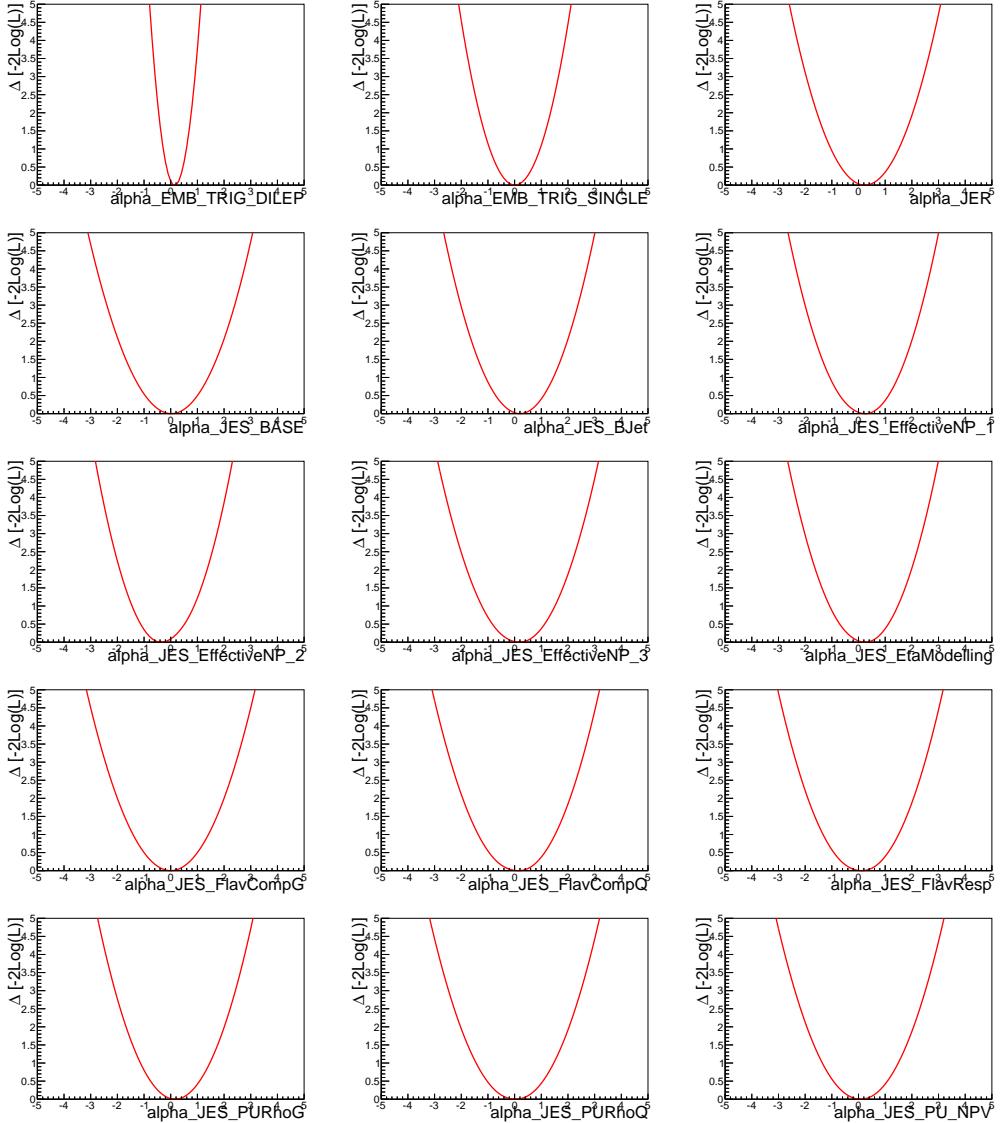


Figure C.8: Likelihood scans for nuisance parameter considered in the fit, mA = 120 GeV, $\tan\beta = 20$, combination between the two channel.

