Search for a Light NMSSM Higgs Boson in the b-Quark Final State at the LHC

Dissertation
zur Erlangung des Doktorgrades
des Fachbereichs Physik
der Universität Hamburg

vorgelegt von

Gregor Mittag, geb. Hellwig

aus Berlin

Hamburg
2015

Gutachter der Dissertation:	Dr. Rainer Mankel Prof. Dr. Peter Schleper Dr. Giacomo Polesello
Gutachter der Disputation:	Dr. Rainer Mankel Prof. Dr. Peter Schleper Prof. Dr. Christian Sander Prof. Dr. Georg Weiglein
Datum der Disputation:	20. Juli 2015
Vorsitzender des Prüfungsausschusses:	Prof. Dr. Jan Louis
Vorsitzender des Promotionsausschusses:	Prof. Dr. Jan Louis
Leiter des Fachbereichs Physik:	Prof. Dr. Peter Hauschildt
Dekan der Fakultät für Mathematik, Informatik und Naturwissenschaften:	Prof. Dr. Heinrich Graener

Abstract

A search for a light neutral Higgs boson decaying into a pair of b quarks is presented. Such Higgs bosons are predicted by certain scenarios of the Next-to-Minimal Supersymmetric Standard Model (NMSSM), whose Higgs sector has two complex Higgs doublets and an additional Higgs singlet. If the lightest Higgs boson has a sizeable singlet admixture, it may have escaped detection in previous searches at LEP, Tevatron, and LHC, but could be produced at a substantial rate in neutralino decays within supersymmetric cascades. This analysis is the first search for Higgs bosons in the mass range below 100 GeV in the b-quark final state at the LHC. It uses a data sample corresponding to an integrated luminosity of $19.7\,\mathrm{fb}^{-1}$, which has been recorded by the CMS experiment in 2012 at a centre-of-mass energy of 8 TeV. Events with a pair of b-tagged jets, at least two very energetic jets, and large missing energy are selected. The invariant mass of the two selected b-tagged jets is used to extract the signal. The observed invariant-mass spectrum is in good agreement with the background expected from Standard-Model processes. The results are interpreted in terms of cross-section limits and within light-Higgs-boson scenarios of the NMSSM. The modified P4 benchmark scenario of the NMSSM is excluded by this analysis.

Kurzdarstellung

Diese Arbeit präsentiert eine Suche nach einem leichten, neutralen Higgs-Boson, welches in ein Paar von b-Quarks zerfällt. Ein derartiges Higgs-Boson wird in gewissen Szenarien des Nächstminimalen Supersymmetrischen Standardmodells (NMSSM) vorhergesagt. Der NMSSM-Higgs-Sektor besitzt zwei komplexe Higgs-Dubletts und ein zusätzliches Higgs-Singulett-Feld. Falls das leichteste Higgs-Boson einen großen Singulett-Anteil hat, könnte es in vorhergehenden Suchen bei LEP, Tevatron und LHC unentdeckt geblieben sein. Es würde jedoch in Neutralino-Zerfällen innerhalb von supersymmetrischen Kaskaden in beachtlichem Maße produziert werden. Diese Analyse ist die erste Suche am LHC nach Higgs-Bosonen im Massenbereich unter 100 GeV im b-Quark-Endzustand. Es werden Daten verwendet, die durch das CMS-Experiment im Jahr 2012 bei einer Schwerpunktsenergie von 8 TeV aufgezeichnet wurden und einer integrierten Luminosität von $19.7\,\mathrm{fb}^{-1}$ entsprechen. Ereignisse mit einem Paar von b-Quark-Jets, mindestens zwei hochenergetischen Jets und großer fehlender Transversalenergie werden selektiert. Die invariante Masse der zwei b-Quark-Jets wird mit dem vom Standardmodell erwarteten Untergrund verglichen und ist gut mit diesem verträglich. Die Ergebnisse werden in Form von oberen Grenzen auf Wirkungsquerschnitte und innerhalb von NMSSM-Szenarien mit leichtem Higgs-Boson interpretiert. Ein Standardszenario des NMSSM, das modifizierte P4-Szenario, wird durch diese Analyse ausgeschlossen.

Contents

Um klar zu sehen, genügt oft
ein Wechsel der Blickrichtung.

Antoine de Saint-Exupéry

Introduction

The driving force behind physics research is curiosity. The desire for an elementary understanding of nature fosters creative thinking and the development of new ideas to answer fundamental questions about the origin and the building blocks of the universe. Particle physics aims to address these points at a very basic level: The interactions between elementary matter particles, denoted as quarks and leptons, by means of mediating *gauge bosons* are studied on the basis of relativistic quantum mechanics. The electroweak and strong interactions are described in the framework of the *Standard Model of Particle Physics* (SM) [1–3]. It has resisted numerous high-precision tests that confirmed its validity over a wide energy-range [4], though achieving this with a relatively small set of parameters. Many phenomenological findings are incorporated in the SM, e.g. the coupling structure of the weak interactions or the violation of the CP symmetry. Its success culminated on the 4th of July, 2012 in the announcement of the discovery of a particle with a mass of 125 GeV [5, 6] that is at the current experimental accuracy consistent with the Higgs boson predicted by the SM [7–9].

However, there are clear indications that the SM is incomplete: Most notably, the fourth known type of interaction, gravity, is not included in the SM. In addition, it provides no candidate to explain the observation of dark matter [10–15]. It is also subject to the hierarchy problem [16–19], i.e. the mass of the Higgs boson requires large fine-tuning in the underlying theory parameters. In particular, the latter is addressed by the concept of *supersymmetry* (SUSY) that connects bosonic and fermionic degrees of freedom and thereby automatically achieves the desired fine tuning. Even a dark-matter candidate can be easily provided, and the most promising approaches to a quantum-mechanical description of gravity implement SUSY [20].

A major consequence of imposing SUSY on the SM is the extension of the particle spectrum. In particular, the Higgs sector is enlarged in any supersymmetric extension of the SM. This has motivated searches for additional Higgs bosons, especially in the range above the mass of the known Higgs boson at 125 GeV because previous searches for the SM Higgs boson have not found a signal in the low mass region. But an additional Higgs boson is not necessarily SM-like. For instance, models with an additional singlet field extend the Higgs sector with particles that have reduced couplings to gauge bosons. If the lightest Higgs boson has a sizeable singlet admixture, the prior constraints on the low Higgs-boson-mass range are evaded and its mass could well be below 125 GeV.

The Large Hadron Collider (LHC) [21] near Geneva allows the study of particle

collisions at unprecedented energy scales, thus enabling new production modes of a hypothetical light Higgs boson. In cascade decays of very heavy SUSY particles, the light Higgs boson might be produced at a sizeable rate. Therefore, a search for such a light Higgs boson in supersymmetric cascades is performed in this thesis.

A phenomenological study is carried out to investigate how the search results can be interpreted in a wider range of light-Higgs-boson scenarios in an appropriate way. The possibility of a simplified model is discussed and compared with a realistic model exhibiting a singlet extension. In particular, the structure of the cascade decays of SUSY particles is studied using various theoretical tools combined in a dedicated software-framework. The outcome of this theoretical analysis is used as basis for an extended interpretation of the results obtained in the central analysis of this thesis.

The experimental analysis, constituting the main part of this work, is a search for a light Higgs boson decaying into a pair of b quarks using the data taken by the Compact Muon Solenoid (CMS) experiment [22] at a centre-of-mass energy of 8 TeV corresponding to $19.7\,\mathrm{fb}^{-1}$. The b-quark final state is chosen because the decay rate into these elementary particles is highest for the Higgs boson under study. The signature required by the analysis includes two very energetic jets, two b-quark jets, and high transverse momentum imbalance. The contributions expected from the SM are either estimated with a data-driven technique, or from simulation and validated in dedicated control regions. The invariant mass of the two selected b-quark jets is used to measure the signal cross-section. In absence of signal, upper cross-section limits are set and exclusion limits are derived on the parameters of the *Next-to-Minimal Supersymmetric Standard Model* (NMSSM) [23–27].

This work is structured as follows. First, the theoretical basis is reviewed in Chapter 1, beginning with the particles, symmetries and mechanisms of the SM, followed by an introduction to SUSY and models implementing it, and concluding with the description of certain light-Higgs-boson scenarios. Chapter 2 presents the phenomenological study described above. The experimental setup to perform the search for light Higgs bosons is described in Chapter 3, i.e. the CMS experiment at the LHC, the event simulation, and reconstruction are reviewed. Chapter 4 is devoted to a specific part of this setup, the alignment of the inner tracking system installed in CMS. This part is crucial for the identification of b-quark jets, and important development in the validation and further related workflows is made in the course of this thesis. The search for a light Higgs boson in SUSY cascades and decaying into a pair of b quarks is presented in Chapter 5. The results of this analysis and their interpretation are described in the subsequent Chapter 6. Chapter 7 concludes this thesis with a short summary of the main results.

1 The Standard Model and its Supersymmetric Extensions

The physics of elementary particles aims to describe nature at its most basic level. The *Standard Model of Particle Physics* (SM) [1–3] is to date the most successful approach to this objective [4]. It describes very precisely all known elementary particles and their interactions, with the exception of gravity. Global fits of SM predictions to data show a good agreement between theory and observation with a $\chi^2/\text{ndf} = 48.3/44$ [4].

Despite this success, there are several clear indications for physics beyond the Standard Model (BSM), e.g. the mentioned description of gravity. But even the SM itself poses the hierarchy problem [16–19] (see Section 1.1.6). Also the missing dark matter [10–15] candidate and the deviation of the SM prediction of the anomalous magnetic moment of the muon from the measured value [4, 28] are a motivation to look for theoretical alternatives.

One promising candidate, which naturally solves the hierarchy problem and also offers solutions to other problematic aspects of the SM, is given by supersymmetry (SUSY).

This chapter gives an overview of the basic ideas of the SM and SUSY. Afterwards, the *Minimal Supersymmetric Standard Model* (MSSM) [29, 30] and its minimal supersymmetric extension, the *Next-to-Minimal Supersymmetric Standard Model* (NMSSM) [23–27] will be briefly described.

1.1 The Standard Model of Particle Physics

The most compelling argument for the SM is its comparatively simple structure and the relatively small particle content. These items will be discussed in the following subsections after a brief phenomenological overview of this field. When quoting experimental results, the system of natural units is used in this section and throughout this thesis, i.e. the fundamental physical constants obey $\hbar = c = 1$. Thus, all units are represented as eV^{n}, whereas $\text{n} = 1$ for masses, momenta, and energies.

1.1.1 Phenomenological Overview

The known elementary particles can be categorised according to their spin quantum number which is an intrinsic property of each particle. Half-integer-spin particles are called *fermions*, and particles with an integer spin are referred to as *bosons*. The known fundamental fermions are called *leptons* and *quarks*, and are grouped into three generations of two leptons and two quarks, each (see Table 1.1). The

first generation leptons are the *electron* (e) and *electron neutrino* (ν_e), whereas the corresponding quarks are denoted as *up quark* (u) and *down quark* (d). Most of the matter in our present universe can be approximately represented by these four particles. The constituents of atomic nuclei, *protons* (p) and *neutrons* (n), can be described as compounds of u quarks and d quarks, while the atomic shell consists of electrons.

Table 1.1: Matter particles of the Standard Model (generations $i = 1, 2, 3$). For each field and gauge symmetry, the respective quantum numbers are listed.

	Field	Colour	Weak Isospin $\tau^3/2$	Hypercharge $Y/2$
leptons	$\ell_{i\mathrm{L}} = (\nu_{i\mathrm{L}} \quad e_{i\mathrm{L}})^T$	0	$\pm 1/2$	$-1/2$
	$e_{i\mathrm{R}}$	0	0	-1
quarks	$q_{i\mathrm{L}} = (u_{i\mathrm{L}} \quad d_{i\mathrm{L}})^T$	1, 2, 3	$\pm 1/2$	$1/6$
	$u_{i\mathrm{R}}$	1, 2, 3	0	$2/3$
	$d_{i\mathrm{R}}$	1, 2, 3	0	$-1/3$

The particles of the second and third generation are heavier twins of the first generation fermions, i.e. they have identical properties, except for their masses. The leptons are denoted as *muon* (μ) and *tau* (τ), and their corresponding neutrinos ν_μ and ν_τ. The quarks are labelled *charm* (c), *strange* (s), *top* (t), and *bottom* (b) quarks. The fermion masses range from $m_\mathrm{e} = (510.998\,928 \pm 0.000\,011)\,\mathrm{keV}$ to $m_\mathrm{t} = (173.21 \pm 0.51 \pm 0.71)\,\mathrm{GeV}$, while electron neutrinos have a mass below a few eV [4]. The origin of this large mass diversity is not explained by the SM.

The electromagnetic interaction can be described in the framework of *Quantum Electrodynamics* (QED), which has been formulated in a relativistically fully covariant way based on [31–38]. The interaction is mediated between electrically charged objects via the exchange of a massless, electrically neutral vector-boson, the *photon* (γ). The neutrinos are neutral as well, while the other leptons carry an electric charge of -1, given in units of the elementary charge. The quarks have a charge of $+2/3$ and $-1/3$ for up-type and the down-type quarks, respectively. In consequence of the massless and chargeless photon, electromagnetic interactions are long ranged.

In addition to the electric charge, all fermions in the SM carry a *weak charge*, and the corresponding interaction is mediated by Z^0- and $\mathrm{W}^\pm$- boson exchange. Contrary to the photon, these bosons are massive with masses of $m_{\mathrm{Z}^0} = (91.1876 \pm 0.0021)\,\mathrm{GeV}$ and $m_\mathrm{W} = (80.385 \pm 0.015)\,\mathrm{GeV}$, respectively. Thus, the weak interaction is suppressed in the energy regime well below the $\mathrm{Z}^0/\mathrm{W}^\pm$ mass scale. Nonetheless, the Z^0 boson is phenomenologically similar to the photon, but the interactions mediated by the $\mathrm{W}^\pm$ boson change the type and charge of the involved fermion. This manifests itself, e.g. in the radioactive β^--decay, which can be traced back to the conversion of a d quark into a u quark under emission of a W^- boson. An important experimental finding is the parity violation of the weak interaction, i.e. it is not symmetric under spatial point-reflections [39]. In addition, certain types of weak decays also violate the CP symmetry [40], i.e. the product of charge conjugation and parity.

The last type of interaction in the context of the SM is the strong interaction described by the *Quantum Chromodynamics* (QCD), where the corresponding vector boson is the massless *gluon* (g), that mediates between objects with *colour charge*, i.e. quarks and gluons themselves, commonly denoted as *partons*. A unique feature of QCD is its increasing strength with increasing distance, referred to as *confinement*. This is a consequence of the number of coloured states and the self interaction of the gluons. The phenomenological implications are that only "colourless" states remain after the process of *hadronisation*. Hadrons can be understood as composite particles that consist, in a simplified picture, of two (*mesons*) or three (*baryons*), or possibly more quarks, referred to as *valence quarks*. These constantly exchange gluons which create for very short time-scales quark-antiquark pairs, referred to as *sea quarks*. The momentum fraction x, that is carried by each parton in a proton, which is a baryon, is phenomenologically described with *parton distribution functions* (PDF) (see Fig. 1.1). These are highest for gluons at low x (solid line in Fig. 1.1), and highest for the proton valence quarks (u, d in Fig. 1.1) at high x.

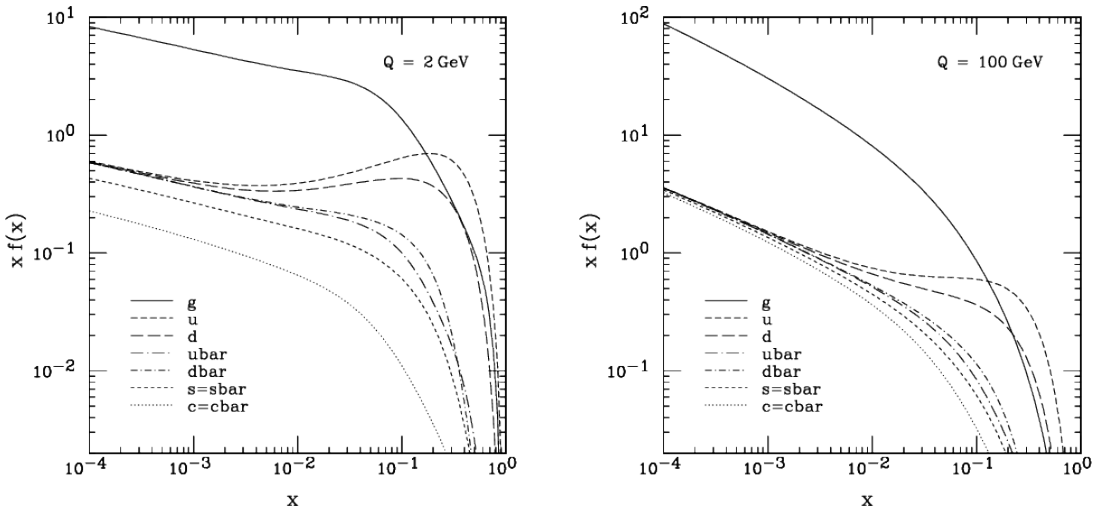

Figure 1.1: CTEQ6M parton-distribution functions, multiplied by x, as a function of x for two different energy scales Q for gluons (solid line) and quarks (dashed lines). Taken from [41].

However, for small distances or high energies, partons can be considered as free particles with respect to strong interactions. This phenomenon is referred to as *asymptotic freedom* [42, 43].

1.1.2 Symmetries and Particle Content

The plurality of phenomena described in the previous section are organised by the symmetries of the SM, which are reviewed in this section. From a theoretical point of view, the SM is a relativistic *Quantum Field Theory* (QFT), i.e. a synthesis of special relativity and quantum theory. For this reason, it is subject to the space-time symmetries of nature, which are translation invariance and Lorentz invariance,

commonly denoted as Poincaré invariance. The interactions of the SM depend on space-time and must be invariant under transformations of the Poincaré group. The generators of the group are the translation operators P^μ, $\mu \in \{0,1,2,3\}$, and the generators of the Lorentz group $J^{\mu\nu}$, $\mu,\nu \in \{0,1,2,3\}$. These fulfil the conditions

$$[P^\mu, P^\nu] = 0 \,, \tag{1.1a}$$
$$[P^\mu, J^{\rho\sigma}] = \mathrm{i}(g^{\mu\rho}P^\sigma - g^{\mu\sigma}P^\rho) \,, \tag{1.1b}$$
$$[J^{\mu\nu}, J^{\rho\sigma}] = \mathrm{i}(g^{\nu\rho}J^{\mu\sigma} - g^{\mu\rho}J^{\nu\sigma} + g^{\mu\sigma}J^{\nu\rho} - g^{\nu\sigma}J^{\mu\rho}) \,, \tag{1.1c}$$

where $g^{\mu\nu}$ is the metric tensor of the Minkowski space. Equation (1.1) describes a Lie algebra, whose concept is explained in more detail in [44]. The particles described by the SM are leptons, quarks, gauge bosons (see Tables 1.1, 1.2), and the Higgs boson. The existence of all of these particles has been experimentally confirmed. In case of the Higgs boson a particle was recently found which is in agreement with the expectations from the SM [7–9]. The interactions of these particles follow from a gauge theory over a corresponding QFT [45]. A common feature is the usage of (special) unitary gauge groups under whose transformations the theory is invariant, a concept referred to as Yang–Mills theory [46].

Table 1.2: Gauge bosons of the Standard Model (indices $j = 1,\ldots,3$, $a = 1,\ldots,8$). The respective gauge group, coupling, and generators are listed.

Field	Gauge Group	Coupling	Generators
B_μ	$U(1)_Y$	g_Y	$Y/2$
W_μ^j	$SU(2)_\mathrm{L}$	g_2	$\tau^j/2$
A_μ^a	$SU(3)_\mathrm{C}$	g_3	$\lambda^a/2$

The gauge group of the SM is $SU(3)_\mathrm{C} \times SU(2)_\mathrm{L} \times U(1)_Y$, where the indices denote colour (C), left chirality (L), and the weak hypercharge (Y). The subgroup $SU(2)_\mathrm{L} \times U(1)_Y$ describes the electroweak interaction and can be considered separately from the strong interaction. Under the gauge group $SU(2)_\mathrm{L}$ the left-chiral fermion fields f_L transform as doublets and the right-chiral fermion fields f_R transform as singlets, i.e. they carry no weak isospin which is the charge corresponding to this gauge group. The complete electroweak gauge transformations read as follows

$$f_\mathrm{L}(x) = \frac{1}{2}(1-\gamma_5)f(x) \to \exp\left(-\mathrm{i}g_Y\alpha_Y(x)\frac{Y}{2} - \mathrm{i}g_2\alpha_2^j(x)\frac{\tau^j}{2}\right) f_\mathrm{L}(x) \,, \tag{1.2a}$$
$$f_\mathrm{R}(x) = \frac{1}{2}(1+\gamma_5)f(x) \to \exp\left(-\mathrm{i}g_Y\alpha_Y(x)\frac{Y}{2}\right) f_\mathrm{R}(x) \,. \tag{1.2b}$$

Here, $(1-\gamma_5)/2$ and $(1+\gamma_5)/2$ are the left-chiral and right-chiral projector, i.e. they extract the respective component from the general fermion field f. The quantities g_Y and g_2 are the gauge couplings of the $U(1)_Y$ and the $SU(2)_\mathrm{L}$, respectively. The $\alpha_Y(x)$ and $\alpha_2^j(x)$ are the corresponding gauge functions. Furthermore, $Y/2$ is the

hypercharge operator and the $\tau^j/2 = T^j$ are the three generators[1] of the $SU(2)_\mathrm{L}$. Here, and in the following, the Einstein summation convention [47] is used, i.e. the sum is taken over all values of an index appearing twice in a single term, if not stated otherwise.

In respect of the $SU(3)_\mathrm{C}$ gauge group, the quarks and leptons transform differently because only the former carry colour charge,

$$q_{L,R}(x) \to \exp\left(-\mathrm{i}g_3\alpha_3^a(x)\frac{\lambda^a}{2}\right) q_{L,R}(x) \;, \tag{1.3a}$$

$$\ell_{L,R}(x) \to \ell_{L,R}(x) \;. \tag{1.3b}$$

Here, g_3 is the $SU(3)_\mathrm{C}$ gauge coupling and $\alpha_3^a(x)$ is the corresponding gauge function. The $\lambda^a/2$ are the eight generators[2] of the $SU(3)_\mathrm{C}$ and form a Lie algebra in the same way as the $\tau^j/2$ of the $SU(2)_\mathrm{L}$,

$$\left[\frac{\tau^j}{2}, \frac{\tau^k}{2}\right] = \mathrm{i}\varepsilon^{jkl}\frac{\tau^l}{2} \;, \qquad \left[\frac{\lambda^a}{2}, \frac{\lambda^b}{2}\right] = \mathrm{i}f^{abc}\frac{\lambda^c}{2} \;, \tag{1.4}$$

where ε^{jkl} and f^{abc} are the structure constants of the corresponding algebra.