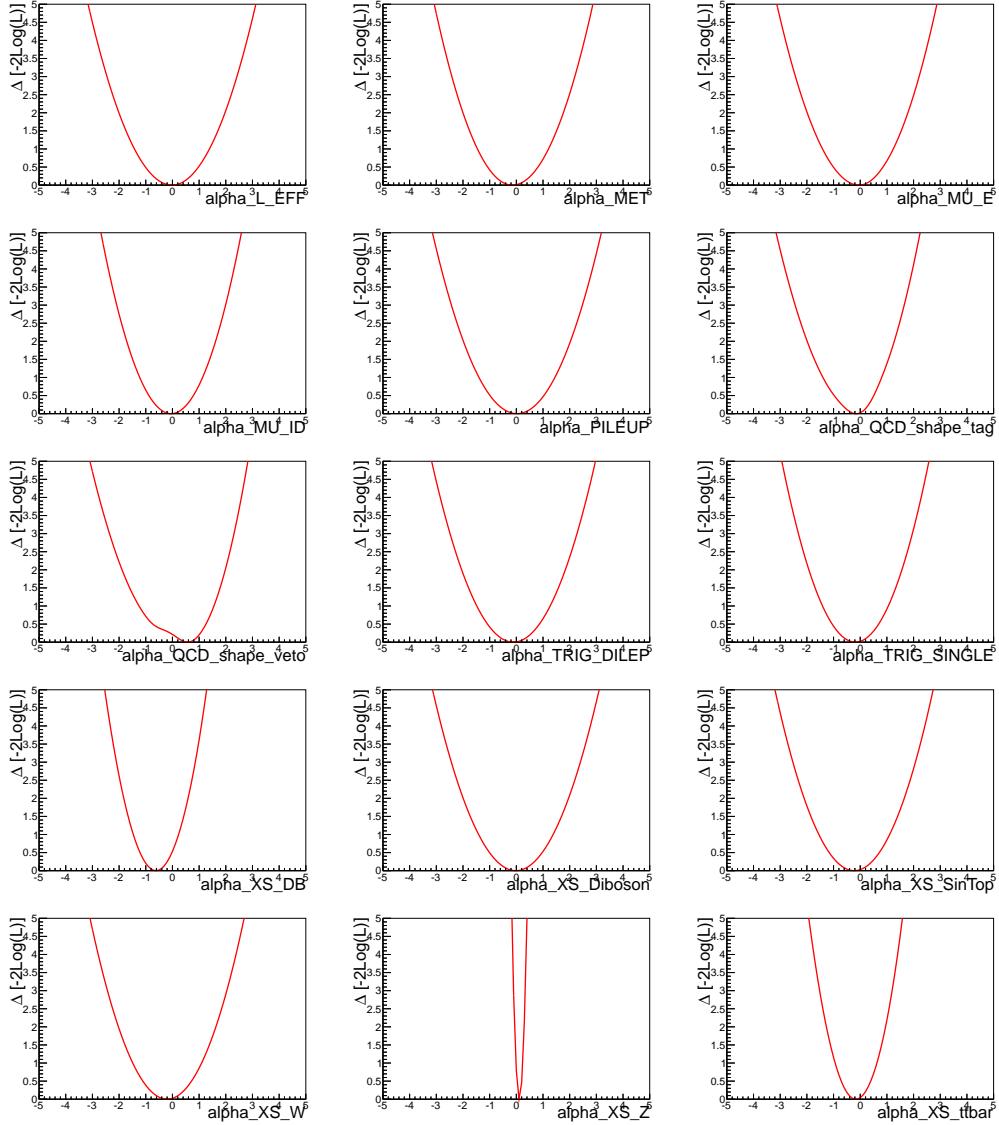


Figure C.9: Likelihood scans for nuisance parameter considered in the fit, mA = 120 GeV, $\tan\beta = 20$, combination between the two channel.

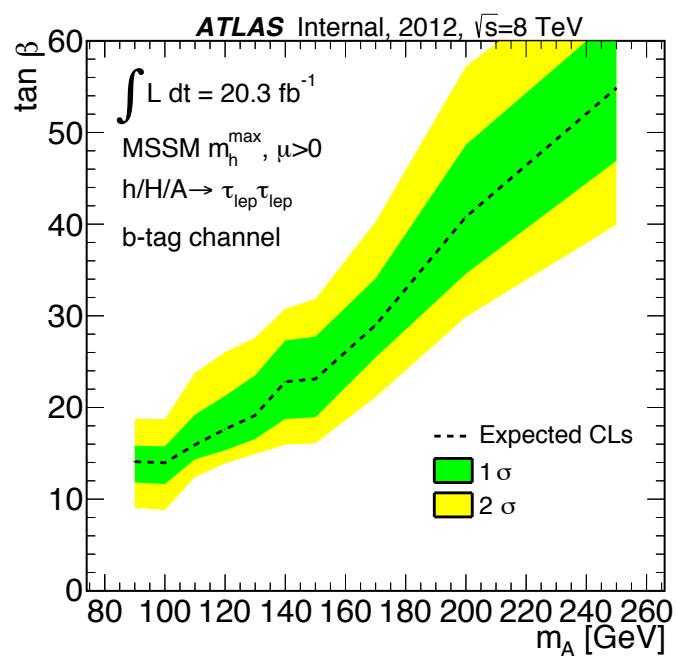


Figure C.10: Expected exclusion limits, using the b-tag channel, for MSSM Higgs boson production in the MSSM m_A vs $\tan \beta$ parameter space.