1.1.3 Lagrangian Density

In the Lagrangian formalism, only one scalar function L is needed to completely describe a system by means of the principle of stationary action. The time integral over L yields the action, which is invariant under global symmetry transformations. If the described variable quantities are fields over the Minkowski space, L equals the space integral over a function of space-time, referred to as Lagrangian density $\mathcal{L}$. After introduction of the gauge transformations (1.2) and (1.3) for all matter fields, the $SU(3)_\mathrm{C} \times SU(2)_\mathrm{L} \times U(1)_Y$ gauge-invariant Lagrangian density reads

$$\mathcal{L} = \mathcal{L}_\mathrm{matter} + \mathcal{L}_\mathrm{gauge} \;, \quad \text{where} \tag{1.5a}$$

$$\mathcal{L}_\mathrm{matter} = \bar{\ell}_{i\mathrm{L}}\mathrm{i}\not{D}\ell_{i\mathrm{L}} + \bar{e}_{i\mathrm{R}}\mathrm{i}\not{D}e_{i\mathrm{R}} + \bar{q}_{i\mathrm{L}}\mathrm{i}\not{D}q_{i\mathrm{L}} + \bar{u}_{i\mathrm{R}}\mathrm{i}\not{D}u_{i\mathrm{R}} + \bar{d}_{i\mathrm{R}}\mathrm{i}\not{D}d_{i\mathrm{R}} \;, \tag{1.5b}$$

$$\mathcal{L}_\mathrm{gauge} = -\frac{1}{4}F_a^{\mu\nu}F^a_{\mu\nu} - \frac{1}{4}W_j^{\mu\nu}W^j_{\mu\nu} - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} \;. \tag{1.5c}$$

The derivative operator $\partial_\mu := \partial/\partial x^\mu$ is extended to the covariant derivative

$$D_\mu = \partial_\mu + \mathrm{i}g_3\frac{\lambda^a}{2}A^a_\mu + \mathrm{i}g_2\frac{\tau^j}{2}W^j_\mu + \mathrm{i}g_Y\frac{Y}{2}B_\mu \;, \tag{1.6}$$

[1] The τ^j are the *Pauli* matrices [44]. T^j is an alternative notation for the generators.
[2] The λ^a are the *Gell-Mann* matrices [44].

which ensures local gauge invariance if the introduced gauge fields A^a_μ, W^j_μ and B_μ obey the following transformations in component fields:

$$A^a_\mu \to A^a_\mu + \partial_\mu \alpha^a_3 - g_3 f^{abc} A^b_\mu \alpha^c_3 \,, \qquad (a,b,c = 1,\ldots,8) \,, \tag{1.7a}$$

$$W^j_\mu \to W^j_\mu + \partial_\mu \alpha^j_2 - g_2 \varepsilon^{jkl} W^k_\mu \alpha^l_2 \,, \qquad (j,k,l = 1,2,3) \,, \tag{1.7b}$$

$$B_\mu \to B_\mu + \partial_\mu \alpha_Y \,. \tag{1.7c}$$

The field strength tensors, which are introduced in (1.5c), are defined as

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g_3 f^{abc} A^b_\mu A^c_\nu \,, \tag{1.8a}$$

$$W^j_{\mu\nu} = \partial_\mu W^j_\nu - \partial_\nu W^j_\mu - g_2 \varepsilon^{jkl} W^k_\mu W^l_\nu \,, \tag{1.8b}$$

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

As stated above, only the Lagrangian density (1.5a) is needed to derive physical observables predicted by a given theory in terms of scattering amplitudes. However, a closed-form expression of these quantities is usually not possible. Hence, a perturbative expansion in terms of the involved couplings is performed. The first non-vanishing order of this expansion is referred to as leading order (LO). Higher orders are usually denoted by prefixing the letter N for every order considered in the calculation, i.e. the next to leading order is abbreviated with NLO, etc.

A diagrammatic approach, proposed by Feynman in 1948 [48], is used to ease the calculations. If these diagrams do not contain closed lines, they are referred to as tree level. Higher order corrections often involve diagrams with closed lines, referred to as loops. These loops give rise to divergent expressions leading to infinities, hence meaningless results. This problem is cured by the procedure of renormalisation, and a theory to which it can be applied is called renormalisable. For details on renormalisability see, e.g. [49, 50].

1.1.4 Electroweak Symmetry Breaking

The gauge symmetries introduced in Section 1.1.2 forbid explicit mass terms for the gauge bosons in the Lagrangian density. However, the gauge bosons $\mathrm{W}^\pm$ and Z^0, which were observed in 1983 for the first time [51–53], are massive. This implies that the $SU(2)_\mathrm{L} \times U(1)_Y$ symmetry has to be broken at the electroweak scale $\sim m_{\mathrm{W/Z}^0}$ to allow for gauge-boson-mass terms. In the SM, the masses are generated by means of the Brout–Englert–Higgs (BEH) mechanism [54–56], which introduces a complex scalar field. A non-vanishing *vacuum expectation value* (VEV) of the field is postulated that spontaneously breaks the $SU(2)_\mathrm{L} \times U(1)_Y$ symmetries. The only remaining symmetry $U(1)_\mathrm{em}$ is attributed to the electromagnetic interaction mediated by the massless photon.

The complex field, called *Higgs* field, introduced to implement *Electroweak Sym-*

metry Breaking (EWSB), is a scalar $SU(2)_\mathrm{L}$-doublet ϕ with $Y = 1$,

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \ . \tag{1.9}$$

The Lagrangian density is then extended by a gauge-invariant term

$$\mathcal{L}_\mathrm{Higgs} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi) \ , \tag{1.10}$$

where $V(\phi)$ is the Higgs potential with the maximum number of terms that fulfil the requirements of gauge invariance and renormalisability,

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \ , \qquad \mu^2, \lambda > 0 \ . \tag{1.11}$$

The postulation of a non-vanishing VEV for the Higgs field,

$$\langle \phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \ , \qquad v = \sqrt{\frac{\mu^2}{\lambda}} \ , \tag{1.12}$$

which is the minimum of $V(\phi)$, leaves the vacuum no longer invariant under transformations of $SU(2)_\mathrm{L} \times U(1)_Y$. The symmetry is spontaneously broken. Expanding the Higgs field around v,

$$\phi = \begin{pmatrix} G^+ \\ \frac{1}{\sqrt{2}} \left[v + h(x) + \mathrm{i} G^0 \right] \end{pmatrix} \ , \tag{1.13}$$

results in bilinear terms for the gauge bosons B_μ and W_μ^j because v is simply a number. Thus, the interaction terms of the gauge bosons with the Higgs field are reduced to a product of v, two gauge boson fields, and the gauge coupling. From these bilinear terms, the gauge-boson mass matrix is constructed, which yields after diagonalisation the boson masses. The real field $h(x)$ is the physical Higgs boson. The fields G^0 and $G^\pm$ are the Goldstone bosons, which arise when breaking a continuous symmetry [57, 58]. These are unphysical and do not contribute to observable quantities. Because of the hereby introduced gauge-boson mixing terms, the gauge eigenstates are no longer the observable mass-eigenstates. The latter can be obtained via a rotation of the fields. Accordingly, one arrives at three massive gauge-bosons and a massless one

$$W_\mu^\pm = \frac{1}{\sqrt{2}} \left(W_\mu^1 \mp \mathrm{i} W_\mu^2 \right) \ , \qquad m_\mathrm{W} = \frac{g_2}{2} v = m_{\mathrm{Z}^0} \cos\theta_W \ , \tag{1.14a}$$

$$Z_\mu = \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu \ , \qquad m_{\mathrm{Z}^0} = \frac{v}{2} \sqrt{g_Y^2 + g_2^2} \ , \tag{1.14b}$$

$$A_\mu = \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu \ , \qquad m_A = 0 \ . \tag{1.14c}$$

Since A_μ is massless, it can be identified with the photon of $U(1)_\mathrm{em}$. The electroweak

mixing angle θ_W has the following relations to the gauge couplings

$$\sin\theta_W = \frac{g_Y}{\sqrt{g_Y^2 + g_2^2}} , \qquad \cos\theta_W = \frac{g_2}{\sqrt{g_Y^2 + g_2^2}} . \tag{1.15}$$

The generator of the electromagnetic charge Q of the remaining $U(1)_{\mathrm{em}}$ relates to the generators of the $SU(2)_{\mathrm{L}}$ and $U(1)_Y$ via the Gell-Mann–Nishijima formula [59, 60],

$$Q = \frac{\tau^3}{2} + \frac{Y}{2} . \tag{1.16}$$

The corresponding gauge coupling e is obtained as

$$e = g_2 \sin\theta_W = g_Y \cos\theta_W . \tag{1.17}$$

The non-vanishing VEV of the Higgs field yields via *Yukawa* interactions also the fermion mass terms of the SM, which would break gauge invariance if introduced explicitly,

$$\mathcal{L}_{\mathrm{Yukawa}} = -y_{ij}^{e*}\bar{\ell}_{i\mathrm{L}}\phi e_{jR} - y_{ij}^{d*}\bar{q}_{i\mathrm{L}}\phi d_{jR} - y_{ij}^{u*}\bar{q}_{i\mathrm{L}}\phi^C u_{jR} + \mathrm{h.\,c.} , \tag{1.18}$$

where $\phi^C = \mathrm{i}\tau^2\phi^*$ is the charge-conjugate Higgs field and y_{ij}^k are the Yukawa couplings. The latter are, together with the already introduced four parameters g_Y, g_2, λ, and μ, free parameters of the SM[3]. The fermion mass matrices are then

$$m_{ij}^e = \frac{v}{\sqrt{2}} y_{ij}^{e*}, \qquad m_{ij}^d = \frac{v}{\sqrt{2}} y_{ij}^{d*}, \qquad m_{ij}^u = \frac{v}{\sqrt{2}} y_{ij}^{u*} . \tag{1.19}$$

These mass matrices can be diagonalised by rotation matrices, which appear within the SM only in flavour-changing interactions. The specific product of the matrices appearing in these interaction terms is combined to the Cabibbo–Kobayashi–Maskawa (CKM) matrix [61, 62] for quarks, whereas for neutrinos the corresponding matrix is named after Pontecorvo, Maki, Nakagawa, and Sakata (PMNS) owing to their work on neutrino oscillation [63, 64].

1.1.5 Complete Standard Model Lagrangian Density

In the previous sections, the main parts of the SM Lagrangian density are described. In addition, there are two other contributions, $\mathcal{L}_{\mathrm{ghost}}$ and $\mathcal{L}_{R_\xi}$, which are introduced by the quantisation procedure applied to the SM. The explicit formulas are not listed here for brevity. They are unphysical and cancel out in calculations of observable quantities like cross sections or decay widths.[4]

The complete SM Lagrangian density is the sum of all before mentioned contribu-

[3] One could also choose another set of equivalent parameterisations, like e.g. e, s_W, m_{Z^0}, and m_h, which can be more directly obtained in measurements.

[4] Details on the quantisation can be found in [45].

tions, including the above-mentioned quantisation related terms, and reads

$$\mathcal{L}_{\mathrm{SM}} = \mathcal{L}_{\mathrm{matter}} + \mathcal{L}_{\mathrm{gauge}} + \mathcal{L}_{\mathrm{Higgs}} + \mathcal{L}_{\mathrm{Yukawa}} + \mathcal{L}_{R_\xi} + \mathcal{L}_{\mathrm{ghost}} \,. \tag{1.20}$$

1.1.6 Hierarchy Problem

Although the SM describes the currently known phenomena remarkably well, it is obviously incomplete since it cannot describe physics at the reduced Planck scale $M_{\mathrm{P}} = 2.43 \times 10^{18}\,\mathrm{GeV}$, where gravitational effects become important. Therefore, at some scale between the electroweak regime $\sim m_{\mathrm{W}}$ and M_{P}, new physics phenomena, e.g. heavy particles, will most probably emerge.

This gives rise to a difficulty which is actually not a problem of the SM itself: The physical mass of the SM Higgs boson is very sensitive to the masses of possible new heavy particles. In higher-order corrections, like depicted in the one-loop Feynman diagrams in Fig. 1.2, quadratic dependencies on the masses of the heavy particles arise.

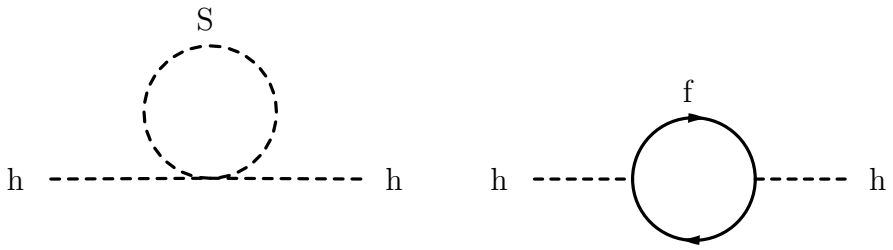

Figure 1.2: One-loop quantum corrections to the Higgs squared mass parameter m_{h}^2, due to a scalar S (left), and a Dirac fermion f (right).

The sign of contributions from scalar loops (see Fig. 1.2, left) is opposite to contributions from fermionic loops (see Fig. 1.2, right). The mass hierarchy between the Higgs boson mass at the order of $10^2\,\mathrm{GeV}$ and the Planck scale at the order of $10^{18}\,\mathrm{GeV}$ would require an enormous fine-tuning of these contributions to obtain a physical mass at the weak scale. This is the hierarchy problem of the SM [16–19]. Because it is considered unnatural to achieve this mass hierarchy, one typically introduces an additional symmetry that connects the otherwise unrelated contributions. A promising candidate for such a symmetry is *supersymmetry*, which is introduced in the next section.

1.2 Supersymmetry

A relatively large group of theories beyond the SM involves the concept of supersymmetry (SUSY), which connects bosonic and fermionic degrees of freedom. From the theoretical point of view, SUSY is motivated by the desire for a consistent combination of the space-time symmetries formulated in the Poincaré algebra (1.1) with

the internal symmetries. The Coleman–Mandula theorem [65] states that all symmetry groups with bosonic generators decompose into a direct product of the Poincaré group and the internal symmetries. A symmetry between bosons and fermions evades this theorem [66]. This symmetry, called supersymmetry, has been observed first in the context of string theories [67–69] and has been later applied to quantum field theories [70, 71].

According to the Haag–Łopuszánski–Sohnius theorem [72], the only non-trivial extension of (1.1) is the introduction of N fermionic generators Q_i, $i = 1, \ldots, N$. In four dimensions and provided that to date only fields with maximum spin 2 are describable, the maximum possible number of generators can be up to $N = 8$. However, only $N = 1$-supersymmetric quantum field theories are suitable for the description of chiral fields [73]. Therefore, only $N = 1$ will be considered in the following. In Weyl spinor notation (see Appendix A.2), this yields in addition to the Poincaré algebra (1.1) the following relations for the introduced fermionic generators (see [72] or [74])

$$\{Q_A, \bar{Q}_{\dot{B}}\} = 2\,(\sigma^\mu)_{A\dot{B}}\,P_\mu\,, \qquad \{\bar{Q}^{\dot{A}}, Q^B\} = 2\,(\bar{\sigma}^\mu)^{\dot{A}B}\,P_\mu\,, \tag{1.21a}$$

$$\{Q_A, Q_B\} = 0\,, \qquad \{\bar{Q}^{\dot{A}}, \bar{Q}^{\dot{B}}\} = 0\,, \tag{1.21b}$$

$$[Q_A, P_\mu] = 0\,, \qquad [\bar{Q}^{\dot{A}}, P_\mu] = 0\,, \tag{1.21c}$$

$$[J_{\mu\nu}, Q_A] = -(\sigma_{\mu\nu})_A{}^B Q_B\,, \qquad [J_{\mu\nu}, \bar{Q}^{\dot{A}}] = -(\bar{\sigma}_{\mu\nu})^{\dot{A}}{}_{\dot{B}} Q^{\dot{B}}\,, \tag{1.21d}$$

with $\sigma^\mu = (\mathcal{I}, \sigma^p)$, $\bar{\sigma}^\mu = (\mathcal{I}, -\sigma^p)$, $\sigma^p (p = 1, 2, 3)$ being the Pauli matrices. The operators Q_A and $\bar{Q}^{\dot{A}}$ transform bosonic states into fermionic states and vice versa. States which transform as irreducible representations of the super-Poincaré algebra in (1.1) and (1.21) form super multiplets. The simplest form requiring $P^2 \neq 0$ is the Wess–Zumino multiplet [75], which contains a two-component Weyl spinor and a complex scalar field. These can be interpreted as their respective superpartners. Since the mass operator P^2 commutes with Q_A and $\bar{Q}^{\dot{A}}$, the Weyl fermion and the scalar field have equal mass in exact SUSY.

1.2.1 Superfield Formalism

In order to construct a supersymmetric theory it is advantageous to introduce the *superspace*, which is a product of the Minkowski space and a four-dimensional Graßmann space (see Appendix A.1). Functions defined on the superspace are called *superfields* and automatically preserve SUSY as is demonstrated in this section. Thus, the particle content of a given supersymmetric theory can be easily extended, if one adds a set of fields, that constitute a complete superfield. The particles known from the SM can then be identified with certain components of a superfield, while the remaining components predict the *superpartners* of the SM particles. Therefore, the superfield formalism is an elegant way to construct supersymmetric theories. The different types of superfields are reviewed that are needed to incorporate the SM particles within SUSY models.

The superspace coordinates are denoted with $(x^\mu, \theta^A, \bar{\theta}_{\dot{A}})$, where x^μ is an element

of the Minkowski space and θ^A, $\bar{\theta}_{\dot{A}}$ $(A = 1, 2)$ are Graßmann coordinates which span the Graßmannian subspace of the superspace. Graßmann variables are nilpotent, i.e. the products $\theta^i\theta^i$ and $\bar{\theta}_i\bar{\theta}_i$ vanish. Thus, superfields can be decomposed into a finite number of component fields,

$$\begin{aligned}\mathcal{F}(z) \equiv \mathcal{F}(x,\theta,\bar{\theta}) = f(x) + \sqrt{2}\theta\xi(x) + \sqrt{2}\bar{\theta}\bar{\chi}(x) + \theta\theta M(x) + \bar{\theta}\bar{\theta}N(x) \\ + \theta\sigma^\mu\bar{\theta}A_\mu(x) + \theta\theta\ \bar{\theta}\bar{\lambda}(x) + \bar{\theta}\bar{\theta}\ \theta\zeta(x) + \frac{1}{2}\theta\theta\ \bar{\theta}\bar{\theta}D(x)\ ,\end{aligned} \tag{1.22}$$

where the notation $\theta\theta \equiv -2\theta^1\theta^2$, $\bar{\theta}\bar{\theta} \equiv -2\bar{\theta}_{\dot{1}}\bar{\theta}_{\dot{2}}$ is used. The fields $f(x)$, $M(x)$, $N(x)$ and $D(x)$ are scalar, $A_\mu(x)$ is a vector field, $\xi_A(x)$, $\zeta_A(x)$ are left-handed and $\bar{\chi}^{\dot{A}}(x)$, $\bar{\lambda}^{\dot{A}}(x)$ are right-handed Weyl spinor fields.

A global SUSY transformation on the superspace changes the superspace coordinates as follows

$$(x^\mu, \theta, \bar{\theta}) \to (x^\mu - \mathrm{i}\theta\sigma^\mu\bar{\epsilon} + \mathrm{i}\epsilon\sigma^\mu\bar{\theta}, \theta + \epsilon, \bar{\theta} + \bar{\epsilon})\ . \tag{1.23}$$

Thereby, one finds a representation of the generators Q_A and $\bar{Q}^{\dot{A}}$ on the superspace,

$$Q_A = -\mathrm{i}\left(\partial_A + \mathrm{i}\sigma^\mu_{A\dot{B}}\bar{\theta}^{\dot{B}}\partial_\mu\right)\ , \qquad \bar{Q}^{\dot{A}} = -\mathrm{i}\left(\bar{\partial}^{\dot{A}} + \mathrm{i}\theta^B\sigma^\mu_{B\dot{B}}\epsilon^{\dot{B}\dot{A}}\partial_\mu\right)\ , \tag{1.24}$$

whereby the transformation of a general superfield (1.22) can be expressed using Q_A and $\bar{Q}^{\dot{A}}$ as

$$\mathcal{F} \to \mathcal{F} + \mathrm{i}(\epsilon Q + \bar{\epsilon}\bar{Q})\mathcal{F}\ . \tag{1.25}$$

If one expands the right hand side of (1.25), one observes that the D-component of $\mathcal{F}$ is transformed into itself and a total derivative with respect to space-time. Since the latter yields no contribution to the action, one can construct supersymmetric actions from D-terms.

Chiral Superfields

In order to construct irreducible chiral multiplets, one defines the superfields Φ and $\Phi^\dagger$ which hold the conditions

$$\bar{\mathcal{D}}_{\dot{A}}\Phi = 0\ , \qquad \mathcal{D}_A\Phi^\dagger = 0\ . \tag{1.26}$$

The fields Φ and $\Phi^\dagger$ are called left- and right-chiral superfields. The derivative $\mathcal{D}_A := \partial_A - \mathrm{i}(\sigma^\mu)_{A\dot{B}}\bar{\theta}^{\dot{B}}\partial_\mu$ is defined such that it transforms covariant under (1.23). If one introduces left- and right-chiral superspace coordinates $y^\mu = x^\mu - \mathrm{i}\theta\sigma^\mu\bar{\theta}$ and $\bar{y}^\mu = x^\mu + \mathrm{i}\theta\sigma^\mu\bar{\theta}$, one can decompose the chiral superfields:

$$\Phi(y,\theta) = \phi(y) + \sqrt{2}\theta\xi(y) + \theta\theta F(y)\ , \tag{1.27a}$$

$$\Phi^\dagger(\bar{y},\bar{\theta}) = \phi^*(\bar{y}) + \sqrt{2}\bar{\theta}\bar{\xi}(\bar{y}) + \bar{\theta}\bar{\theta}F^*(\bar{y})\ . \tag{1.27b}$$

The Weyl spinor $\xi(y)$ and the complex scalar field $\phi(y)$ are regarded as superpartners, and $F(y)$ is an auxiliary field. An important property of the F-component of a chiral superfield is that it transforms under (1.23) into itself and a total derivative with respect to space-time. This facilitates the construction of a general supersymmetric Lagrangian density from chiral superfields,

$$\mathcal{L} = \int \mathrm{d}^4\theta\ \Phi_i^\dagger \Phi_i + \int \mathrm{d}^2\theta\ \Big(\mathcal{W}(\Phi_i) + \mathrm{h.\,c.}\Big)\ , \tag{1.28a}$$

$$\mathcal{W}(\Phi_i) = a_i\Phi_i + \frac{1}{2} m_{ij}\Phi_i\Phi_j + \frac{1}{3!} y_{ijk}\Phi_i\Phi_j\Phi_k\ . \tag{1.28b}$$

The function $\mathcal{W}(\Phi_i)$ is the superpotential. The invariance of $\mathcal{L}$ under SUSY transformations implies that $\mathcal{W}$ is an analytic function of the left-chiral superfields Φ, i.e. it does not contain any of the $\Phi^\dagger$ fields.

Vector Superfields and Gauge Couplings

Vector superfields $V(x,\theta,\bar\theta)$ are superfields that fulfil the condition $V^\dagger = V$. This means, that the components f, A_μ and D are real-valued and it holds $\bar\chi = \bar\xi$, $N = M^*$, $\zeta = \lambda$.

One can construct vector superfields out of chiral superfields. For instance, if $\mathrm{i}\Lambda$ is a chiral superfield, the combination $(\mathrm{i}\Lambda - \mathrm{i}\Lambda^\dagger)$ is a vector superfield. This relation in connection with the property that the sum of vector superfields is again a vector superfield, allows the definition of an Abelian supergauge-transformation [76],

$$V \to V + \mathrm{i}\Lambda - \mathrm{i}\Lambda^\dagger\ . \tag{1.29}$$

The name "supergauge transformation" is justified because the A_μ component of V transforms as under a classical Abelian gauge transformation $A_\mu \to A_\mu - 2\partial_\mu\, \Im\mathrm{m}\,\phi$. A vector superfield adopts in the Wess–Zumino gauge [75] a relatively simple form with respect to the general superfield (1.22):

$$V(z) = \theta\sigma^\mu\bar\theta A_\mu(x) + \theta\theta\ \bar\theta\bar\lambda(x) + \bar\theta\bar\theta\ \theta\lambda(x) + \frac{1}{2}\theta\theta\ \bar\theta\bar\theta D(x)\ , \tag{1.30}$$

where the components f, $\bar\xi$, M are gauged to zero. The component fields $\lambda(x)$ and $\bar\lambda(x)$ can be interpreted as superpartners of the real-valued vector field $A_\mu(x)$ whereas $D(x)$ is another auxiliary field.