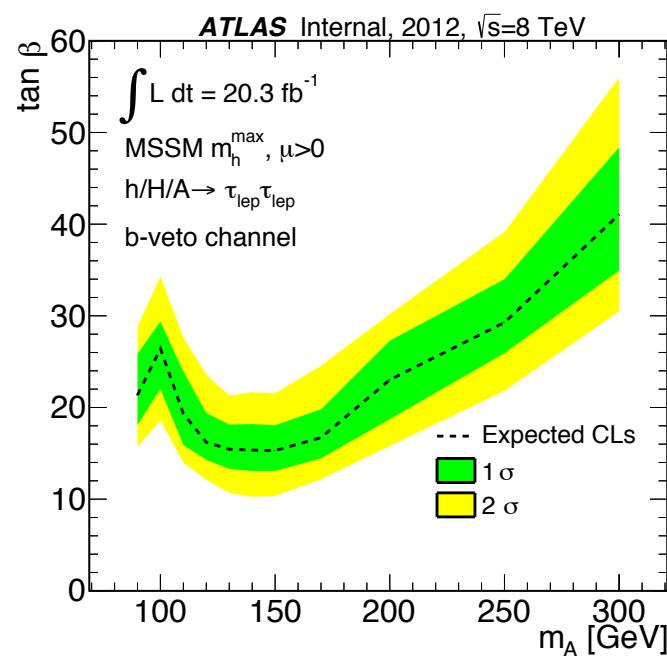


Figure C.11: Expected exclusion limits, using the b-veto channel, for MSSM Higgs boson production in the MSSM m_A vs $\tan \beta$ parameter space.

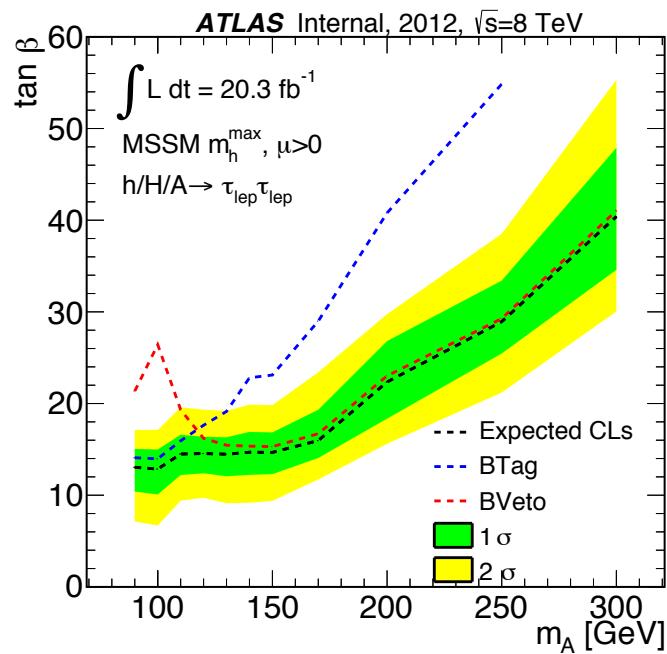
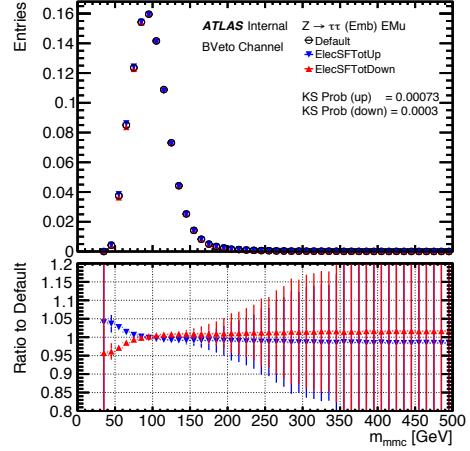
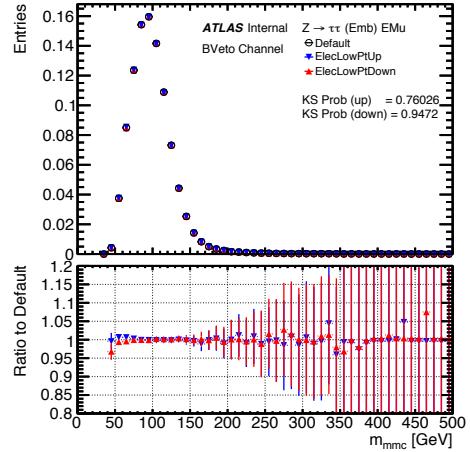


Figure C.12: Expected exclusion limits for MSSM Higgs boson production in the MSSM m_A vs $\tan \beta$ parameter space. Limits are compared for the b-tag and b-veto channel with the combined limit from both channels.

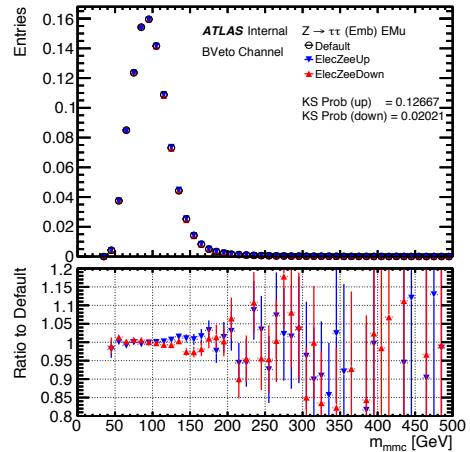
C.3 Shape Systematics



(a)



(b)



(c)

Figure C.13: Effect on the $m_{\tau\tau}^{MMC}$ distribution of the embedding sample due to the (a) the electron reconstruction and identification systematics, (b) the electron low p_T energy scale systematic and (c) the electron Zee energy scale systematic. The plots are made after the full b-veto category selection.