1.2.2 Supergauge Invariant Lagrangian Density

A supergauge transformation for a non-Abelian gauge group is defined as

$$\Phi \to \exp\left(-\mathrm{i}\Lambda(z)\right)\Phi\ , \qquad \bar{\mathcal{D}}^{\dot A}\Lambda^a(z)T^a = 0\ , \tag{1.31a}$$

$$\Phi^\dagger \to \Phi^\dagger \exp\Big(+\mathrm{i}\Lambda^\dagger(z)\Big)\ , \qquad \mathcal{D}_A\Lambda^{a\dagger}(z)T^a = 0\ , \tag{1.31b}$$

where $\Lambda(z) \equiv 2g\Lambda^a(z)T^a$, $\Lambda^a(z)$ is a vector superfield and T^a are the generators of the gauge group. If one introduces a gauge vector superfield $V^a(z)$ which transforms as

$$\mathrm{e}^{V} \to \mathrm{e}^{-\mathrm{i}\Lambda^\dagger}\mathrm{e}^{V}\mathrm{e}^{\mathrm{i}\Lambda} \,, \tag{1.32a}$$

$$\mathrm{e}^{-V} \to \mathrm{e}^{-\mathrm{i}\Lambda}\mathrm{e}^{-V}\mathrm{e}^{\mathrm{i}\Lambda^\dagger} \,, \qquad \text{with } V \equiv 2gV^aT^a \,, \tag{1.32b}$$

then terms like $\Phi^\dagger \mathrm{e}^{V}\Phi$ are supergauge invariant. In order to construct kinetic terms for the gauge fields, one introduces the supersymmetric field strength tensor:

$$W_A = -\frac{1}{4}\bar{\mathcal{D}}\bar{\mathcal{D}}\mathrm{e}^{-V}\mathcal{D}_A\mathrm{e}^{V} \,, \qquad W_A = 2gW_A^aT^a \,, \tag{1.33a}$$

$$\bar{W}^{\dot{A}} = -\frac{1}{4}\mathcal{D}\mathcal{D}\mathrm{e}^{V}\bar{\mathcal{D}}^{\dot{A}}\mathrm{e}^{-V} \,, \qquad \bar{W}^{\dot{A}} = 2g\bar{W}^{\dot{A}a}T^a \,. \tag{1.33b}$$

It is evident from (1.32) that the hereby defined field strength tensors transform according to

$$W_A \to \mathrm{e}^{-\mathrm{i}\Lambda}W_A\mathrm{e}^{\mathrm{i}\Lambda} \,, \qquad \bar{W}^{\dot{A}} \to \mathrm{e}^{-\mathrm{i}\Lambda^\dagger}\bar{W}^{\dot{A}}\mathrm{e}^{\mathrm{i}\Lambda^\dagger} \,. \tag{1.34}$$

This means that terms of the form $\mathrm{Tr}[W^AW_A]$ and $\mathrm{Tr}[\bar{W}_{\dot{A}}\bar{W}^{\dot{A}}]$ are gauge invariant. Thereby, one can construct the general supersymmetric and supergauge-invariant Lagrangian density

$$\begin{aligned}\mathcal{L} = \frac{1}{16g^2C_r}\int \mathrm{d}^2\theta \ \mathrm{Tr}\left[W^AW_A + \bar{W}_{\dot{A}}\bar{W}^{\dot{A}}\right] \\ + \int \mathrm{d}^4\theta \ \Phi^\dagger \mathrm{e}^{V}\Phi + \int \mathrm{d}^2\theta \left(\mathcal{W}(\Phi) + \mathrm{h.\,c.}\right) ,\end{aligned} \tag{1.35}$$

where the superpotential $\mathcal{W}(\Phi)$ has to be gauge invariant. The constant C_r is the representation invariant of the gauge group in the adjoint representation.

From (1.35) one can see that a supersymmetric theory is completely described by the particle content in terms of superfields, the gauge symmetries, and the interactions of the superfields contained in the superpotential. The much more complex structure of SUSY models with their many particles can be derived from the compact formulation in (1.35) by performing the integration and decomposing the superfields.

1.2.3 Soft Supersymmetry Breaking

As pointed out in the beginning of Section 1.2, the mass operator P^2 commutes with all other generators of the super-Poincaré algebra. That is, all component fields of a chiral or a vector superfield have the same mass. However, no superpartners of quarks, leptons, and gauge bosons with the corresponding masses have been observed [4]. Therefore, SUSY has to be a broken symmetry. However, an appropriate SUSY breaking needs to respect gauge symmetry and renormalisability, and should avoid the reintroduction of quadratic divergences, which cancel out in exact SUSY.

A solution in agreement with phenomenological constraints is the introduction of terms into the Lagrangian density, which are free parameters and explicitly break SUSY. Especially because the reintroduction of quadratic divergences is avoided, this mechanism is referred to as *soft supersymmetry breaking* (SSB). The general SSB Lagrangian density then reads [77, 78]

$$\begin{aligned}\mathcal{L}_{\text{SSB}} = -\phi_i^*(m^2)_{ij}\phi_j + \Big(&\frac{1}{3!}\mathcal{A}_{ijk}\phi_i\phi_j\phi_k - \frac{1}{2}\mathcal{B}_{ij}\phi_i\phi_j + \mathcal{C}_i\phi_i \\ &- \frac{1}{2}M\tilde{\lambda}^a\tilde{\lambda}^a + \text{h. c.}\Big)\,,\end{aligned} \tag{1.36}$$

where ϕ is the scalar component of a chiral superfield and $\tilde{\lambda}^a$ is the fermionic component of a vector superfield. In the literature it is common to introduce the parameters A_{ijk}, B_{ij}, and C_i by defining $\mathcal{A}_{ijk}$, $\mathcal{B}_{ij}$, and $\mathcal{C}_i$ as the product of the newly introduced parameters and the corresponding superpotential parameters:

$$\begin{aligned}\mathcal{L}_{\text{SSB}} = -\phi_i^*(m^2)_{ij}\phi_j + \Big(&\frac{1}{3!}(yA)_{ijk}\phi_i\phi_j\phi_k - \frac{1}{2}(\mu B)_{ij}\phi_i\phi_j + (hC)_i\phi_i \\ &- \frac{1}{2}M\tilde{\lambda}^a\tilde{\lambda}^a + \text{h. c.}\Big)\,.\end{aligned} \tag{1.37}$$

The parameters A_{ijk}, B_{ij}, and C_i are called trilinear, bilinear, and linear soft-breaking parameters and are introduced for each superpotential parameter. The mass parameters M are called gaugino mass parameters and are introduced for each gauge-boson superpartner. The terms in (1.37) can be generated in various ways typically involving a hidden sector. Interactions between this hidden sector and the visible sector result in the soft SUSY breaking terms, e.g. via gravitational interactions in the gravity-mediated supersymmetry-breaking mechanism [29, 79].

1.3 The Minimal Supersymmetric Standard Model

With the superfield formalism introduced in the previous section, one has a recipe at hand to construct any supersymmetric model. As long as one uses superfields, the theory is automatically supersymmetric. This is analogous to using fields defined on Minkowski space to automatically construct a relativistic theory.

As pointed out in the beginning of this chapter, the SM describes the known phenomena with high precision. Therefore, it should be embedded in any realistic supersymmetric theory. The minimal approach to such a model with respect to its particle content is the *Minimal Supersymmetric Standard Model* (MSSM) [29, 30], which will be introduced in this section.

1.3.1 Symmetries and Particle Content

The MSSM obeys a supersymmetry with $N = 1$. The SM fermions correspond to Weyl spinors of chiral superfields, and the SM gauge bosons are described as vector

fields from vector superfields. The remaining particles of the respective superfields define the superpartners of the SM particles.

In this subsection, the superfield formalism is used to construct the MSSM as a concrete model. Since the gauge group $SU(3)_\mathrm{C} \times SU(2)_\mathrm{L} \times U(1)_Y$ of the SM is kept, one needs a vector superfield for each generator of this gauge group. This automatically introduces the superpartners of the gauge bosons, called *gauginos*. These are summarised at the bottom of Tables 1.3 and 1.4, respectively.

Table 1.3: Superfields of the MSSM and their respective component fields (generation indices $i \in 1,2,3$, $a \in 1,\ldots,8$ and $j \in 1,2,3$). The gauge superfields transform in the adjoint representation of the respective gauge group and are singlets with respect to the two other gauge groups.

Superfield	Component Fields		
	spin 0	spin $\frac{1}{2}$	spin 1
$Q_i = (Q_{u_i} \;\; Q_{d_i})^T$	$\tilde{\mathrm{q}}_{i\mathrm{L}} = (\tilde{u}_{i\mathrm{L}} \;\; \tilde{d}_{i\mathrm{L}})^T$	$q_{i\mathrm{L}} = (u_{i\mathrm{L}} \;\; d_{i\mathrm{L}})^T$	
$\bar{U}_i$	$\tilde{u}^*_{i\mathrm{R}}$	$u^C_{i\mathrm{R}}$	
$\bar{D}_i$	$\tilde{d}^*_{i\mathrm{R}}$	$d^C_{i\mathrm{R}}$	
$L_i = (L_{\nu_i} \;\; L_{e_i})^T$	$\tilde{\ell}_{i\mathrm{L}} = (\tilde{\nu}_{i\mathrm{L}} \;\; \tilde{e}_{i\mathrm{L}})^T$	$\ell_{i\mathrm{L}} = (\nu_{i\mathrm{L}} \;\; e_{i\mathrm{L}})^T$	
$\bar{E}_i$	$\tilde{e}^*_{i\mathrm{R}}$	$e^C_{i\mathrm{R}}$	
$H_1 = (H_1^0 \;\; H_1^-)^T$	$h_1 = (h_1^0 \;\; h_1^-)^T$	$\tilde{h}_{1L} = (\tilde{h}_{1L}^0 \;\; \tilde{h}_{1L}^-)^T$	
$H_2 = (H_2^+ \;\; H_2^0)^T$	$h_2 = (h_2^+ \;\; h_2^0)^T$	$\tilde{h}_{2L} = (\tilde{h}_{2L}^+ \;\; \tilde{h}_{2L}^0)^T$	
V_g^a		$\tilde{g}^a$	A_μ^a
V_W^j		$\tilde{\lambda}_W^j$	W_μ^j
V_Y		$\tilde{\lambda}_Y^0$	B_μ

For each matter particle, of the SM a superfield with the corresponding quantum numbers is introduced, i.e. the SM fermions get superpartners with spin 0 and the Higgs fields get spin-1/2 superpartners. In addition, the Higgs sector needs to be extended in supersymmetric theories, as will be discussed in the next subsection.

Higgs Sector of the MSSM

Similar to the SM, electroweak symmetry breaking in the MSSM is realised via the BEH mechanism. However, unlike the SM, the MSSM exhibits two Higgs doublets. This extended Higgs sector is required by the property of the superpotential (1.28b) to contain only left-chiral superfields. Therefore, one cannot simply generalise the Yukawa terms $y_{ij}^{u*}\overline{q_{i\mathrm{L}}}\phi^C u_{jR}$ of the SM introduced for the up-type fermions. This term contains the charge-conjugate Higgs field which has to be a component of a right-chiral field, while the other two would be components of left-chiral fields. Thus, one always needs two Higgs doublets, one is coupling to up-type fermions and one is coupling to down-type fermions, respectively.

Table 1.4: Dimensions of the representations of the MSSM superfields with respect to each of the MSSM gauge groups. For the abelian group $U(1)_Y$ the quantum number $Y/2$ is listed.

Superfield	$SU(3)_\mathrm{C}$	$SU(2)_\mathrm{L}$	$U(1)_Y$
Q_i	$\mathbf{3}$	$\mathbf{2}$	$\frac{1}{6}$
$\bar{U}_i$	$\bar{\mathbf{3}}$	$\mathbf{1}$	$-\frac{2}{3}$
$\bar{D}_i$	$\bar{\mathbf{3}}$	$\mathbf{1}$	$\frac{1}{3}$
L_i	$\mathbf{1}$	$\mathbf{2}$	$-\frac{1}{2}$
$\bar{E}_i$	$\mathbf{1}$	$\mathbf{1}$	1
H_1	$\mathbf{1}$	$\mathbf{2}$	$-\frac{1}{2}$
H_2	$\mathbf{1}$	$\mathbf{2}$	$\frac{1}{2}$
V_g^a	$\mathbf{8}$	$\mathbf{1}$	0
V_W^j	$\mathbf{1}$	$\mathbf{3}$	0
V_Y	$\mathbf{1}$	$\mathbf{1}$	0

Another reason for having two doublets are gauge anomalies arising in triangular diagrams of three abelian gauge bosons as external particles. At lowest order of perturbation theory this diagram is forbidden by the gauge symmetries. Therefore, it must not be generated at higher orders in perturbation theory, i.e. the sum of diagrams like in Fig. 1.3 must vanish. In the case of the $U(1)_Y$ symmetry, this means that $\sum_n Y_n \stackrel{!}{=} 0$ must hold. In the SM, this is accidentally fulfilled, but the additional particles in the MSSM would violate this condition without an additional Higgs doublet. Having justified the need of two Higgs doublets, the implications on the phenomenology of the scalar part of the Higgs doublets are described in the following.

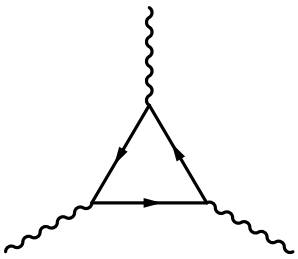

Figure 1.3: Triangular diagram of three abelian gauge bosons at one-loop order.

The scalar $SU(2)_\mathrm{L}$-doublet h_1 couples only to down-type fermions, whereas the second Higgs doublet h_2 couples only to up-type fermions. The two doublets read in

components

$$h_1 = \begin{pmatrix} h_1^0 \\ h_1^- \end{pmatrix} , \qquad h_2 = \begin{pmatrix} h_2^+ \\ h_2^0 \end{pmatrix} \tag{1.38}$$

and can be expanded analogous to the SM to

$$h_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \left(v_1 + \phi_1^0 - \mathrm{i}\gamma_1^0\right) \\ \phi_1^- \end{pmatrix} , \qquad h_2 = \begin{pmatrix} \phi_2^+ \\ \frac{1}{\sqrt{2}} \left(v_2 + \phi_2^0 - \mathrm{i}\gamma_2^0\right) \end{pmatrix} , \tag{1.39}$$

around their VEVs v_1 and v_2. Assuming CP conservation, one gets after diagonalisation of the Higgs mass matrices five physical Higgs bosons. Two CP-even states h and H, one CP-odd state A, and two charged states $\mathrm{H}^\pm$. The remaining degrees of freedom are absorbed by the gauge bosons, which acquire mass analogous to the SM. The ratio of the two VEVs is defined as

$$\tan\beta = \frac{v_2}{v_1} . \tag{1.40}$$

The mass parameter of the CP-odd Higgs boson, M_A, and $\tan\beta$ are free parameters of the MSSM and govern the phenomenology of the Higgs sector at tree level. The supersymmetric partners of the Higgs bosons, the Higgsinos, will be covered in the next subsection.

Neutralinos and Charginos

Due to EWSB the fermionic components of the gauge superfields and the Higgs superfields can mix with each other. Since $U(1)_\mathrm{em}$ remains unbroken, charged and neutral particles mix separately. The former are referred to as *charginos* and their corresponding mass term in the Lagrangian density reads

$$\mathcal{L}_\mathrm{char,mass} = -(\psi^-)^T \mathbf{X} \psi^+ + \mathrm{h.\,c.} , \tag{1.41}$$

where the two-dimensional vectors ψ^-, ψ^+ are composed of two-component Weyl spinors of the charged gauginos and higgsinos,

$$\psi^+ = \begin{pmatrix} \tilde{\lambda}^+ \\ \tilde{h}_2^+ \end{pmatrix} , \qquad \psi^- = \begin{pmatrix} \tilde{\lambda}^- \\ \tilde{h}_1^- \end{pmatrix} , \qquad \tilde{\lambda}^\pm = \frac{1}{\sqrt{2}} \left(\tilde{\lambda}_W^1 \mp \mathrm{i}\tilde{\lambda}_W^2\right) . \tag{1.42}$$

The mixing matrix $\mathbf{X}$ is given by the soft-breaking parameters in (1.37) and terms induced by EWSB,

$$\mathbf{X} = \begin{pmatrix} M_2 & \sqrt{2} m_\mathrm{W} \sin\beta \\ \sqrt{2} m_\mathrm{W} \cos\beta & \mu \end{pmatrix} . \tag{1.43}$$

Here, m_W is the W-boson mass and M_2 is the Wino mass parameter, which is the soft-breaking parameter of the superpartner of the W boson. The parameter μ is a dimensionful superpotential parameter, which is also referred to as Higgsino mass

parameter. This parameter poses certain problems that motivate an extension of the MSSM as discussed in Section 1.3.3.

The matrix $\mathbf{X}$ can be diagonalised by two rotation matrices U and V, which yields four two-component chargino mass-eigenstates. These can be combined to two four-component Dirac spinors $\tilde{\chi}^{\pm}_{1,2}$ and (1.41) then reads

$$\mathcal{L}_{\text{char,mass}} = -\sum_{i=1}^{2} m_{\tilde{\chi}^{\pm}_i} \overline{\chi^{+}_i} \chi^{+}_i \ , \tag{1.44}$$

with the chargino masses

$$\begin{aligned} m^2_{\chi^{\pm}_{1,2}} = \frac{1}{2}\Big[& M_2^2 + \mu^2 + 2m_{\text{W}}^2 \\ & \pm \sqrt{(M_2^2 + \mu^2 + 2m_{\text{W}}^2)^2 - 4(M_2\mu - m_{\text{W}}^2 \sin 2\beta)^2} \Big] \ . \end{aligned} \tag{1.45}$$

The Lagrangian density of the mixing neutral gauginos and Higgsinos can be obtained analogous to the charginos:

$$\mathcal{L}_{\text{neu,mass}} = -\frac{1}{2}(\psi^0)^T \mathbf{Y} \psi^0 + \text{h. c.} \ , \tag{1.46}$$

where the neutral fields involved can be written in the vector

$$\psi^0 = (\tilde{\lambda}^0_Y, \tilde{\lambda}^3_W, \tilde{h}^0_1, \tilde{h}^0_2)^T \ . \tag{1.47}$$

The mass matrix $\mathbf{Y}$ for these fields reads

$$\mathbf{Y} = \begin{pmatrix} M_1 & 0 & -\frac{1}{2}g_1 v_1 & \frac{1}{2}g_1 v_2 \\ 0 & M_2 & \frac{1}{2}g_2 v_1 & -\frac{1}{2}g_2 v_2 \\ -\frac{1}{2}g_1 v_1 & \frac{1}{2}g_2 v_1 & 0 & -\mu \\ \frac{1}{2}g_1 v_2 & -\frac{1}{2}g_2 v_2 & -\mu & 0 \end{pmatrix} \ . \tag{1.48}$$

In addition to the already introduced parameters, there is now the soft-breaking Bino mass parameter M_1. After diagonalisation of $\mathbf{Y}$ one obtains four Weyl-spinor mass-eigenstates, which can be combined to four Majorana spinors $\tilde{\chi}^0_i$ $(i = 1, \ldots, 4)$, i.e. they are their own antiparticles. Equation (1.46) then reads

$$\mathcal{L}_{\text{neu,mass}} = -\frac{1}{2}\sum_{i=1}^{4} m_{\tilde{\chi}^0_i} \overline{\tilde{\chi}^0_i} \tilde{\chi}^0_i \ . \tag{1.49}$$

Sfermions

The remaining sector of the MSSM particle spectrum are the superpartners of the SM fermions, called squarks and sleptons and commonly denoted as sfermions. Their masses are given by the Yukawa couplings of the SM fermions, the corresponding soft-breaking parameters, and also by Higgs-related parameters as $\tan\beta$ via the BEH

mechanism. Thus, a thorough understanding of these relations is beneficial to comprehend the interplay between the individual particle sectors of the MSSM and related models.

The mass terms of the sfermions in the MSSM Lagrangian density can be written in the compact form

$$\mathcal{L}_{\tilde{\mathbf{f}},\text{bil}} = -\sum_{\tilde{\mathbf{f}}} \tilde{\mathbf{f}}^\dagger \mathcal{M}^2_{\tilde{\mathbf{f}}} \tilde{\mathbf{f}} \ , \tag{1.50}$$

where $\mathcal{M}^2_{\tilde{\mathbf{f}}}$ is the sfermion mass matrix and $\tilde{\mathbf{f}}$ is the vector which contains all mixing sfermions for a certain mixing sector. In the MSSM one has

$$\text{up-sleptons :} \qquad \tilde{\mathbf{f}} = (\tilde{\nu}_e, \tilde{\nu}_\mu, \tilde{\nu}_\tau)^T \ , \tag{1.51a}$$

$$\text{down-sleptons :} \qquad \tilde{\mathbf{f}} = (\tilde{e}_\text{L}, \tilde{\mu}_\text{L}, \tilde{\tau}_\text{L}, \tilde{e}_\text{R}, \tilde{\mu}_\text{R}, \tilde{\tau}_\text{R})^T \ , \tag{1.51b}$$

$$\text{up-squarks :} \qquad \tilde{\mathbf{f}} = (\tilde{u}_\text{L}, \tilde{c}_\text{L}, \tilde{t}_\text{L}, \tilde{u}_\text{R}, \tilde{c}_\text{R}, \tilde{t}_\text{R})^T \ , \tag{1.51c}$$

$$\text{down-squarks :} \qquad \tilde{\mathbf{f}} = (\tilde{d}_\text{L}, \tilde{s}_\text{L}, \tilde{b}_\text{L}, \tilde{d}_\text{R}, \tilde{s}_\text{R}, \tilde{b}_\text{R})^T \ . \tag{1.51d}$$

For example, the mass matrix of up-squarks is

$$\mathcal{M}^2_{\tilde{\mathbf{u}}} = \begin{pmatrix} \mathcal{M}_{11} & m^u_{ii}\left(A^u_{ii} - \frac{\mu \tan\beta}{\sqrt{2}}\right) \\ m^{u*}_{ii}\left(A^{u*}_{ii} - \frac{\mu \tan\beta}{\sqrt{2}}\right) & \mathcal{M}_{22} \end{pmatrix} \ , \tag{1.52a}$$

$$\mathcal{M}_{11} = m^2_{\tilde{\text{q}}_{i\text{L}}} + m^2_{\text{Z}^0} \cos 2\beta \left(\frac{1}{2} - \frac{2}{3} s^2_W\right) + |m^u_{ii}|^2 \ , \tag{1.52b}$$

$$\mathcal{M}_{22} = m^2_{\tilde{u}_{i\text{R}}} + \frac{2}{3} m^2_{\text{Z}^0} \cos 2\beta s^2_W + |m^u_{ii}|^2 \ , \tag{1.52c}$$

where mixing of sfermions of different flavours has been neglected because those terms are very small compared to the soft-breaking terms in $\mathcal{M}^2_{\tilde{\mathbf{u}}}$. These are, in addition to the already introduced parameters, the soft-breaking mass terms for the squarks $m^2_{\tilde{\text{q}}_{i\text{L}}}$, $m^2_{\tilde{u}_{i\text{R}}}$, and the soft-breaking terms A^u_{ii} corresponding to the Yukawa couplings. The quark mass terms $|m^u_{ii}|^2$ are defined as in (1.19). The mass matrix $\mathcal{M}^2_{\tilde{\mathbf{f}}}$ can be diagonalised with a unitary matrix $W^{\tilde{f}}$, which yields

$$\mathcal{L}_{\tilde{\mathbf{f}},\text{bil}} = -\sum_{\tilde{\mathbf{f}}} \tilde{\mathbf{f}}^\dagger \mathcal{M}^2_{\tilde{\mathbf{f}}} \tilde{\mathbf{f}} = -\sum_{\tilde{\mathbf{f}}} \tilde{\mathbf{f}}^{m\dagger} M^2_{\tilde{\mathbf{f}}} \tilde{\mathbf{f}}^m \ , \tag{1.53}$$

where $M^2_{\tilde{\mathbf{f}}} = W^{\tilde{f}\dagger} \mathcal{M}^2_{\tilde{\mathbf{f}}} W^{\tilde{f}}$ is diagonal and $\tilde{\mathbf{f}}^m = W^{\tilde{f}\dagger} \tilde{\mathbf{f}}$ are mass eigenstates. In the following, flavour mixing is neglected like in (1.52) such that only left- and right-handed sfermions of the same type mix. The matrix $W^{\tilde{f}}$ then reads for one generation

$$W^{\tilde{f}} = \begin{pmatrix} \cos\theta_{\tilde{f}} & -\sin\theta_{\tilde{f}} \\ \sin\theta_{\tilde{f}} & \cos\theta_{\tilde{f}} \end{pmatrix} \ , \tag{1.54}$$

where $\theta_{\tilde{f}}$ is the mixing angle for the respective sfermion and generation. The mass eigenstates of a certain type are then denoted with the numbers 1, 2 according to their mass

$$\text{up-sleptons :} \qquad \tilde{\mathbf{f}}^m = (\tilde{\nu}_e, \tilde{\nu}_\mu, \tilde{\nu}_\tau)^T \ , \tag{1.55a}$$

$$\text{down-sleptons :} \qquad \tilde{\mathbf{f}}^m = (\tilde{e}_1, \tilde{\mu}_1, \tilde{\tau}_1, \tilde{e}_2, \tilde{\mu}_2, \tilde{\tau}_2)^T \ , \tag{1.55b}$$

$$\text{up-squarks :} \qquad \tilde{\mathbf{f}}^m = (\tilde{u}_1, \tilde{c}_1, \tilde{t}_1, \tilde{u}_2, \tilde{c}_2, \tilde{t}_2)^T \ , \tag{1.55c}$$

$$\text{down-squarks :} \qquad \tilde{\mathbf{f}}^m = (\tilde{d}_1, \tilde{s}_1, \tilde{b}_1, \tilde{d}_2, \tilde{s}_2, \tilde{b}_2)^T \ . \tag{1.55d}$$

1.3.2 Complete MSSM Lagrangian Density

Using the general supersymmetric Lagrangian density (1.35), one just needs to insert the above-mentioned superfields to obtain the MSSM Lagrangian density. The only non-trivial part is the superpotential, which in case of the MSSM is given as

$$\mathcal{W}_{\text{MSSM}} = y_u (Q \cdot H_2) \bar{U} - y_d (Q \cdot H_1) \bar{D} - y_e (L \cdot H_1) \bar{E} - \mu (H_1 \cdot H_2) \ . \tag{1.56}$$

Here, $(A \cdot B)$ is an $SU(2)_{\text{L}}$-invariant product. The y are the generalised supersymmetric Yukawa couplings and are dimensionless. The only dimensionful parameter is μ, which will be discussed in more detail in Section 1.3.3.

The soft-breaking part of the MSSM Lagrangian density, following the general formula (1.37), reads

$$\begin{aligned}
\mathcal{L}_{\text{SSB}}^{\text{MSSM}} = & - m^2_{\tilde{\text{q}}_{i\text{L}}} |\tilde{\text{q}}_{i\text{L}}|^2 - m^2_{\tilde{u}_{i\text{R}}} |\tilde{u}_{i\text{R}}|^2 - m^2_{\tilde{d}_{i\text{R}}} |\tilde{d}_{i\text{R}}|^2 - m^2_{\tilde{\ell}_{i\text{L}}} |\tilde{\ell}_{i\text{L}}|^2 - m^2_{\tilde{e}_{i\text{R}}} |\tilde{e}_{i\text{R}}|^2 \\
& - m^2_{h_{1i}} |h_{1i}|^2 - m^2_{h_{2i}} |h_{2i}|^2 \\
& - \Big[\mu B_\mu (h_1 \cdot h_2) + (y^e A^e)(h_1 \cdot \tilde{\ell}_L) \tilde{e}^*_R \\
& \quad + (y^d A^d)(h_1 \cdot \tilde{\text{q}}_L) \tilde{d}^*_R + (y^u A^u)(\tilde{\text{q}}_L \cdot h_2) \tilde{u}^*_R + \text{h. c.} \Big] \\
& - \frac{1}{2} \left(M_3 \bar{\tilde{g}} \tilde{g} + M_2 \bar{\tilde{\lambda}}^j_W \tilde{\lambda}^j_W + M_1 \bar{\tilde{\lambda}}^0_Y \tilde{\lambda}^0_Y + \text{h. c.} \right) \ .
\end{aligned} \tag{1.57}$$

The first two lines contain the soft-breaking squark-, slepton-, and Higgs-masses, the third and fourth line contain the bilinear and trilinear soft-breaking terms which are associated to the μ term and the Yukawa couplings, and the last line contains the gaugino masses, i.e. gluino-, wino-, and bino- mass parameter.

R-Parity

The MSSM superpotential (1.56) is the minimal version needed to produce a phenomenological viable model. However, the general MSSM superpotential would contain terms which violate either conservation of baryon number B or lepton number L.

These terms would read as follows

$$\mathcal{W}_{\Delta L=1} = \frac{1}{2}\lambda^{ijk}(L_i \cdot L_j)\bar{E}_k + \lambda'^{ijk}(L_i \cdot Q_j)\bar{D}_k + \mu'^{i}(L_i \cdot H_2) \ , \tag{1.58a}$$

$$\mathcal{W}_{\Delta B=1} = \frac{1}{2}\lambda''^{ijk}\bar{U}_i\bar{D}_j\bar{D}_k \ , \tag{1.58b}$$

where i, j, k are family indices. The superfields carry baryon number $B = +1/3$ for Q_i, $B = -1/3$ for $\bar{U}_i$, $\bar{D}_i$ and $B = 0$ for all other fields. The lepton numbers are $L = +1$ for L_i, $L = -1$ for $\bar{E}_i$ and $L = 0$ for all others. That means, the terms in (1.58) violate the conservation of baryon number and lepton number by one unit, each.

The existence of these terms is problematic because they would lead to rapid proton decay via diagrams like the one given in Fig. 1.4 [80]. In addition to the depicted process, there are further experimental constraints on the violation of lepton and baryon number [81–92].

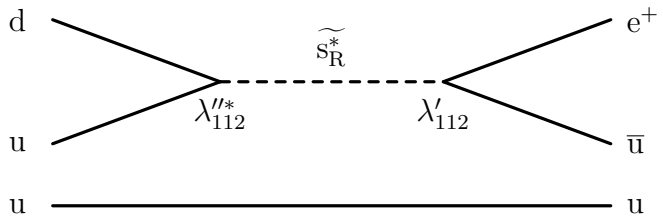

Figure 1.4: Tree-level Feynman diagram contributing to the proton decay. The decay is mediated via R-parity-violating couplings.

In principle one could postulate exact conservation of baryon number and lepton number. However, it is already known that they are violated by non-perturbative electroweak effects [93]. Therefore, a new symmetry is added to the MSSM, which eliminates the terms in (1.58), while allowing the terms in (1.56). This symmetry is called *R-parity* and is formulated as a multiplicatively conserved quantum number defined by

$$P_R = (-1)^{3(B-L)+2s} \ , \tag{1.59}$$

where s is the spin of the particle. By that definition, all SM particles have $P_R = +1$ and all supersymmetric particles carry $P_R = -1$. This symmetry has three important consequences for the phenomenology at collider experiments:

- Supersymmetric particles are produced in even numbers.
- Each supersymmetric particle must decay into a final state with an odd number of LSPs.
- The lightest supersymmetric particle (LSP) must be stable. If it is neutral and interacts only weakly, it is eligible to serve as a dark matter candidate.

1.3.3 μ Problem

Although the MSSM offers a solution to the hierarchy problem and other issues, it introduces another problem connected with hierarchy. The superpotential (1.56) contains the dimensionful parameter μ, which gives rise to the μ problem of the MSSM [94]. Since μ is a parameter of the superpotential, which might be generated via a mechanism at the grand unification (GUT) scale M_{GUT}, it should be of the the order of M_{GUT}. However, EWSB in the MSSM yields a relation between μ, the EWSB parameter $\tan\beta$, and the SUSY-breaking parameters $m_{h_1}^2$, $m_{h_2}^2$,

$$\mu^2 = \frac{m_{h_1}^2 - m_{h_2}^2 \tan^2\beta}{\tan^2\beta - 1} - \frac{1}{2} m_Z^2 \,. \tag{1.60}$$

If there are no large fine-tuned cancellations, μ has to be either at the order of the electroweak scale or at the order of the SUSY-breaking scale. Removing the parameter by an imposed symmetry, e.g. Z_3, and thereby hiding the problem, would lead to massless Higgsinos, which is in conflict with phenomenological constraints. Therefore, one has again the question for the natural scale of a model parameter. This question is addressed by the model described in the following section.

1.4 The Next-to-Minimal Supersymmetric Standard Model

A main motivation for a non-minimal extension of the SM is the μ problem mentioned in the previous section. A very promising approach that solves the problem, but still preserves the supersymmetric properties is the *Next-to-Minimal Supersymmetric Standard Model* (NMSSM) [23–27].

1.4.1 Symmetries and Particle Content

The model keeps the (MS)SM gauge group $SU(3)_{\mathrm{C}} \times SU(2)_{\mathrm{L}} \times U(1)_Y$. The extension comprises an additional superfield which transforms as a singlet with respect to the gauge groups of the model, i.e. it does not couple to the fields of the gauge bosons and gauginos. The interactions with the other fields are given in the superpotential of the NMSSM[5],

$$\begin{aligned} \mathcal{W}^{\mathrm{NMSSM}} = {} & y_u (Q \cdot H_2) \bar{U} - y_d (Q \cdot H_1) \bar{D} - y_e (L \cdot H_1) \bar{E} \\ & - \lambda S (H_1 \cdot H_2) + \frac{1}{3} \kappa S^3 \,, \end{aligned} \tag{1.61}$$

where the singlet superfield couples only to the two Higgs doublets with λ and to itself via κ. Note, that the μ parameter is not present in the superpotential anymore, but it is generated dynamically similar to the way SM fermions acquire their masses.

[5] To be more precise, it is the Z_3-symmetric NMSSM. See [95, 96] for details.

Also the scalar component of the singlet superfield acquires a VEV leading to an effective μ term

$$\mu_{\text{eff}} = \lambda v_s \ . \tag{1.62}$$

This way, a fixed scale for this parameter is avoided.

The new superfield and the accompanying parameters of the superpotential also lead to further SUSY-breaking parameters, that influence the mass matrices of the Higgs and neutralino sector,

$$\mathcal{L}_{\text{SSB}}^{\text{NMSSM}} = \mathcal{L}_{\text{SSB}}^{\text{MSSM}} + m_S^2 s^2 - \lambda A_\lambda s(h_1 \cdot h_2) + \frac{1}{3}\kappa A_\kappa s^3 \ , \tag{1.63}$$

where s is the scalar component of the singlet superfield.

The newly introduced singlet leaves the largest part of the MSSM untouched, but results in three additional particles in the Higgs and neutralino sector of the model. Their masses and interactions with the other fields are described in the following.

Higgs Sector of the NMSSM

Electroweak symmetry breaking in the NMSSM is implemented similarly to the MSSM, i.e. the two doublets are expanded around their VEVs as in (1.39), and the singlet field has the expansion

$$s = v_s + \frac{1}{\sqrt{2}}\left(\phi_s + \mathrm{i}\gamma_s\right) \ . \tag{1.64}$$

Assuming CP conservation in the Higgs sector, one obtains after EWSB and diagonalisation of the Higgs mass matrices seven physical Higgs bosons: three CP-even states h_1, h_2 and h_3, two CP-odd states a_1 and a_2, and two charged states $\mathrm{H}^\pm$. The remaining degrees of freedom are again absorbed by the gauge bosons, which acquire mass analogous to the SM and MSSM. The parameter $\tan\beta$ is defined as for the MSSM in (1.40).

Typically, one of the CP-even Higgs bosons can have a sizeable singlet admixture, which leads to very interesting phenomenological aspects as will be discussed in Section 1.4.2.

Neutralinos and Charginos

The fermionic component $\tilde{s}$ of the newly introduced singlet superfield S is called *singlino* and mixes with the existing four neutralinos known from the MSSM. The neutralino mass term in the Lagrangian density has the same form as for the MSSM in (1.46). However, the vector ψ^0 is five-dimensional in the NMSSM,

$$\psi^0 = (\tilde{\lambda}_Y^0, \tilde{\lambda}_W^3, \tilde{h}_1^0, \tilde{h}_2^0, \tilde{s})^T \ . \tag{1.65}$$

The mass matrix $\mathbf{Y}$ for these fields reads

$$\mathbf{Y} = \begin{pmatrix} M_1 & 0 & -\frac{1}{2}g_1 v_1 & \frac{1}{2}g_1 v_2 & 0 \\ 0 & M_2 & \frac{1}{2}g_2 v_1 & -\frac{1}{2}g_2 v_2 & 0 \\ -\frac{1}{2}g_1 v_1 & \frac{1}{2}g_2 v_1 & 0 & -\mu_{\text{eff}} & -\lambda v_2 \\ \frac{1}{2}g_1 v_2 & -\frac{1}{2}g_2 v_2 & -\mu_{\text{eff}} & 0 & -\lambda v_1 \\ 0 & 0 & -\lambda v_2 & -\lambda v_1 & 2\kappa v_s \end{pmatrix} . \tag{1.66}$$

The upper left part of (1.66) is identical to (1.48). The matrix in (1.66) can be diagonalised which yields five neutralino states $\tilde{\chi}_i^0$ $(i = 1, \ldots, 5)$.

The chargino sector of the NMSSM is identical to the chargino sector of the MSSM with μ_{eff} instead of μ, i.e. one has the chargino states $\tilde{\chi}_{1,2}^{\pm}$.

1.4.2 NMSSM Benchmark Scenarios

The extension of the particle content described above leads also to a rather generic feature of the NMSSM and other singlet extensions of the minimal model. The lightest CP-even Higgs boson h_1 has suppressed couplings to gauge bosons if it is mainly singlet-like, i.e. if the physical state h_1 obtained after diagonalisation of the CP-even mass matrix has a high admixture of the gauge singlet s. This suppression evades the experimental lower bounds on the Higgs boson mass given by the existing Higgs searches [97] and a Higgs boson with a mass even below m_{Z^0} is therefore not excluded. Light-Higgs-boson benchmark scenarios, i.e. NMSSM parameter settings representative for the above described variant of the model, are discussed in the following.

Modified P4 Benchmark Scenario

The *P4 benchmark scenario* of the NMSSM [98] is of special importance for light-Higgs-boson searches as pointed out in [99]. Adjustments of the soft-breaking parameter A_κ in this scenario are made in order to obtain different h_1 masses leading to the *modified P4 benchmark scenario*, whose masses and parameters are listed in Tables 1.5, 1.6. The h_1-mass values considered in this thesis are within $30 - 100\,\mathrm{GeV}$.

It features two important aspects: It has a light Higgs boson h_1, that decays predominantly into b quarks because these are the heaviest particles in the available phase space. The superpotential (1.61) contains no couplings between the pure singlet and the matter fields, thus the h_1 decay into SM fermions is only possible because of the non-zero Higgs-doublet component of h_1. The second interesting feature is the second lightest CP-even Higgs boson h_2, which is SM-like with a mass at $\approx 125\,\mathrm{GeV}$. Thus, it can be identified with the observed Higgs boson. The latter point is achieved by adjusting the trilinear soft-breaking parameters A_t, A_b, and A_τ (see Table 1.6). The tuning of the h_2 mass to the mass of the observed Higgs boson is a modification of the scenario described [99], which also contains only one h_1-mass point with $m_{\mathrm{h}_1} = 40\,\mathrm{GeV}$, in contrast to this thesis.

Table 1.5: Mass spectrum of the Higgs sector (top) and neutralino/chargino sector (bottom) in the modified NMSSM P4 scenario for example parameter points.

Mass (GeV)					
h_1	h_2	h_3	a_1	a_2	$H^\pm$
40	125.3	577	185	577	565
65	125.3	577	163	577	566
85	125.3	577	132	577	566

Mass (GeV)						
$\tilde{\chi}_1^0$	$\tilde{\chi}_2^0$	$\tilde{\chi}_3^0$	$\tilde{\chi}_4^0$	$\tilde{\chi}_5^0$	$\tilde{\chi}_1^\pm$	$\tilde{\chi}_2^\pm$
97.8	227	228	304	622	208	622

Table 1.6: Relevant parameters of the Higgs sector (top) and other sectors (bottom) in the modified NMSSM P4 scenario for an example parameter point ($m_{h_1} = 65\,\text{GeV}$).

Higgs Sector Parameter					
$\tan\beta$	λ	κ	A_λ	A_κ	μ_{eff}
2.6	0.6	0.12	$-510\,\text{GeV}$	$161.3\,\text{GeV}$	$-200\,\text{GeV}$

Non-Higgs Sector Parameter						
M_1	M_2	M_3	A_τ	A_b	A_t	M_{SUSY}
300 GeV	600 GeV	1.0 TeV	1.1 TeV	1.1 TeV	1.1 TeV	1.0 TeV

In this scenario, the squark and gluino masses are at the same values and therefore, the production of coloured SUSY particles includes all four production mechanisms

- gluino-gluino ($pp \to \tilde{g}\tilde{g}$, see Appendix B, Figs. B.1, B.2),
- squark-gluino ($pp \to \tilde{q}\tilde{g}$, see Appendix B, Fig. B.3),
- squark-squark ($pp \to \tilde{q}\tilde{q}$, see Appendix B, Figs. B.4, B.5),
- squark-antisquark ($pp \to \tilde{q}\bar{\tilde{q}}$, see Appendix B, Figs. B.4, B.5).

The respective cross sections, calculated with PROSPINO [100], are listed in Table 1.7. These processes have an impact on the h_1 production cross sections in SUSY cascades as described in Section 1.5.2.

Decoupled Squarks Scenario

A scenario similar to the above-mentioned modified P4 scenario is the *decoupled squarks scenario*, which has slightly different physics assumptions, leading to a different signal topology yet preserving a light Higgs boson. The main difference is that

Table 1.7: Cross sections for the various processes of coloured SUSY particle production in the modified P4 scenario. The values are calculated with PROSPINO 2.1. Their uncertainties vary within 10–30 %, depending on the process.

Subprocess		σ^{LO} (pb)	σ^{NLO} (pb)
gluino-gluino	$(pp \to \tilde{g}\tilde{g})$	0.184×10^{-2}	0.624×10^{-2}
squark-gluino	$(pp \to \tilde{q}\tilde{g})$	0.290×10^{-1}	0.499×10^{-1}
squark-squark	$(pp \to \tilde{q}\tilde{q})$	0.454×10^{-1}	0.530×10^{-1}
squark-antisquark	$(pp \to \tilde{q}\bar{\tilde{q}})$	0.752×10^{-2}	0.115×10^{-1}

the common sfermion-mass parameter M_{SUSY} is set to 2 TeV, thus leading to practically decoupled squarks. In order to maintain an h_2 boson at a mass of 125 GeV, the trilinear soft-breaking terms have to be adjusted (see Tables 1.8, 1.9). This scenario has been devised for this thesis to study the dependence of the Higgs production rate on the neutralino spectrum, while keeping the number of relevant parameters as low as possible. Therefore, the gaugino mass parameters M_1, M_2 and the gluino mass parameter M_3 vary as listed in Table 1.10.

Table 1.8: Mass spectrum of the Higgs sector (top) and neutralino/chargino sector (bottom) in the decoupled squarks scenario of the NMSSM for an example parameter point ($m_{h_1} = 65\,\text{GeV}$, $M_1 = 350\,\text{GeV}$, $M_2 = 600\,\text{GeV}$, $M_3 = 1\,\text{TeV}$).

Mass (GeV)					
h_1	h_2	h_3	a_1	a_2	$H^\pm$
65	125	575	175	574	566

Mass (GeV)						
$\tilde{\chi}_1^0$	$\tilde{\chi}_2^0$	$\tilde{\chi}_3^0$	$\tilde{\chi}_4^0$	$\tilde{\chi}_5^0$	$\tilde{\chi}_1^\pm$	$\tilde{\chi}_2^\pm$
98.1	228	230	352	634	209	634

In the modified P4 scenario, squark production of the first two generations is by far the dominant production mode because these originate from the quarks of the first two generations, which have the largest PDF values for high x values (see Fig 1.1). These are required for the production of the very heavy squarks and gluinos. This mode is suppressed in the decoupled squarks scenario, hence gluino-gluino production ($pp \to \tilde{g}\tilde{g}$) is by far the dominant mode for production of coloured SUSY particles. The respective cross sections, calculated with NLL-FAST [100–106], are listed for different gluino mass parameters in Table 1.11. In total, the inclusive production cross section of coloured SUSY particles is reduced

$$\sigma_{\text{P4}}(M_3 = M_{\text{SUSY}} = 1\,\text{TeV}) = 0.12\,\text{pb}\,, \tag{1.67a}$$

$$\sigma_{\text{decoupled}}(M_3 = 1\,\text{TeV}) = 0.024\,\text{pb}\,. \tag{1.67b}$$

Table 1.9: Relevant parameters of the Higgs sector (top) and other sectors (bottom) in the decoupled squarks scenario of the NMSSM for an example parameter point ($m_{\mathrm{h}_1} = 65\,\mathrm{GeV}$, $M_1 = 350\,\mathrm{GeV}$, $M_2 = 600\,\mathrm{GeV}$, $M_3 = 1\,\mathrm{TeV}$).

Higgs Sector Parameter					
$\tan\beta$	λ	κ	A_λ	A_κ	μ_{eff}
2.6	0.6	0.12	$-510\,\mathrm{GeV}$	$209.7\,\mathrm{GeV}$	$-200\,\mathrm{GeV}$

Non-Higgs Sector Parameter						
M_1	M_2	M_3	A_τ	A_{b}	A_{t}	M_{SUSY}
$350\,\mathrm{GeV}$	$600\,\mathrm{GeV}$	$1.0\,\mathrm{TeV}$	$1.9\,\mathrm{TeV}$	$1.9\,\mathrm{TeV}$	$1.9\,\mathrm{TeV}$	$2.0\,\mathrm{TeV}$

Table 1.10: Scanned values of NMSSM parameters. All other parameters are in the same regime as in the modified P4 scenario (see Tables 1.6, 1.9), except for the squark masses, which are set to 2 TeV leading to decoupled squarks, and the trilinear couplings which are raised to 1.9 TeV to adjust the mass of h_2.

NMSSM Parameter	Range (GeV)
m_{h_1}	$60 - 95$
M_1	$100 - 600$
M_2	$100 - 600$
M_3	$800 - 1400$

1.5 Light Higgs Production Mechanisms

The production modes of light Higgs bosons predicted by the benchmark scenarios described in Section 1.4.2 are summarised in this section, and the relative importance of these channels for light-Higgs-boson searches is discussed.

1.5.1 Standard Production Channels

Direct production mechanisms (see Fig. 1.5) like vector boson fusion (VBF) and production in association with vector bosons (VH) are suppressed because h_1 has reduced

Table 1.11: Cross sections for gluino-gluino production for different gluino masses in the decoupled squarks scenario. The values are calculated with NLL-FAST.

M_3 (GeV)	$\sigma^{NLO}_{\mathrm{pp}\to\widetilde{\mathrm{g}}\widetilde{\mathrm{g}}}$ (pb)
800	0.157
1000	0.244×10^{-1}
1200	0.440×10^{-2}
1400	0.871×10^{-3}

couplings to gauge bosons due to its large singlet component. The pure singlet does also not couple to SM fermions. Therefore, h_1 production in gluon fusion, which is proportional to its reduced squared couplings to quarks, is also suppressed. The light Higgs boson couples only to the other Higgs bosons and neutralinos/charginos at a sizeable rate.

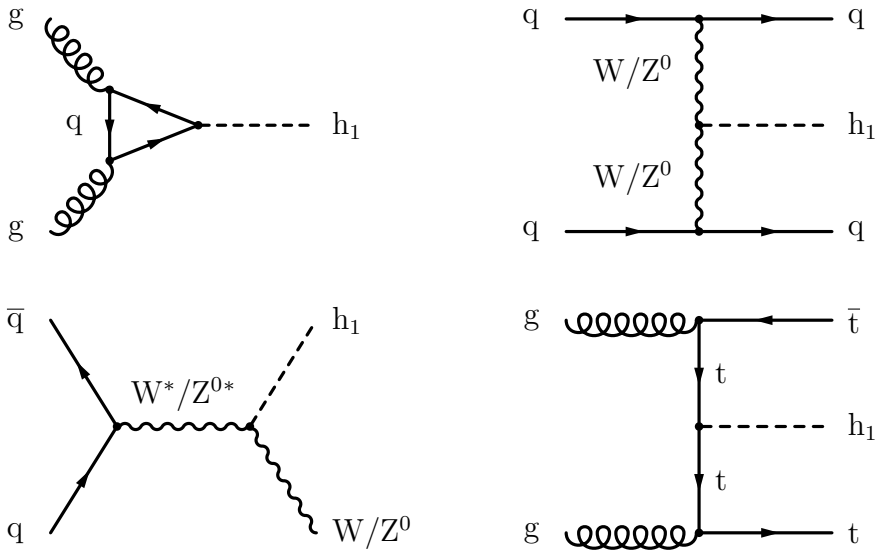

Figure 1.5: Standard Higgs production mechanisms at the LHC: Gluon fusion (top left), vector boson fusion (top right), vector boson associated production (bottom left) and top associated production (bottom right).

In summary, it is unlikely to produce a light h_1 in any of the standard channels with discernible signal strength, which is why cascade decays of supersymmetric particles (SUSY cascades) are taken into consideration.

1.5.2 Production in SUSY Cascades

As discussed in the previous section, the standard production mechanisms do not yield a substantial rate of events where an h_1 is produced. However, h_1 couples considerably to neutralinos, and therefore could be copiously produced in cascade decays of squarks ($\tilde{q}$) and gluinos ($\tilde{g}$), the supersymmetric partners of quarks and gluons, respectively.

The production cross section of these strongly interacting SUSY particles mainly depends on their masses, i.e. the parameters M_{SUSY} (common parameter for mass of squarks ,$m_{\tilde{q}}$, and sleptons, $m_{\tilde{\ell}}$) and M_3 (gluino mass parameter), as can be seen from Fig. 1.6. In case of $M_3 = M_{\mathrm{SUSY}}$, the production of squarks of the first two generations dominates as mentioned above. Once these are produced, they subsequently decay into neutralinos and charginos. Because of the small Yukawa couplings of the squarks of the first two generations they decay preferably into the gaugino-like neutralinos. The latter fact puts constraints on the mass hierarchy of the neutralinos, which is accounted for in the above-mentioned modified P4 scenario. Since the singlet couples only to the Higgs doublets, the lightest CP-even Higgs boson h_1 couples due to its

large singlet component preferably to the higgsino-like neutralinos. Therefore, in order to produce the light Higgs boson in SUSY cascades at significant rates, one needs the gaugino-like neutralinos be heaviest, followed in decreasing order by the higgsino-like neutralinos and finally the singlino-like one, which represents the LSP. In order to achieve this hierarchy, one or both of the gaugino mass parameters M_1 and M_2 need to be larger than the higgsino mass parameter μ_{eff}. The dependence of the h_1 production on the gaugino mass parameters will be experimentally studied in this thesis using the decoupled squarks scenario. An example h_1-production diagram is shown in Fig. 1.7, where h_1 is produced and decays into a pair of b quarks.

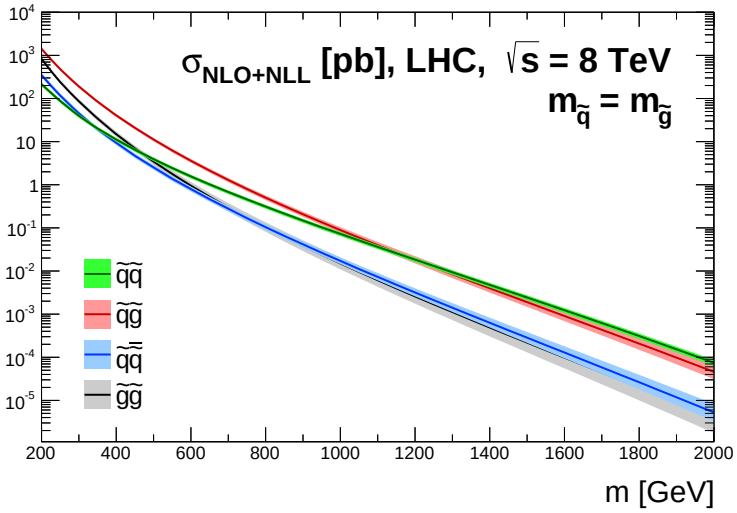

Figure 1.6: Production cross section of coloured SUSY particles at the LHC at $\sqrt{s} = 8\,\text{TeV}$ (calculated with NLL-FAST [100–106]).

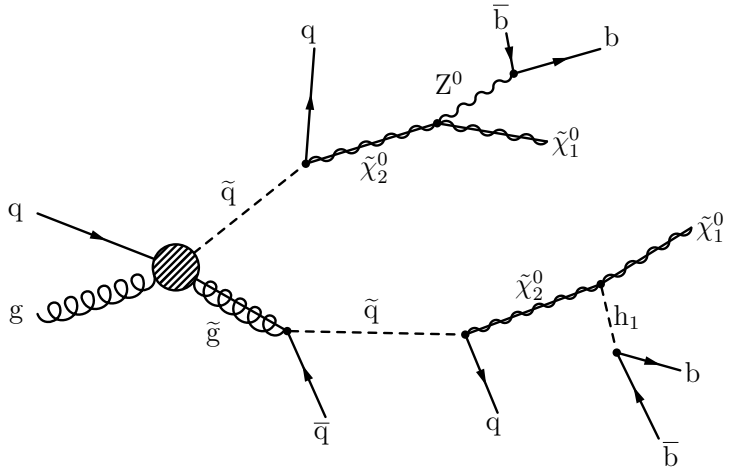

Figure 1.7: Feynman diagram of an example SUSY cascade with a neutralino decaying into an h_1, subsequently decaying into a pair of b quarks.

2 Phenomenological Studies of Light Higgs Scenarios

It is desirable to interpret experimental results as model-independent as possible. However, it is useful to investigate the implications of the measurements on the parameter space of theoretical frameworks. For instance, as pointed out in Section 1.3.1, the MSSM Higgs phenomenology at tree level has two free parameters, $\tan\beta$ and M_{A}. Therefore, model-independent results like, e.g. production cross section limits are frequently also translated into exclusion limits for these two quantities [107, 108]. Such model-specific interpretations can then be used to combine different experimental signatures.

The experimental analysis described in Chapter 5 of this thesis is motivated by the modified P4 scenario of the NMSSM, where light Higgs bosons are produced in supersymmetric cascade decays of coloured SUSY particles. This suggests to interpret the experimental results in terms of parameters which influence certain properties of a supersymmetric model. The production cross section of coloured SUSY particles is driven by their mass scale, i.e. M_{SUSY} in case of the squarks, and M_3 in case of the gluinos. The phase space available for the production of the light CP-even Higgs boson is mainly determined by the boson's mass m_{h_1}. However, light Higgs boson production in the NMSSM is not only driven by m_{h_1}, but also by the mass hierarchy of the neutralinos: The gaugino mass parameters M_1, M_2 and the higgsino mass term μ_{eff} determine the mass splitting between the neutralinos, thus influencing the phase space available for a particle produced in a neutralino decay and the respective branching fraction of a neutralino into this particle. The gaugino mass parameters also define the admixture of the gaugino and higgsino states in the neutralinos. The latter is important because in case of $M_{\mathrm{SUSY}} = M_3$ the predominantly produced coloured SUSY particles are squarks of the first two generations (see Section 1.5.2) which couple mainly to gaugino-like neutralinos [99]. However, the light Higgs boson with its dominant singlet admixture couples primarily to higgsino-like neutralinos leading to a three-step decay of the squarks into light Higgs bosons.

In this chapter, a study is presented investigating how experimental search results can be interpreted in a wide range of light-Higgs-boson scenarios in an appropriate way. Special interest is paid to the question, how well a *simplified model spectrum* (SMS) with its reduced particle content represents a full model with a sizeable production rate of light Higgs bosons like the NMSSM. In particular, it is investigated how well the two-step decay of the squarks in the SUSY cascades of the SMS approximates a full model where light Higgs bosons are mainly produced in three-step decays.

2.1 Interpretation Approaches

This section discusses two different approaches: Parameter scans within SMS and within the neutralino-chargino sector of the NMSSM.

The former is widely used to present the results of SUSY searches in a generic way [109–112]. The latter is a realistic and full model, where the production of light Higgs bosons is well motivated in a sizeable fraction of the parameter space. It also serves as a prototype for other models with singlet extensions like the E_6SSM [113].

2.1.1 Parameter Scan within a Simplified Model Spectrum

The hypothetical SMS presented in this section takes into account ideas of [114] and is designed as object of comparison for the purpose of the studies described in this chapter. For simplicity, only the production of gluinos is considered in this model. The most simple cascade decay from gluinos into a light Higgs boson h and the electroweak vector bosons Z^0, $W^\pm$ is cast into an SMS, which is denoted in this thesis as T5VH (see Fig. 2.1) following CMS terminology.

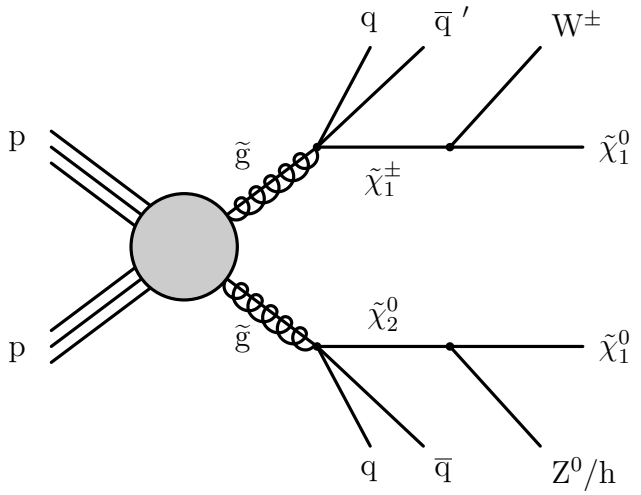

Figure 2.1: Signal event topology in the simplified model T5VH.

The parameters of this model are the light Higgs mass m_h, the gluino mass parameter M_3, and the common mass $m_{\tilde{\chi}}$ of the second lightest neutralino $\tilde{\chi}_2^0$ and the lightest chargino $\tilde{\chi}_1^\pm$. The mass of the lightest supersymmetric particle m_LSP is fixed at a constant value and is therefore no free parameter of the SMS. The branching fractions into each of the bosonic final states are fixed at one third each.

Although the sketch in Fig. 2.1 illustrates the topology of an SMS signal event, it does not serve as a Feynman graph of the decay process. It omits the decay of the gluino into a squark and a quark, which subsequently decays into a neutralino or chargino and another quark. This is only shown as an effective three body decay of the gluino, for which no vertex exists because the gluino only couples to coloured particles, while a neutralino or chargino is colourless. Therefore, in general every

SUSY cascade includes the decay chain from a squark to the LSP and additional particles.

2.1.2 Scan of the Neutralino and Chargino Sectors of the NMSSM

The neutralino and chargino sectors of the NMSSM are introduced in Section 1.4.1. Their structure influences the length of the squark-decay chain into a light Higgs boson. As pointed out above, the squarks of the first two generations preferentially decay via a three-step cascade (see Fig. 2.2, left) into an h_1 boson after emitting a particle X, e.g. a Z^0 boson, that couples to both, gauginos and higgsinos. However, the particle X is of minor interest in the presented study because only bosons produced in the final decay into an LSP are considered. This provides a fair comparison with the T5VH model of the previous section because in this SMS the bosons Z^0, $W^{\pm}$, and h are only produced in the final decay of the cascades.

Although the three-step case is preferred, a two-step decay might be a good approximation, if there is a strong mixing of higgsinos and gauginos, i.e. the neutralinos have sizeable admixtures from both type of particles. This is the case, if the mass parameters μ_{eff}, M_1, and M_2 are numerically similar, i.e. the higgsino-like and gaugino-like neutralinos are nearly mass-degenerate (see Fig. 2.2, right). This hypothesis is to be tested in the study presented in this chapter, taking also cases into account where the cascades involves charginos, resulting in W boson production.

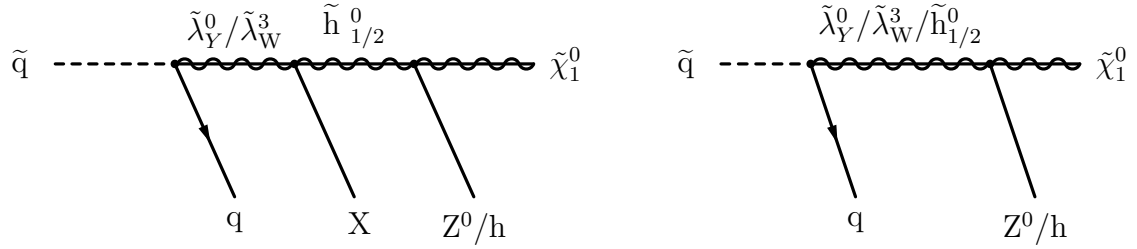

Figure 2.2: The three-step squark decay (left) and the two-step squark decay (right) include intermediate steps via gaugino-like ($\tilde{\lambda}^0_Y/\tilde{\lambda}^3_W$) and higgsino-like ($\tilde{h}^0_{1/2}$) neutralinos. The gaugino masses are driven by M_1, M_2, while the higgsino masses are mainly determined by μ_{eff}. If these parameters are numerically similar, the intermediate neutralino has a sizeable admixture of both after diagonalisation of (1.66). The particle X in the three-step cascade is of minor interest in the study presented in this chapter because for a fair comparison with the SMS from Section 2.1.1 only bosons produced in the final decay into an LSP are considered.

2.2 Implementation Details

In this section, the technical details of this study and the definition of relevant quantities are given. They serve as criteria to judge the applicability of the above introduced

SMS to light Higgs boson scenarios of the NMSSM. As base scenario the modified P4 scenario is used (see Section 1.4.2). The last coloured SUSY particle in a cascade decay is always a squark because gluinos couple only to coloured particles, thus not directly to neutralinos and charginos. Therefore, it is sufficient to consider only squarks as starting point of the decay chain. For the analysis in this chapter, a further restriction is applied by considering only the squarks of the first two generations because they are the dominant production mode of coloured SUSY particles in this scenario.

The investigated quantities are the branching fractions of the squarks of the first two generations into the three possible bosonic final states h_1, Z^0 and $\mathrm{W}^\pm$ which are represented in the SMS described in Section 2.1.1:

Two-step decays:

$$\mathcal{B}_2(\tilde{\mathrm{q}} \to \mathrm{h}_1) = \sum \mathcal{B}(\tilde{\mathrm{q}} \to \chi + \ldots \to \mathrm{LSP} + \mathrm{h}_1) \,, \tag{2.1a}$$
$$\mathcal{B}_2(\tilde{\mathrm{q}} \to \mathrm{Z}^0) = \sum \mathcal{B}(\tilde{\mathrm{q}} \to \chi + \ldots \to \mathrm{LSP} + \mathrm{Z}^0) \,, \tag{2.1b}$$
$$\mathcal{B}_2(\tilde{\mathrm{q}} \to \mathrm{W}^\pm) = \sum \mathcal{B}(\tilde{\mathrm{q}} \to \chi + \ldots \to \mathrm{LSP} + \mathrm{W}^\pm) \,, \tag{2.1c}$$
$$\mathcal{B}_2^{\mathrm{tot}} = \mathcal{B}_2(\tilde{\mathrm{q}} \to \mathrm{h}_1) + \mathcal{B}_2(\tilde{\mathrm{q}} \to \mathrm{Z}^0) + \mathcal{B}_2(\tilde{\mathrm{q}} \to \mathrm{W}^\pm) \,. \tag{2.1d}$$

Three-step decays:

$$\mathcal{B}_3(\tilde{\mathrm{q}} \to \mathrm{h}_1) = \sum \mathcal{B}(\tilde{\mathrm{q}} \to \chi_i + \ldots \to \chi_j + \ldots \to \mathrm{LSP} + \mathrm{h}_1) \,, \tag{2.2a}$$
$$\mathcal{B}_3(\tilde{\mathrm{q}} \to \mathrm{Z}^0) = \sum \mathcal{B}(\tilde{\mathrm{q}} \to \chi_i + \ldots \to \chi_j + \ldots \to \mathrm{LSP} + \mathrm{Z}^0) \,, \tag{2.2b}$$
$$\mathcal{B}_3(\tilde{\mathrm{q}} \to \mathrm{W}^\pm) = \sum \mathcal{B}(\tilde{\mathrm{q}} \to \chi_i + \ldots \to \chi_j + \ldots \to \mathrm{LSP} + \mathrm{W}^\pm) \,, \tag{2.2c}$$
$$\mathcal{B}_3^{\mathrm{tot}} = \mathcal{B}_3(\tilde{\mathrm{q}} \to \mathrm{h}_1) + \mathcal{B}_3(\tilde{\mathrm{q}} \to \mathrm{Z}^0) + \mathcal{B}_3(\tilde{\mathrm{q}} \to \mathrm{W}^\pm) \,. \tag{2.2d}$$

Here, χ represents a neutralino or chargino, provided the decay is kinematically allowed. The ellipses in (2.1a)–(2.1c) and (2.2a)–(2.2c) represent particles produced in addition to χ, which are irrelevant for this study. The branching fractions for the two-step and three-step decay are defined separately in order to analyse the probability of two- or three-step decays in randomly generated modifications of the base scenario. For these modifications values of the following parameters are drawn from a uniform distribution within the intervals

$$\mu_{\mathrm{eff}}: \quad [-500\,\mathrm{GeV}, -100\,\mathrm{GeV}] \,, \tag{2.3a}$$
$$M_1: \quad [100\,\mathrm{GeV}, 600\,\mathrm{GeV}] \,, \tag{2.3b}$$
$$M_2: \quad [100\,\mathrm{GeV}, 600\,\mathrm{GeV}] \,. \tag{2.3c}$$

For this purpose a workflow (see Fig. 2.3) using NMSSMTools 4.2.1 [115–121] at its core is implemented to calculate masses and branching fractions from a given set of parameters. The input is generated using `python`'s built-in `random` module, which uses the Mersenne Twister [122] to generate pseudo-random numbers. The ROOT framework [123] is used to store, analyse, and display the output from the

previous steps. `PySLHA 2.1.3` [124] is used to interface the above listed programs via the SUSY Les Houches Accord (SLHA) [125, 126]. This way 2.5 million parameter points with physical particle masses are generated.

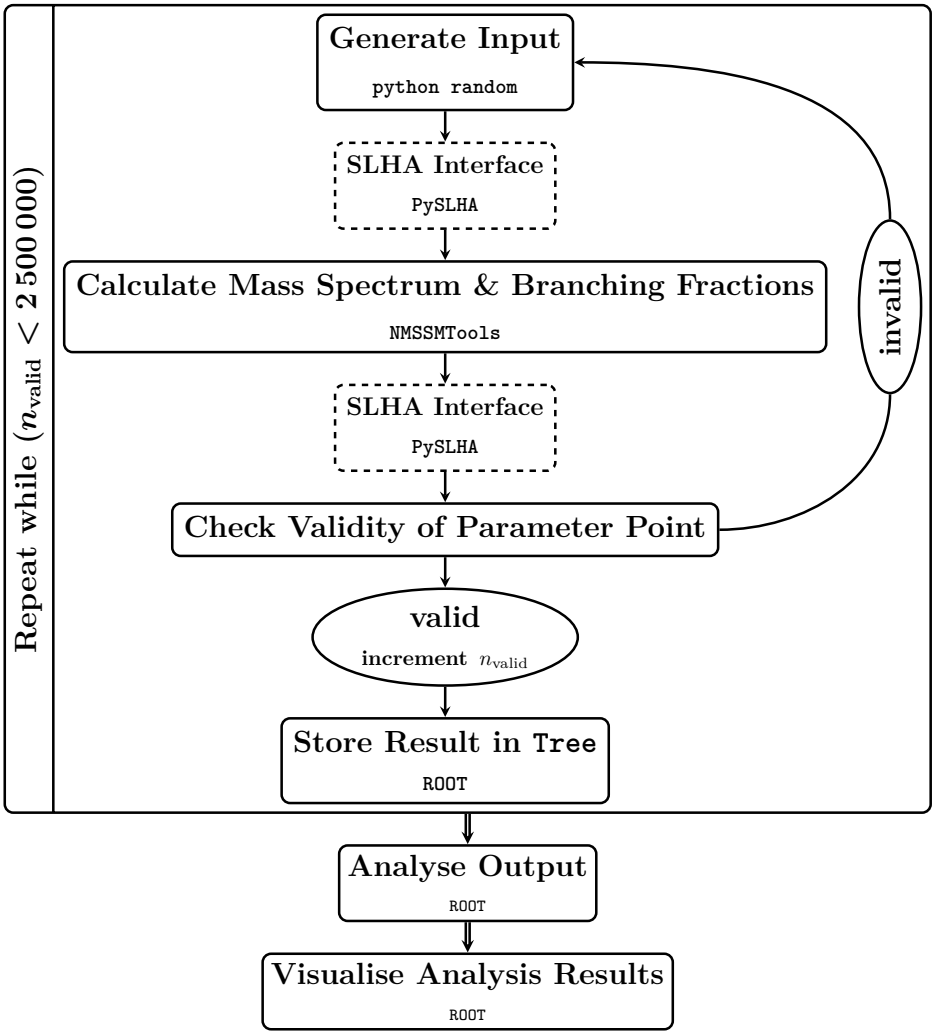

Figure 2.3: Flowchart of the NMSSM-parameter-space scan over 2.5 million valid parameter points. External programs or libraries are indicated in the respective steps.

2.3 Results and Conclusions

For each of the chosen parameter points, the quantities introduced in (2.1) and (2.2) are calculated. Figure 2.4 shows the branching fraction of two-step versus three-step squark decays for all considered bosons (Fig. 2.4, left) and for h_1 only (Fig. 2.4, right). The left plot demonstrates that the majority of possible models prefers the three-step decay into bosons ($\mathcal{B}_3^{\text{tot}} \approx 0.4$), while there is still a sizeable fraction for the two-step case ($\mathcal{B}_2^{\text{tot}} \approx 0.1$). The right plot confirms that the squarks predominantly decay in

three steps into h_1, which can be seen from the highly occupied region close to the abscissa.

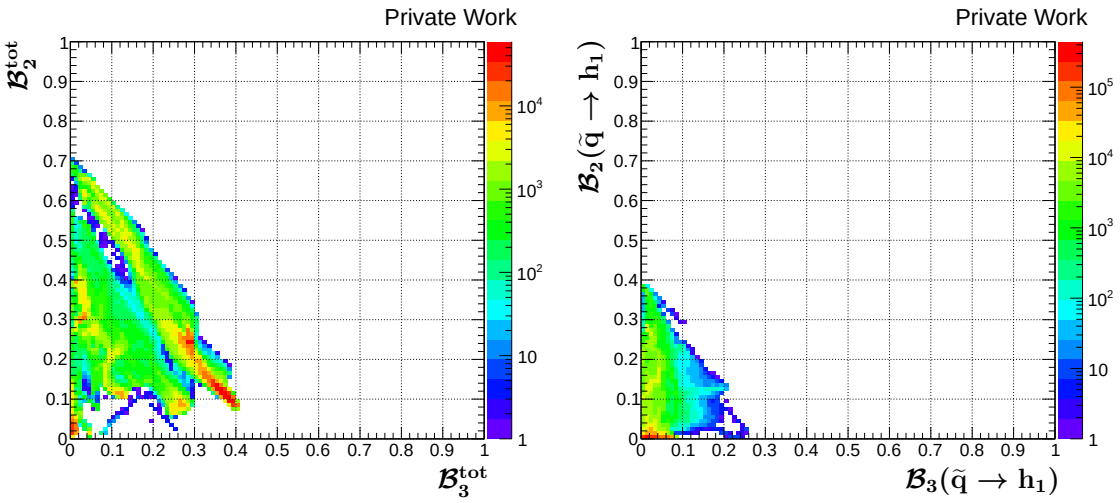

Figure 2.4: Distribution of the generated parameter points in the $\mathcal{B}_2^{\text{tot}}$–$\mathcal{B}_3^{\text{tot}}$ plane (left) and in the $\mathcal{B}_2(\tilde{q} \to h_1)$–$\mathcal{B}_3(\tilde{q} \to h_1)$ plane (right).

In Fig. 2.5, the differences between μ_{eff} and each of the gaugino mass parameters M_1 (Fig. 2.5, left) and M_2 (Fig. 2.5, right), defined as

$$\Delta\left(M_1, |\mu_{\text{eff}}|\right) = M_1 - |\mu_{\text{eff}}| \,, \tag{2.4a}$$

$$\Delta\left(M_2, |\mu_{\text{eff}}|\right) = M_2 - |\mu_{\text{eff}}| \,, \tag{2.4b}$$

are shown vs. the two-step branching fraction into h_1 bosons. Consistent with Fig. 2.4 (right), the majority of points is distributed close to the abscissa, i.e. at low values of $\mathcal{B}_2(\tilde{q} \to h_1)$. However, $\mathcal{B}_2(\tilde{q} \to h_1)$ increases in both cases to values up to 10 % or more in the interval from $-100\,\text{GeV}$ to $100\,\text{GeV}$. This supports the hypothesis stated in Section 2.1.2, that the approximation of the full NMSSM-light-Higgs-boson scenario with the two-step squark decay of the SMS improves in case of nearly mass-degenerate higgsinos and gauginos. In this case, these particles mix strongly, i.e. the neutralinos have a sizeable gaugino admixture thereby coupling to squarks, and a sizeable higgsino admixture providing a coupling to the h_1 boson. Thus, the intermediate decay-step from a gaugino-like neutralino to a higgsino-like neutralino is not needed. To further verify this statement an additional selection is applied on the parameter points. The plots in Fig. 2.6 contain only points where the requirement

$$\left|\Delta\left(M_1, |\mu_{\text{eff}}|\right)\right| < 100\,\text{GeV} \tag{2.5}$$

is fulfilled. The efficiency of this selection is 37.44 %, leading to enhanced maxima in the two-step regime with respect to Fig. 2.4. As a next step the additional constraint

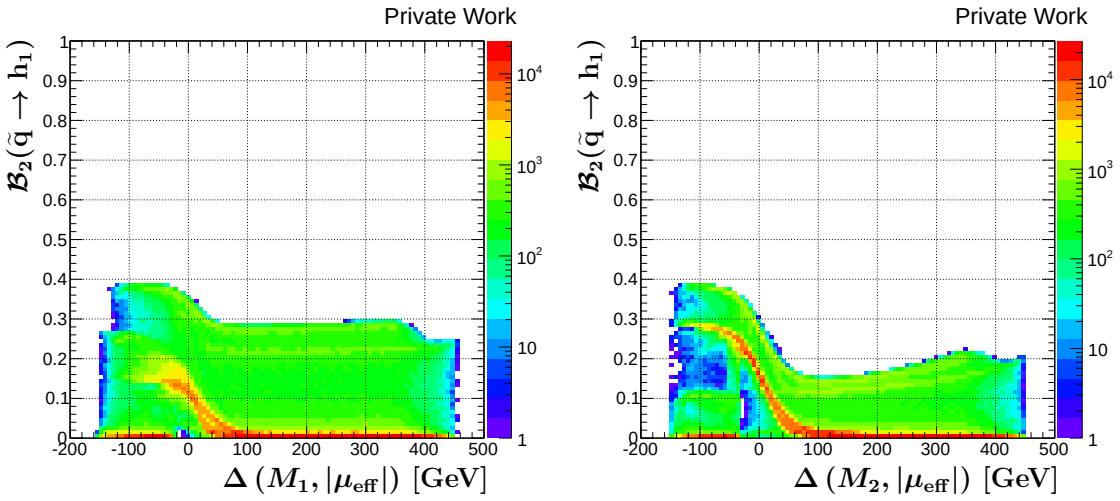

Figure 2.5: Distribution of the generated parameter points with respect to their two-step branching fractions into h_1 and the difference between $|\mu_{\text{eff}}|$ and the gaugino mass parameters M_1 (left) and M_2 (right). The two-step branching fractions into h_1 in general increase if the gaugino mass parameters are close to $|\mu_{\text{eff}}|$, i.e. in case of large mixing between the higgsinos and the respective gaugino.

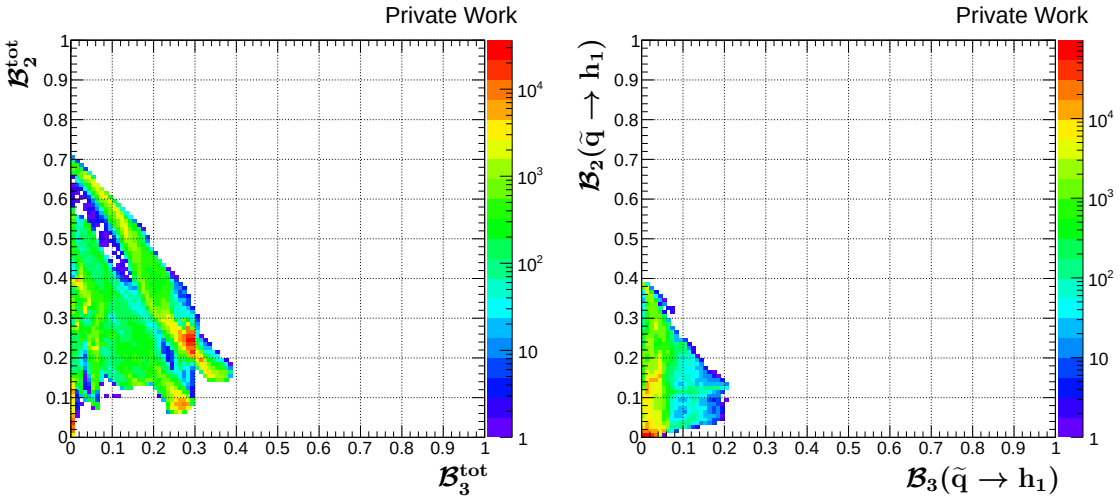

Figure 2.6: Distribution of the generated parameter points in the $\mathcal{B}_2^{\text{tot}}$–$\mathcal{B}_3^{\text{tot}}$ plane (left) and in the $\mathcal{B}_2(\tilde{q} \to h_1)$–$\mathcal{B}_3(\tilde{q} \to h_1)$ plane (right) fulfilling the requirement in (2.5) with an efficiency of $\epsilon = N_{\text{pass}}/N_{\text{total}} = 37.44\,\%$.

$$|\Delta\,(M_2, |\mu_{\text{eff}}|)| < 100\,\text{GeV} \tag{2.6}$$

is added with an efficiency of 14.10 % with respect to the unconstrained case. This results in further enhanced maxima in the two-step regime (see Fig. 2.7).

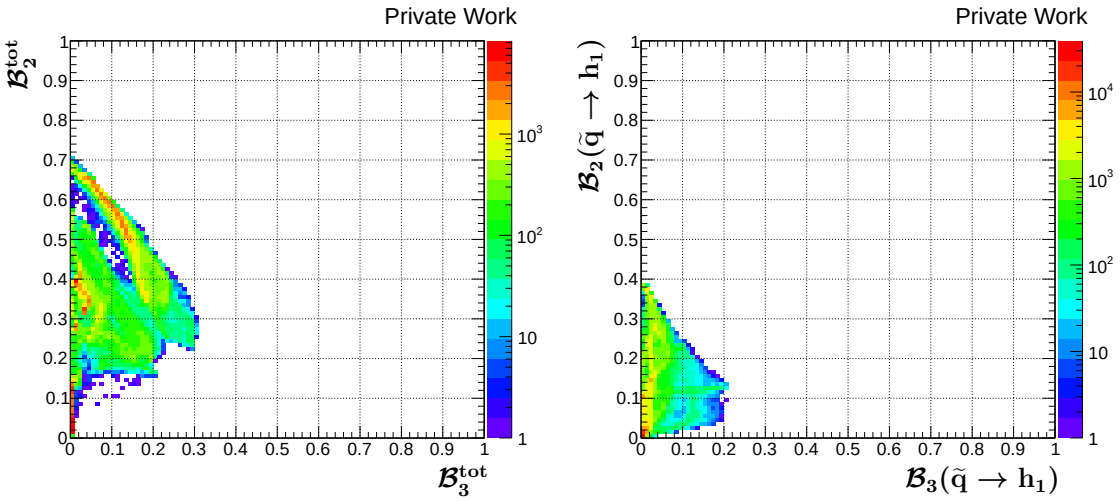

Figure 2.7: Distribution of the generated parameter points in the $\mathcal{B}_2^{\text{tot}}$–$\mathcal{B}_3^{\text{tot}}$ plane (left) and in the $\mathcal{B}_2(\tilde{q} \to h_1)$–$\mathcal{B}_3(\tilde{q} \to h_1)$ plane (right) fulfilling the requirements in (2.5) and (2.6) with an efficiency of $\epsilon = N_{\text{pass}}/N_{\text{total}} = 14.10\,\%$.

Although the left plots in Figs. 2.4, 2.6, 2.7 suggest that the two-step SMS covers partly the NMSSM parameter space, an interpretation within the NMSSM is favoured in the view of a light-Higgs-boson analysis. Figure 2.8 shows the two-step branching fraction versus the total branching fraction into h_1 ($\mathcal{B}(\tilde{q} \to h_1)$). The latter is defined as the sum of two-step, three-step decays, and h_1 production in intermediate steps of the cascade, i.e. not necessarily together with the LSP. The latter contribution is ignored in the SMS, but in the NMSSM it provides a substantial contribution. This can be seen from a comparison of Fig. 2.4 (right), where the majority of points is at values below 5 % for the three-step case and even lower for the two-step case. However, in Fig. 2.8, $\mathcal{B}(\tilde{q} \to h_1)$ has a high density of points up to 30 %. From the populated area close to the diagonal in Fig. 2.8 one can conclude, that there is indeed a part of the NMSSM parameter space where the two-step branching fraction is the dominant mode for h_1 production, but the vast majority of parameter points has low values of $\mathcal{B}_2(\tilde{q} \to h_1)$.

In summary, not only three-step decays would be neglected by the SMS, but also h_1 production in intermediate decay steps. Therefore, an SMS with two-step squark decays would by far underestimate h_1 production in case of NMSSM scenarios where squark production of the first two generations is the dominant production mode for SUSY cascades, as is the case for $M_{\text{SUSY}} = M_3$.

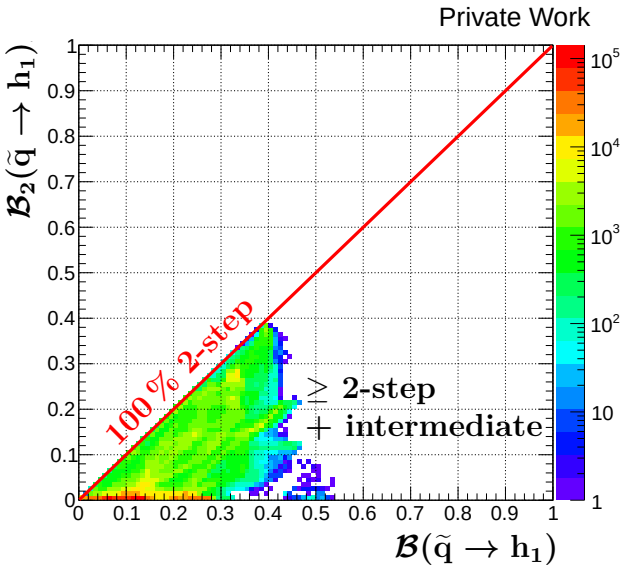

Figure 2.8: Distribution of the generated parameter points in the $\mathcal{B}_2(\tilde{q} \to h_1)$–$\mathcal{B}(\tilde{q} \to h_1)$ plane. In the area close to the diagonal the two-step squark decay is the dominating h_1 production mode. However, the majority of parameter points is at low values of $\mathcal{B}_2(\tilde{q} \to h_1)$. In general there is a sizeable fraction of squark decays in which an h_1 is produced, e.g. after three or more steps or in intermediate decay steps.

3 The CMS Experiment at the LHC

This chapter summarises the experimental apparatus employed to acquire the data used in the Higgs-boson search of this thesis. After a brief description of the particle-accelerator complex and the Compact Muon Solenoid (CMS) experiment, the methods and tools for event generation and reconstruction are presented.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [21] is a superconducting synchrotron accelerator and storage ring built to explore physics at the TeV scale. It is operated by the *European Organization for Nuclear Research* (CERN) near Geneva, and was built in a 26.7 km circular tunnel that was constructed between 1984 and 1989 for the Large Electron-Positron Collider (LEP). The LHC is linked via two transfer tunnels with the other part of the CERN accelerator complex (see Fig. 3.1) that is used as injector of protons or lead-nuclei.

The LHC-design is chosen such that a wide range of physics topics is covered. In the proton-proton mode, Higgs-physics is a prominent application along with searches for supersymmetric particles and extra dimensions. It also serves as laboratory to study top-quark properties and SM processes in general. In addition, the lead-lead mode of the LHC allows the investigation of the quark-gluon plasma.

The particle beams collide in four interaction points where the four main experiments are situated (see Fig. 3.1). Two of them are multipurpose experiments which address a wide range of physics cases including the above-mentioned topics. One is A Toroidal LHC Apparatus (ATLAS) [128] and the other is the CMS experiment [22] which recorded the data used for this analysis. A Large Ion Collider Experiment (ALICE) [129] and LHC beauty (LHCb) [130] are built for specific measurements. LHCb targets heavy-flavour physics, in particular CP-violating and rare decays. ALICE uses lead-lead collisions to gain insight into the origin of confinement and the mechanism of mass generation. Two further experiments probing small-x QCD are situated close to ATLAS and CMS: LHC forward (LHCf) [131] and Total Elastic and Diffractive Cross Section Measurement (TOTEM) [132].

The instantaneous luminosity, which is a measure of the performance of an collider, is given for the LHC by

$$\mathcal{L} = f \cdot \frac{k_B N_p^2}{a} \,. \tag{3.1}$$

where k_B denotes the number of bunches per beam, N_p the number of protons per bunch and f the bunch revolution frequency. The factor a describes the geometric size of the bunch in the plane transverse to the proton beam direction [133]. The

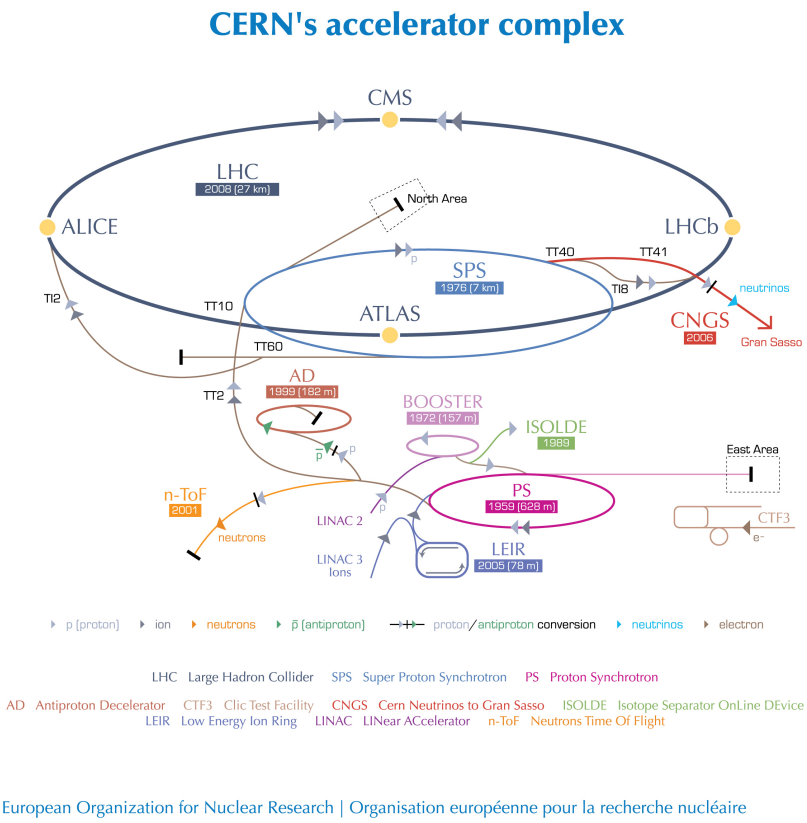

Figure 3.1: Schematic drawing of the CERN accelerator facility. The LHC is shown with its pre-accelerator complex [127].

time duration of $\approx 93\,\mathrm{s}$, corresponding to 2^{20} LHC orbits, is referred to as lumi section [134]. This time interval is of importance because the quality of the data recorded by the CMS experiment is evaluated with the granularity of one lumi section. The number of events for a given process within the time interval T is proportional to the *integrated luminosity*

$$L = \int_0^T \mathrm{d}t\mathcal{L} \,. \tag{3.2}$$

Therefore, the design of the LHC aims at a high luminosity to provide a high rate of interesting but rare processes.

The LHC started to record data in the end of 2009 with centre-of-mass energies of $\sqrt{s} = 0.9 - 2.36\,\mathrm{TeV}$. In March 2010 the energy was increased to $\sqrt{s} = 7\,\mathrm{TeV}$ which was kept till the end of 2011, recording an integrated luminosity of $40.8\,\mathrm{pb}^{-1}$ and $5.6\,\mathrm{fb}^{-1}$ in 2010 and 2011, respectively [135]. In 2012 data was taken at $\sqrt{s} = 8\,\mathrm{TeV}$ and an integrated luminosity of $21.8\,\mathrm{fb}^{-1}$ was recorded by CMS (see Fig. 3.2). The complete 2012 dataset was used for the analysis described in this thesis.

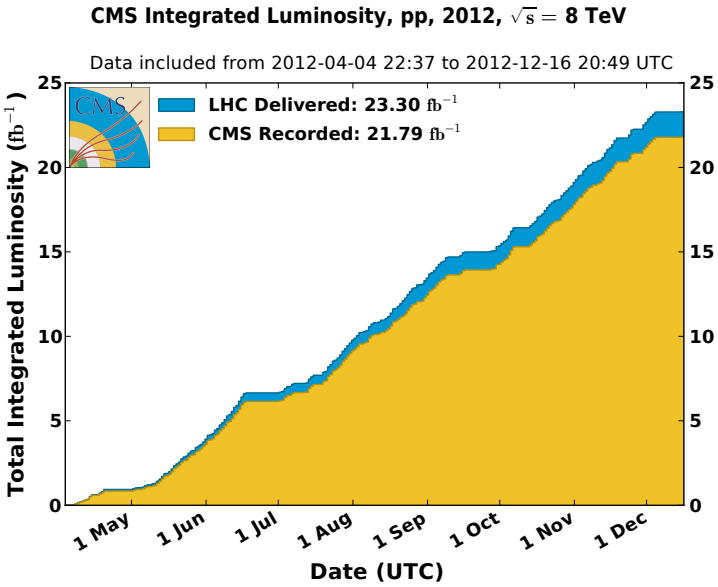

Figure 3.2: Integrated luminosity delivered by the LHC (upper line) and recorded by CMS (lower line) for the data-taking period 2012 [135].

3.2 The Compact Muon Solenoid

The CMS experiment is one of the multipurpose experiments at the LHC. It has the typical onion-like structure common for colliding-beam detectors in the field of high energy physics (see Fig. 3.3). The different layers consist of subdetectors designed to measure different properties of the particles created in hadron collisions.

The positions and dimensions of the subdetectors are given with respect to the coordinate systems used within CMS. The cartesian coordinates are defined with respect to the nominal interaction point of the proton beams. The x axis points to the LHC ring centre, while the y axis points to the surface of the earth. Thus, the z axis of the right-handed system is oriented in the anti-clockwise beam direction. Due to the cylindrical design of the detector, cylindrical coordinates are used in addition to the cartesian system. The radial distance r is defined in the x-y plane, where an azimuthal angle $\phi = 0$ corresponds to the x axis and the polar angle θ is defined with respect to the z axis. In high-energy physics, it is convenient to recast θ into the *pseudorapidity* η, defined as

$$\eta = -\ln\left\{\tan\left(\frac{\theta}{2}\right)\right\} \ . \tag{3.3}$$

In the relativistic limit, applicable to collisions at high centre-of-mass energies, this quantity equals the rapidity. The advantage of this observable is that rapidity differences are invariant under Lorentz transformations in z direction. It is therefore independent of the reference frame used to describe the recorded events. The following sections will go into the details of the subdetector systems and are based on information given in [133].

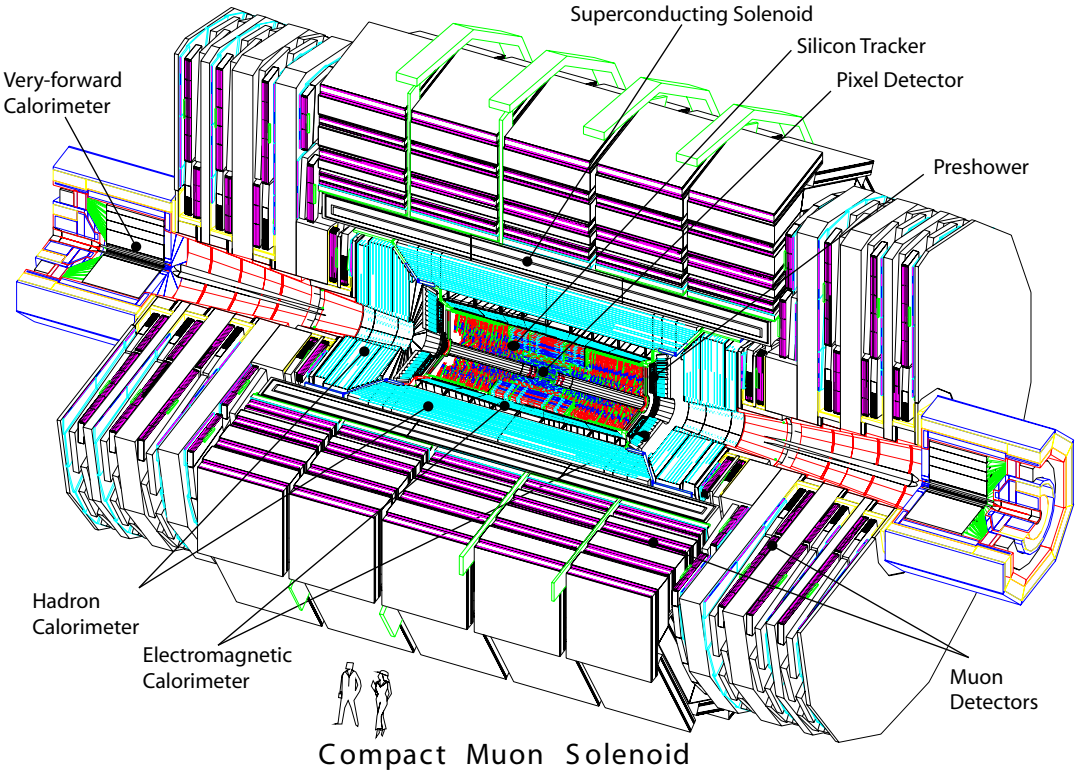

Figure 3.3: Three-dimensional sketch of the CMS detector. Taken from [133].

3.2.1 Tracker

The tracking system [22,133,136,137] is the innermost subdetector of CMS. It detects the positions of charged particles traversing through its layers and measures the curvature of the trajectories caused by the magnetic field of the solenoid (see Fig. 3.3). A schematic drawing is shown in Fig. 3.4. The tracker consists of a silicon pixel and a silicon strip detector part. The former is situated directly around the beam pipe and is divided into a barrel with three layers at radii of 4.4, 7.3 and 10.2 cm, and two endcaps. It provides three-dimensional (3-D) position measurements of the hits with a resolution of approximately 10 µm in the transverse coordinate and 20 – 40 µm in the longitudinal coordinate, while the third coordinate is given by the radial sensor-plane position. In total, its 1440 modules cover an area of about 1 m^2 and have 66 million pixels.

The strip detector has 15 148 silicon modules, which in total have 9.3 million strips covering an active area of about 198 m^2. It occupies the radial region between 20 and 116 cm and is composed of three different subsystems. The Tracker Inner Barrel and Disks (TIB/TID) extend to a radius of 55 cm and consist of four barrel layers and three disks at each end. These provide position measurements in r-ϕ with a resolution of approximately 13 – 38 µm. The TIB/TID system is surrounded by the Tracker Outer Barrel (TOB) consisting of six layers. It extends in r to 116 cm and in z between ±118 cm. Beyond this range the Tracker EndCaps (TEC) cover the region

with 124 cm $< |z| <$ 282 cm and 22.5 cm $< |r| <$ 113.5 cm. Each TEC is composed of nine disks. TOB and TEC provide position measurements in r-ϕ with a resolution of approximately 18 – 47 µm.

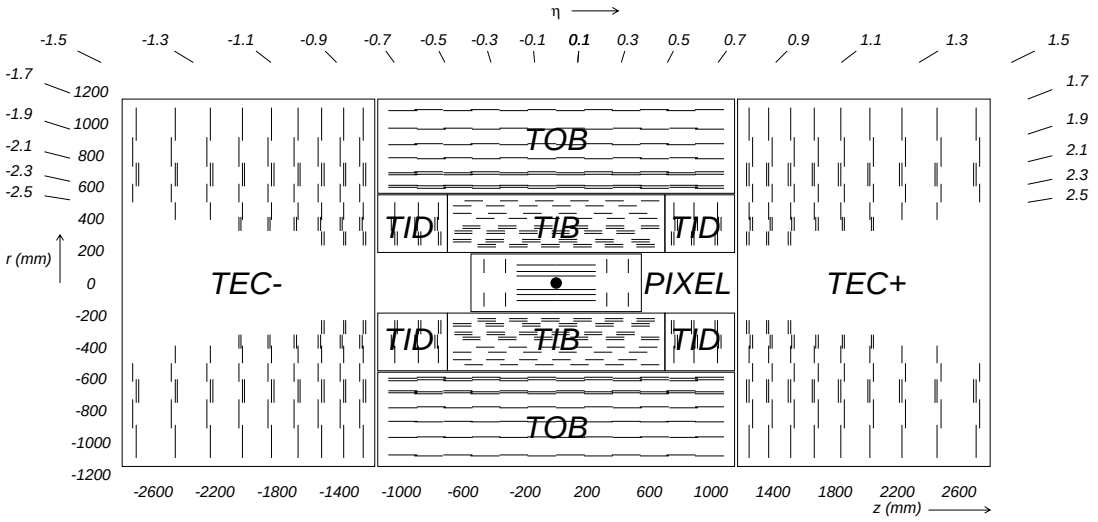

Figure 3.4: Layout of the CMS Tracker. Each line represents a detector module. Double lines represent back-to-back modules, yielding stereo hits [22].

The very good position resolution of the tracking system facilitates the precise reconstruction of primary vertices, whose position resolution approaches 10 µm in x/y and 12 µm in z in jet-enriched samples using at least 50 tracks [136]. Excellent tracking is also important for the reconstruction of secondary vertices, which are typical for decays of bottom hadrons and are an important ingredient for the identification of b-quark jets (b tagging).

3.2.2 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) encloses the inner tracker (see Fig. 3.3). It uses scintillating lead tungstate ($PbWO_4$) crystals covering a pseudorapidity range up to $|\eta| < 3.0$ (see Fig. 3.5) and is subdivided into a barrel part (EB) and endcaps (EE) at each end. The EB consists of 61 200 crystals and the EE contain 7324 crystals each. The dimensions of the crystals are $\Delta\eta \times \Delta\phi = 0.0174 \times 0.0174$. A preshower detector (ES) is placed in front of the endcap crystals, to augment prompt-photon identification by rejecting π^0 mesons. The $PbWO_4$ properties, like high density (8.28 g/cm^3), short radiation length ($X_0 = 0.89$ cm) and small Molière radius ($R_M = 2.2$ cm), allow a compact construction with fine granularity.

In the EB the crystals are arranged symmetrically to the interaction point. They provide 360-fold granularity in ϕ and 85-fold granularity in the positive and negative pseudorapidity range up to $|\eta| = 1.48$. The EE extends the coverage to $|\eta| = 3.0$, and the fiducial area of the ES is approximately $1.65 < |\eta| < 2.6$.

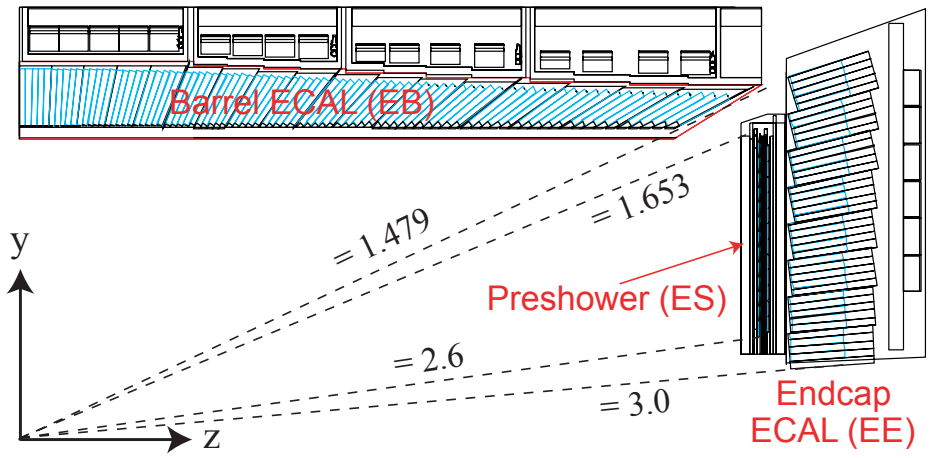

Figure 3.5: One quarter of the electromagnetic calorimeter in r-z view. Taken from [133].

The depth of the material corresponds to a radiation length of $25.8\,X_0$, thus well containing the electromagnetic showers, and enabling electron and photon identification. Their energy is measured by their scintillation light which is detected by avalanche photo-diodes (APD) [138, 139] in the EB and by vacuum photo-triodes (VPT) [140] in the EE. The crystal light yield and the APD gain are temperature dependent ($-2\,\%/°\mathrm{C}$), therefore demanding stable temperatures within $0.05\,°\mathrm{C}$ in the EB and within $0.1\,°\mathrm{C}$ in the EE. This goal is met by the CMS ECAL [141], thus leading to no degradation of the energy resolution with respect to the design specification. The ECAL barrel energy resolution σ_E for electrons in test beams is measured to be [142]

$$\frac{\sigma_E}{E} = \frac{2.8\,\%}{\sqrt{E(\mathrm{GeV})}} \oplus \frac{12\,\%}{E(\mathrm{GeV})} \oplus 0.3\,\%\,, \tag{3.4}$$

where the energy E is given in GeV. The symbol $\oplus$ represents quadratic addition of the individual uncertainties. The three contributions are the stochastic term, the noise term and a constant term. The constant term arises from energy leakage, non-uniformity of the light collection, and calibration uncertainties.

Combining ECAL and tracker information, electrons and photons can be distinguished up to a pseudorapidity of $|\eta| = 2.5$, i.e. the limit of the tracker coverage.

3.2.3 Hadronic Calorimeter

The ECAL is enclosed by the hadronic calorimeter (HCAL), which consists of four sections (see Fig. 3.6). The HCAL barrel (HB) and endcaps (HE) cover a region of $0 < |\eta| < 1.3$ and $1.3 < |\eta| < 3.0$, respectively. Both are sampling calorimeters with alternating layers of brass absorber and plastic scintillator tiles. The scintillation light is detected by hybrid photo-diodes (HPD) that can operate in high axial magnetic fields. The outer HCAL (HO) is made of plastic scintillator placed outside

the solenoid magnet, using the coil as absorber, also with HPD readout. The HO is designed as a tail catcher in the barrel region to ensure that hadronic showers are sampled with up to eleven hadronic interaction lengths. The forward section of the HCAL (HF) is an especially radiation-hard component made of iron absorbers and quartz fibres as active material. It is placed at $|z| = 11\,\mathrm{m}$ and covers the range $2.9 < |\eta| < 5.0$.

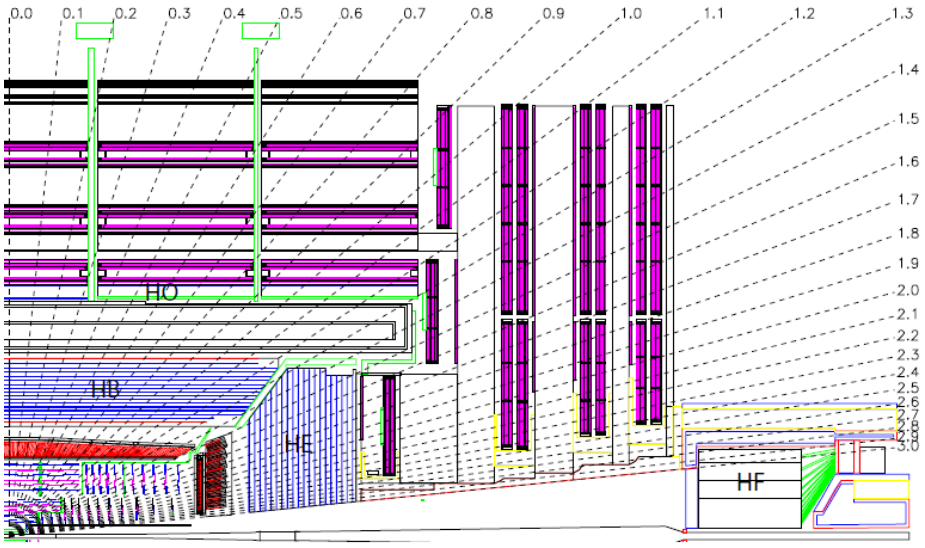

Figure 3.6: One quarter of the hadronic calorimeters in r-z view. Taken from [22].

Besides the identification of hadronic particles and jets, the large pseudorapidity coverage is needed to detect imbalances in the transverse energy. This would be an indication for particles escaping detection like neutrinos or potential unknown particles that do not interact with the detector. The jet-energy resolution obtained from HCAL measurements combined with the ECAL is

$$\frac{\sigma_E}{E} \approx \frac{100\,\%}{\sqrt{E(\mathrm{GeV})}} \oplus 5\,\%\,, \tag{3.5}$$

where the energy E is given in GeV [143]. The symbol $\oplus$ represents quadratic addition of the individual uncertainties. The actual energy resolution of jets used in CMS physics analyses is improved with respect to (3.5) by adding tracking information (see Section 3.4.2).

3.2.4 Muon System

The outermost part of the CMS detector is the muon system. Since many signatures of interesting physics processes feature muons, it is a very powerful tool to discriminate these processes from the high background rate at the LHC. In general, the muon system is used to identify muons and to measure their momentum in combination with tracker information. The high magnetic field of the solenoidal magnet and its

return yoke enable good momentum resolution and make it a key ingredient of the trigger system (see Section 3.2.5).

Muon reconstruction in CMS is done using the inner tracking system, and with up to four stations of gaseous detectors situated outside the solenoid and interspersed with steel layers of the magnet's flux-return yoke, which also serve as hadron absorber. The pseudorapidity coverage of the muon system extends to $|\eta| = 2.4$ and thereby assists the tracker in reconstructing the momentum and charge of muons with high transverse momentum p_T.

The muon system has a total area of detection planes of $25\,000\,m^2$, thus the used technology needs to be inexpensive, but also reliable and robust. Therefore, a combination of different detector systems is used to meet these design goals. Drift tube (DT) chambers and cathode strip chambers (CSC) are used in $|\eta| < 1.2$ and $0.9 < |\eta| < 2.4$, respectively. These two types of muon detectors are complemented by a system of resistive plate chambers (RPC) covering the pseudorapidity range $|\eta| < 1.6$ (see Fig. 3.7) [144]. The basic element of the DT chambers is the drift cell with a size of $42 \times 13\,mm^2$. The gas mixture of 85 % Ar and 15 % CO_2 provides good quenching properties and a saturated drift velocity of about 55 µm/ns with a maximum drift time of up to 380 ns. The CSCs consist of six layers and are standard multi-wire proportional counters (MWPC). The cathode strips are oriented in radial direction, thus providing a precise measurement in r-ϕ and a coarse measurement in the radial direction. This detector type has a fast response and a high radiation resistance, thus it is well suited for the high muon flux in the covered pseudorapidity region. The gas mixture used in all chambers consists of 50 % CO_2, 40 % Ar, and 10 % CF_4. The RPCs are double-gap chambers, operated in avalanche mode to ensure reliable operation at high rates, and provide an independent and fast trigger.

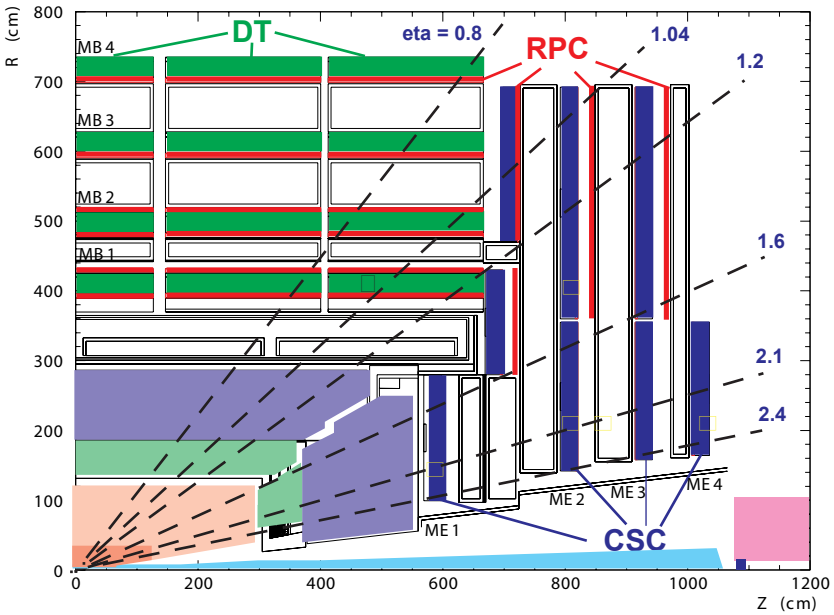

Figure 3.7: One quarter of the muon system in r-z view. Taken from [133].

The spatial resolution per chamber in the first LHC run was 80 – 120 µm in the DTs, 40 – 150 µm in the CSCs and 0 – 1.2 cm in the RPCs. The time resolution was for all three subsystems 3 ns or better. The identification efficiency measured using the tag-and-probe technique with $Z^0 \rightarrow \mu\mu$ decays [145] is typically greater than 90 % and the simulation models well the data. Hadron-to-muon misidentification probabilities are a few per mille for pions, which are dominated by decays in flight, and well below one per mille for protons, which are dominated by punch-through and random matching [146]. Data and simulation are in good agreement. Further details on the muon system and its performance can be found in [147].

3.2.5 Trigger and Data Acquisition System

Proton-proton and heavy ion collisions are provided by the LHC at high interaction rates of up to 40 MHz. Several collisions occur per crossing of the proton bunches resulting in an event data size of about 1 MByte. Hence, it is impossible to store and process all the data delivered by the LHC and a pre-selection of interesting events is necessary. The trigger system of CMS reduces the event rate in a two-step approach: The level-1 trigger (L1T) filters events in less than 4 µs/event [148] with an output rate of around 100 kHz followed by the high-level trigger (HLT) selecting around 400 Hz for storage. The pre-selection is configured via trigger menus, which are adjusted to the instantaneous luminosity in order to achieve the maximum output. Hence, trigger menus need to be updated continuously using either tighter selections or pre-scaling, which means that the trigger does not accept every event, but only every Nth event, where N is the pre-scale factor.

For maximum flexibility, the L1T is implemented in Field Programmable Gate Arrays (FPGA) where possible, and Application Specific Integrated Circuits (ASIC). The L1T decision is based on low-granularity calorimeter and muon-system information which is organised in local, regional and global components (see Fig. 3.8).

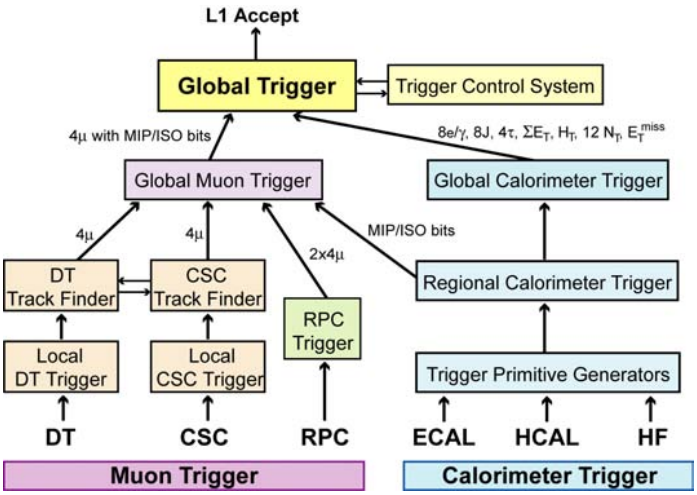

Figure 3.8: Architecture of the level-1 trigger (L1T) system. Taken from [22].

The HLT step is purely software-based, employing a computing farm. It performs a fast event reconstruction using the full detector information, including the inner tracking system, to decide, whether an event is discarded or kept for storage and full event reconstruction.

3.2.6 Data Quality Monitoring

Data Quality Monitoring (DQM) [149] is an important part of the CMS data processing. The data is scanned for problems in the operation of the detector and certified in different stages. The DQM runs online with a parallel data stream to check the status of the subdetectors and to monitor different physics objects, hence allowing fast feedback on the running conditions of CMS. DQM is used after the full offline reconstruction to certify that all subsystems worked properly, thus ensuring good quality for each data reconstruction campaign with a granularity of one lumi section.

This work uses only data centrally certified as good for physics analyses, which means that LHC provides stable beams and all CMS components work properly during the respective lumi section.

3.3 Event Simulation in CMS

The simulation of physics processes and the detector response play a key role in particle physics analyses. They allow the prediction of the processes occurring in particle collisions, and their measurements within the detector. Thus, information from simulation is frequently used for detector calibration and the prediction of signal processes and their backgrounds. The latter is essential for tuning the procedure and understanding the sensitivity of the analysis, but also for estimating the requirements of future experiments. The methods and software tools important for this thesis are briefly summarised in the following.

3.3.1 Monte Carlo Method

The pp collisions at the LHC are complex processes due to the composite internal structure of hadrons and the high particle multiplicity in the final state (see Fig. 3.9). These properties cannot be fully described analytically from first principles. Thus, a stochastic technique, referred to as Monte Carlo (MC), is used, i.e. results are numerically derived by repeated random sampling.

3.3.2 Event Generation

Based on the factorisation theorem [150], one can divide the event generation procedure into subsequent separate processes. These individual steps are illustrated in the event in Fig. 3.9 which shows the production of a Higgs boson in association with a t-quark pair. First, the matrix element of the hard process, depicted as red spot in Fig. 3.9, is calculated at a fixed order of the strong coupling α_s. The momenta of

the incoming partons are randomly sampled from the input PDFs (see Section 1.1.1) and the outgoing partons are randomly distributed in the available phase space. However, the factorisation approach enforces a redefinition of the PDFs because the partonic cross sections exhibit collinear divergences connected to soft long-distance interactions. These divergences are factored out and absorbed in the PDFs thereby introducing an arbitrary *factorisation scale* that separates soft long-distance physics from the hard short-distance process. The factorisation scale is typically set to the *renormalisation scale* at which α_s is computed. This scale is usually chosen equal to the mass or transverse momentum of the final-state system.

After calculating the hard process, higher-order QCD-effects, visualised as blue curly and straight lines in Fig. 3.9, are taken into account using parton-shower models. In these models, initial-state radiation (ISR) and final-state radiation (FSR) of the partons are added to the event as successive parton emissions from the hard-interaction scale down to the hadronisation scale at $\sim 1\,\text{GeV}$.

The final step is the hadronisation of the partons, illustrated as light green ovals in Fig. 3.9, for which different models are used [151–153]. These are purely phenomenological and are tuned to match the data because the hadronisation occurs at scales $\sim \Lambda_{\text{QCD}}$ where perturbative calculations are invalid. At this stage, additional multi-parton interactions (MPI) are taken into account, since these are characterised by small momentum transfers and therefore depend on the hadronisation description. These are represented in Fig. 3.9 by the purple spot.

In the following, the event generators employed to simulate the samples used in this thesis are briefly described.

PYTHIA

PYTHIA [155] is a multi-purpose event generator, widely used in high-energy particle-physics. It provides all generation steps described above and can simulate all relevant initial states and all Standard-Model processes as well as processes in theories beyond the Standard Model. The hard-scattering calculation is limited to tree-level $2 \to 1$ and $2 \to 2$ processes. Higher orders are approximated with parton showers and the hadronisation is based on a string model [151]. The version used throughout this thesis is PYTHIA 6.4 using the parameter settings of the Z2^* tune [156–158]. The interface of PYTHIA allows the communication with other programs, e.g. mass-spectrum generators, via the SLHA interface [125, 126].

MadGraph

MadGraph [159, 160] calculates matrix elements on tree level to arbitrary order. Different from PYTHIA, the hard-gluon radiation in ISR and FSR is also calculated on matrix element level. In order to avoid divergent soft-gluon radiation, a minimum p_{T} threshold is set. This method leads to a precise description of the event topology, but the cross section exhibits a strong scale dependence. Hence, it is normalised to higher-order predictions from MCFM [161] for SM processes and PROSPINO or NLL-FAST [100–106] for SUSY processes. Parton showering and hadronisation are not

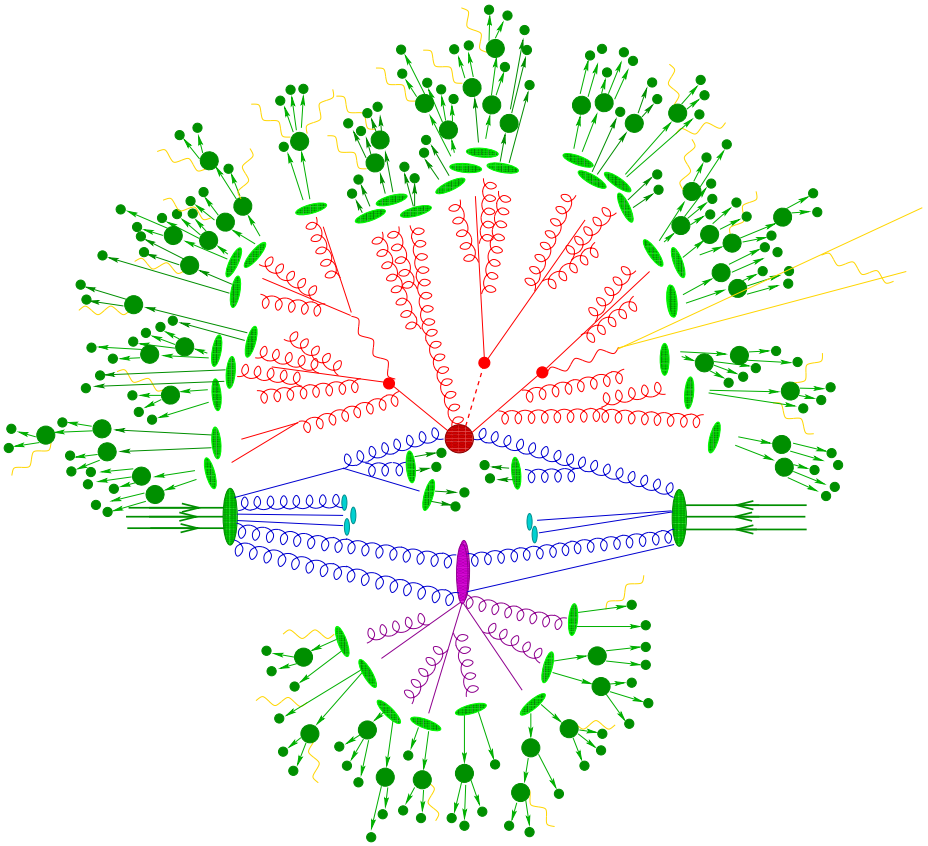

Figure 3.9: Illustration of a $pp \to t\bar{t}h + X$ event [154]. The hard process is shown as red spot, and the purple spot depicts the underlying event. ISR and FSR are illustrated as straight and curly lines. Hadrons generated during hadronisation have a light green colour, while the final hadrons are dark green.

implemented in MadGraph, but pythia is interfaced to fulfil this task.

POWHEG

powheg [162–164] is optimised for heavy-quark production in hadronic collisions. This generator calculates the hard scattering at NLO, whereas parton showering and hadronisation are delegated to pythia in this thesis.

TAUOLA

tauola [165] is a dedicated package for the τ-decay simulation. It takes spin information and QED corrections into account.

3.3.3 Detector Simulation

Event generators model only the steps depicted in Fig. 3.9, however further decays or interactions with the detector material are simulated in a subsequent step. This task is performed with two approaches in CMS. A precise MC-based simulation is

performed using the GEometry ANd Tracking (GEANT4) software [166,167], which is interfaced to the CMS software. The full CMS geometry including support structures and active material with signal readout, noise, and cross-talk is simulated, while taking into account the magnetic-field effects on the detector response. Finally, a trigger emulation is added, and the output has the same format as real collision events. The detector simulation performs the transition from generator level to detector level, thus facilitating a direct comparison of simulation and measurement. As for the event generators, a tuning of the detector-simulation parameters is required in order to fully match the data.

The computing resources required for the detailed detector simulation described above constitute the major part of the event simulation. Therefore, CMS developed a fast simulation of the detector [168], that uses simplifying assumptions, parameterisations and optimised reconstruction algorithms. The fast simulation is tuned to the GEANT4-based simulation, and its output is based on the same data format as the reconstruction of real collision and fully simulated events.

3.4 Event Reconstruction in CMS

Physics analyses performed with CMS data involve the identification of particles emerging from the hadron collisions in the centre of the detector. These particles leave typical signatures in the subdetectors, thus enabling their reconstruction. This is described in this section, which concludes with a definition of the physics objects used in this thesis.

3.4.1 Particle Signatures

The typical signatures of the particles produced in the CMS detector with energies above a few 100 MeV are illustrated in Fig. 3.10.

Photons

Photons mainly cause electromagnetic showers in the ECAL. Due to the high ECAL depth, the photon energy is completely contained in the ECAL, distributed longitudinally over a few crystals. In most cases, more than 94 % of the energy is deposited within a 3×3 matrix of crystals [170]. Depending on the pseudorapidity, between 20 and 60 % of the photons are not directly measured in the ECAL because they converted already into e^+e^- pairs while traversing the inner tracking system, which has a depth of $0.4-1.8\,X_0$ [22].

Electrons

Electrons leave the majority of their energy in the ECAL via electromagnetic showers, but some energy is lost prior via bremsstrahlung while traversing the tracking detector. Because of their electric charge, electrons are bent in the magnetic field and induce a signal in the tracker.

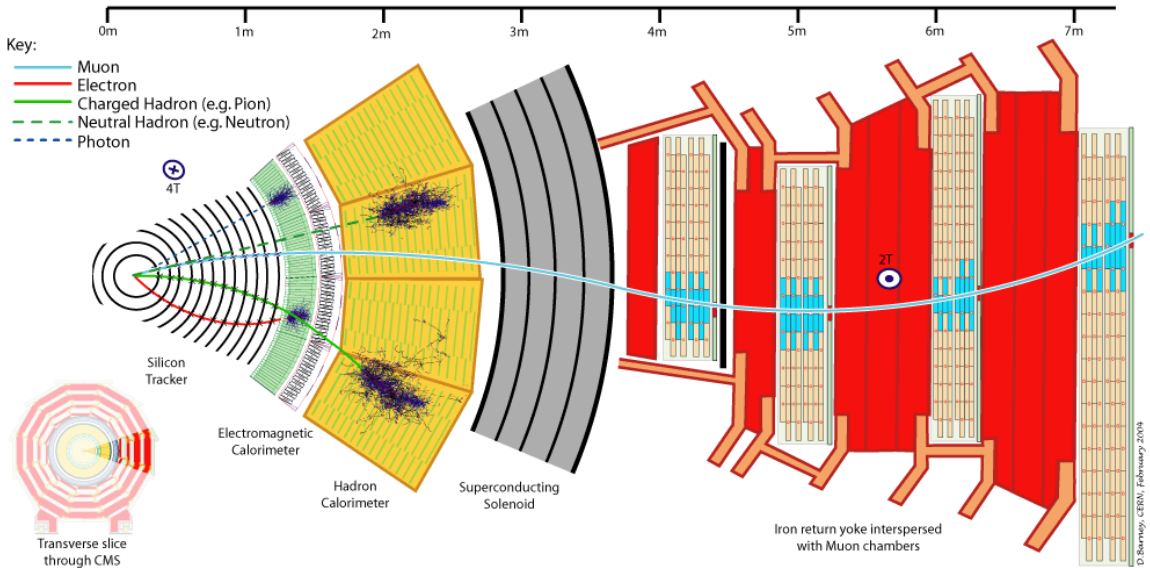

Figure 3.10: Transverse slice through the CMS detector. The interaction point is located on the left. Signatures of different particles are illustrated [169].

Muons

Since muons are more than 200 times heavier than electrons, they lose their energy in matter predominantly by ionisation [4, 171]. At the energies relevant in CMS, they are minimum ionising and therefore traverse the whole detector with little energy loss: Approximately only 2 GeV are deposited in the calorimeters [133]. Therefore, muon detection is based on hits in the inner tracking system and in the muon system.

Hadrons

Charged hadrons lose energy in matter via ionisation and through strong interactions with the detector material, hereby converting them into a number of new particles [4, 171]. Thus, they induce hits in the tracker and hadronic showers in the calorimeters, mostly the HCAL due to its larger hadronic interaction length. Neutral hadrons produce only hadronic showers. The hadrons typically reaching the sensitive layers of the detector are charged pions, charged kaons, protons, and neutrons.

3.4.2 Particle Identification with Particle-Flow

CMS employs a global event-reconstruction, referred to as *Particle-Flow* (PF), to identify the types and reconstruct the four-momenta of all stable and meta-stable particles, by combining information from all subdetectors [172–174]. A lightweight version of the algorithm is also employed at trigger level.

The procedure starts with the identification of charged-particle tracks traversing the inner tracking system and the muon system. In addition, adjacent calorimeter cells with energy deposits are combined to clusters. In general, single particles pro-

duce several of these signatures. They are therefore linked to blocks used for the full reconstruction of the particle while avoiding double-counting of the elementary signatures in different subdetectors.

In the final reconstruction of the particles from the linked blocks, dedicated quality criteria are applied. After muon identification and removal of their tracks from the blocks, electrons and associated bremsstrahlung photons are reconstructed from tracks and linked ECAL clusters with a subsequent removal of the associated blocks. Each remaining track is examined under the charged hadron hypothesis and its momentum is assigned by combining track momentum and linked cluster energies taking into account the detector resolutions. If the cluster energies overshoot the track momentum beyond the calorimeter resolution, the energy excess is used to form photons and neutral hadrons. Similarly, clusters without associated tracks are considered as photons or neutral hadrons. In both of the cases, the ECAL energy is fully associated to photons because on average 25 % of the jet energy is carried by photons and only 3 % of the ECAL energy deposit is caused by neutral hadrons.

The calorimetric part of the energy linked to charged hadrons is calibrated to correct the energy-dependent response of the HCAL to hadrons, the different electromagnetic to hadronic response ratios of HCAL and ECAL, and instrumental effects like readout thresholds. The overall corrections vary within 20 – 30 %. The ECAL has been calibrated for photons and electrons, therefore, clusters associated to these particles need only small adjustments of the order of 1 % to correct for residual instrumental effects. The calorimeter calibration is derived from simulation and validated with collision data [174]. The response to single hadrons in simulation is found to be up to 5 % larger than in data for p_{T} values below 30 GeV. This effect reduces to 2 % for jets because the calibration affects mainly neutral hadrons which are responsible for only 15 % of the jet energy [175]. The residual effect is corrected for by the jet-energy-calibration procedure described in Section 3.4.3. The photon calibration yields an accuracy in the measurement of the π^0 mass of better than 1 % in the ECAL barrel for data and for simulation [176].

The high magnetic field and the high granularity of the different subdetectors, especially the ECAL, are important prerequisites for the PF algorithm at CMS. For instance, the jet-measurement performance is remarkably improved [172, 177].

3.4.3 Physics Objects Reconstruction

The physics objects used in the analysis described in Chapter 5 are reconstructed using the standard algorithms of CMS. This section describes the reconstruction and identification of the objects relevant for the analysis.

Primary Vertices

The interaction of the most energetic partons, referred to as hard interaction, is of high interest in an event. However, at the LHC, the number of pp interactions in the collision of two proton bunches is typically not restricted to the hard interaction. Due to the large number of protons per bunch, needed to achieve a high instantaneous

luminosity, multiple interactions occur. This leads to a high-occupancy environment for track reconstruction, which is referred to as *pileup* (PU). The number of PU interactions, shown in Fig. 3.11, is on average 21 during the 2012 running conditions.

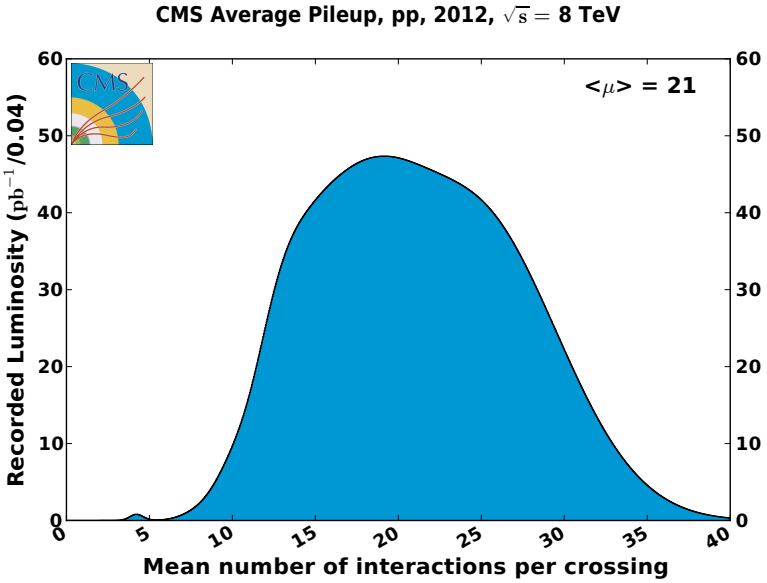

Figure 3.11: Average pileup in 2012, i.e. the mean number of interactions per bunch crossing [178].

Primary vertices are reconstructed from tracks consistent with being produced in the primary interaction region and are identified using the Deterministic Annealing (DA) clustering algorithm [179]. The vertices are required to have a z position within 24 cm of the nominal detector centre, a radial position within 2 cm from the interaction point, and there must be more than four degrees of freedom in the vertex fit. The primary vertex with the largest value of the p_{T}^2 sum of tracks associated with it is chosen as the one originating from the hard interaction.

Muons and Electrons

Muons and electrons are employed in the analysis presented in this thesis to define control regions used to predict the SM backgrounds. Their identification is briefly described in the following.

Muons are required to be global, i.e. coincident signatures in the muon system and the inner tracking system must match. Further quality criteria are applied following the "tight" selection of muons in CMS [180], i.e. the misidentification rate is lower than 0.1 %: Hadronic punch-through and muons from decays in flight are suppressed by requiring one muon chamber hit in the global fit, muon segments in two matched muon segments, and a normalised χ^2 of the muon-track fit of less than 10. The tracker-track transverse impact parameter needs to be less than 2 mm and the longitudinal distance is required to be less than 5 mm with respect to the primary vertex to suppress cosmic muons and to further suppress decays in flight. The longitudinal distance requirement

also removes tracks from PU. An additional removal of muons from decays in flight is achieved by requiring a minimum of one pixel hit and more than five tracker layers with hits. The latter also guarantees a good p_{T} measurement. For this analysis, all muons have to fulfil isolation criteria. The absolute isolation is defined as the sum of the energies of all charged particles, photons and neutral hadrons reconstructed by the PF algorithm within a radius of $\Delta R < 0.3$ in the η-ϕ plane. The relative isolation is defined as the ratio of the muon's absolute isolation to its p_{T}, and is required to be less than 0.12. Finally, the kinematic variables are restricted to $|\eta| < 2.4$ and $p_{\mathrm{T}} > 10\,\mathrm{GeV}$.

As opposed to muons, electrons produce electromagnetic showers in the ECAL and typically do not reach outer detector layers. Based on a charged-particle track and ECAL energy clusters associated with it, electrons are reconstructed from superclusters using the PF algorithm. The selected electrons have to fulfil the same kinematic requirements as the muons used in this analysis, while electrons with superclusters in the ECAL barrel-endcap transition region ($1.4442 < |\eta| < 1.5660$) are excluded. Further quality criteria are applied according to the "tight" selection of electrons in CMS [181]: The association of the track and the supercluster is enhanced by requiring the extrapolated track at the ECAL surface and the barycentre of the supercluster to be close in the η-ϕ plane. Misidentified electrons are rejected using the transverse extension of the supercluster, and the ratio of the energy deposit in the HCAL to the ECAL deposit has to be less than 0.12 (0.10) for electrons in the barrel (endcaps). Moreover, imposing $|1/E - 1/p| < 0.05$ further ensures the association of the track with the supercluster. A conversion-vertex finder rejects electrons coming from photon conversion. To provide electrons originating from a well-reconstructed primary vertex, electron tracks have to be within 0.2 mm around the closest primary vertex in the transverse plane and closer than 1 mm in the z direction to the same vertex. The isolation is defined similar to the muon case and the relative isolation has to be smaller than 0.15 for all electrons but for endcap electrons with a $p_{\mathrm{T}} < 20\,\mathrm{GeV}$ for which the requirement is tightened to be less than 0.10. Finally, electrons are rejected that are closer than 0.1 in the η-ϕ plane from any selected muon.

Jets

Jets provide information about partons in the hard interaction in pp collisions. Outgoing quarks and gluons manifest themselves as jets since their hadronisation typically results in the creation of many hadrons. The dynamics of QCD promote the collimation of these hadrons along the direction of the original parton. These bundles are reconstructed as compound objects, thereby reducing the number of physics object in the final states.

The composite measurement of the particles associated to a jet provides information on the parameters of the original parton. However, the assignment of a parton to a jet is not obvious, but is defined by the jet algorithm used to associate the jet constituents. Different jet types are characterised by the choice of subdetector information and their combination. The most relevant jet properties are the energy and direction. These parameters differ systematically between the reconstructed jet and

the underlying parton. Therefore, phase-space-dependent jet energy corrections are applied.

Jet Algorithms Jets are measurable objects, defined by the underlying clustering algorithm. General aspects of such algorithms are described in [182]. The design goals of the clustering methods are infrared safety and collinear safety, i.e. the result must have no dependence on additionally radiated soft partons and splitting of the parton, respectively. An important practical aspect is the speed of the algorithm, especially at trigger level. CMS employs several jet algorithms, which can be distinguished into two categories: Cone jet algorithms, defining a jet as a cone around the directions of high energy-flow, and sequential recombination algorithms, which cluster the jet constituents according to a defined metric. Important algorithms use as metric

$$d_{ij} = \min\left(k_{\mathrm{T},i}^{2p}, k_{\mathrm{T},j}^{2p}\right) \frac{\Delta_{ij}^2}{R^2} \,, \tag{3.6a}$$

$$d_{iB} = k_{\mathrm{T},i}^{2p} \,, \tag{3.6b}$$

where d_{ij} is the distance between entities (particles, pseudojets) i and j, d_{iB} is the distance between entity i and the proton beam axis, $\Delta_{ij} = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, and $k_{\mathrm{T},i}$, y_i and ϕ_i are the transverse momentum, rapidity and azimuthal angle of the entity i. First, all possible d_{ij} and d_{iB} are calculated and the smallest distance is chosen as the starting point. If the smallest distance is of type d_{iB} the entity i is a final jet, else the entities i and j are merged and the algorithm starts its next iteration until all particles are clustered into jets. Different settings for the parameters p and R define specific algorithms. The case with $p = 1$ is referred to as k_{T} algorithm, $p = 0$ is called Cambridge–Aachen algorithm, and $p = -1$ is the anti-k_{T} algorithm which is used in the analysis presented in this thesis with $R = 0.5$. Among other advantages [183], the shape of anti-k_{T} jets is usually more cone-like than the other two algorithms which facilitates experimental calibration and PU handling. Sequential recombination algorithms are usually collinear safe and infrared safe.

Jet Types The above discussed algorithms are applied to different input objects, specifying the jet type. In simulated events, the generated particles can be clustered to "generator jets", while neutrinos are excluded or not, depending on the specific analysis. For reconstructed objects, CMS uses four jet types. "Calo jets" use calorimeter information only, "track jets" use solely the inner tracking system and a third approach is to combine both sets of information to "jet-plus-track (JPT) jets". However, the type used in this thesis are the "PF jets" which are based on the objects reconstructed by the PF algorithm (see Section 3.4.2). This jet type performs best [177], since it employs the whole detector information, and thus can compensate shortcomings of one detector with measurements from another.

Jet Energy Corrections The measured jet energy differs from the underlying parton's energy, mainly caused by the non-uniform and non-linear calorimeter re-

sponse. Moreover, additional energy is measured due to PU and electronics noise. This energy mismatch is corrected by a multiplicative factor on the four-vectors of the mean jet-energy response which is specifically derived for all four jet types. The corrections of the jet energy scale (JES) in CMS [177,184] are phase-space dependent and factorise in subsequent levels as illustrated in Fig. 3.12. The input collection are the uncorrected reconstructed jets where charged particles associated with other than the reference primary vertex have been discarded in the jet reconstruction. This reduces the effect of PU and is referred to as *charged hadron subtraction* [185]. The output of the jet-energy-correction (JEC) procedure are the calibrated jets. The first three steps correct for the difference between the measured response and the true parton energy obtained in simulated events. The first level removes the energy offset from PU and electronics noise to give the correct energy on average. The second step corrects the jet response in all η regions relative to the $|\eta| < 1.3$ region, providing a flat response with respect to the pseudorapidity of the jets. The third correction balances the response with respect to p_{T} and corrects the absolute JES. Small residual differences to the data response are removed by applying an additional correction on real data only. The jets obtained after these calibration steps are used in this thesis.

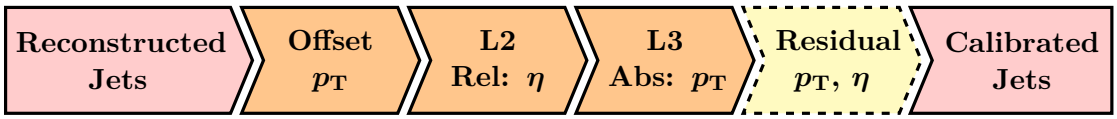

Figure 3.12: Jet energy correction scheme. The calibrated jets are obtained from the uncorrected reconstructed jets by rescaling in a p_{T}- and η-dependent multi-step procedure. For the corrections in 2012, the steps "L2" and "L3" are combined to a single p_{T}- and η-dependent correction.

Jet Identification The PF jets used in this thesis are clustered from all PF particle-candidates, e.g. even an isolated muon is reconstructed as jet containing one charged particle. In addition, calorimeter noise can induce falsely reconstructed jets. Hence, "loose" quality criteria are applied to all jets used in this work to remove the above-mentioned objects, while sustaining high selection efficiencies. The jets must have a neutral hadronic fraction below 99 %, a neutral electromagnetic fraction below 99 % and at least two constituents. For jets within the tracker acceptance, i.e. $|\eta| < 2.4$, additional requirements must be met. The charged hadronic fraction has to be greater than 0 %, the charged electromagnetic fraction must be below 99 % and at least one charged constituent is required. Moreover, jets clustered from particles originating from PU interactions, are identified and rejected with a multivariate approach [185], thus further mitigating the adverse effect of PU on the detector performance.

Heavy Flavour Identification

Jets originating from heavy-flavour quarks (c, b) are in general distinguishable from jets initialised by gluons or light quarks (u, d, s) because bottom and charm hadrons

have relatively large masses, long lifetimes and high-momentum decay products. The algorithms used by CMS to identify b-quark jets (b tagging) take advantage of these properties. A brief review of techniques commonly used in CMS analyses is provided below and follows the description in [186, 187].

In general each of the algorithms returns for each jet a single discriminator. Jets with a discriminator value above a given threshold are identified as b jet. These minimum thresholds define "loose" (L), "medium" (M), and "tight" (T) operating points for each of the algorithms with a misidentification probability for "light"-flavour (udsg) jets of 10 %, 1 %, and 0.1 %, respectively, at a jet p_{T} of 80 GeV. The operating point is indicated by appending the corresponding label to the acronym of the algorithm's name.

The *Track Counting* (TC) algorithm takes advantage of the long lifetime of bottom hadrons and uses as single discriminating variable the impact parameter (IP) significance of tracks. It is defined as the ratio of the signed impact parameter to its uncertainty estimate. The algorithm sorts the tracks by decreasing values of this variable, thus the IP significance of the first track is biased to higher values. However, it is less likely to have several tracks with high IP significance from non-b jets. Hence, the two versions of the algorithm, called *Track Counting High Efficiency* (TCHE) and *Track Counting High Purity* (TCHP), use the second and third track, respectively.

The long lifetime of bottom hadrons is also utilised by the *Simple Secondary Vertex* (SSV) algorithm, but the significance of the flight distance is used as discriminator, and is estimated from the vector between primary and secondary vertices. The algorithm's efficiency is limited to about 65 % by the secondary-vertex reconstruction efficiency, which in turn depends on the required number of tracks associated with the secondary vertex. Hence, analogous to the TC algorithm, a *High efficiency* (SSVHE) and a *High Purity* (SSVHP) version exist, which require at least two and at least three tracks, respectively.

The b-tagging algorithm used in this thesis is an extension of the SSV algorithm. The *Combined Secondary Vertex* (CSV) algorithm employs additional kinematic properties of the bottom-hadron decay. It uses track-based lifetime information, thus providing a higher maximum efficiency than SSV if no secondary vertex was reconstructed. The final discriminator is obtained with a multivariate approach and its performance is shown in Fig. 3.13 in multijet (left) and $\mathrm{t\bar{t}}$ (right) enriched samples. The medium working point with a discriminator threshold of 0.679 is used in the presented analysis, corresponding to an efficiency of $\sim 85\,\%$.

Hadronic Activity

The hadronic activity H_{T} of an event is an important variable for analyses with a signal topology exhibiting a high jet multiplicity. It is defined in this thesis as the scalar p_{T}-sum of all PF jets fulfilling $p_{\mathrm{T}} > 40\,\mathrm{GeV}$ and $|\eta| < 3.0$:

$$H_{\mathrm{T}} = \sum_{\mathrm{jets}} p_{\mathrm{T}}^{\mathrm{jet}} \,, \qquad \text{if } p_{\mathrm{T}}^{\mathrm{jet}} > 40\,\mathrm{GeV}, |\eta^{\mathrm{jet}}| < 3.0 \,. \tag{3.7}$$

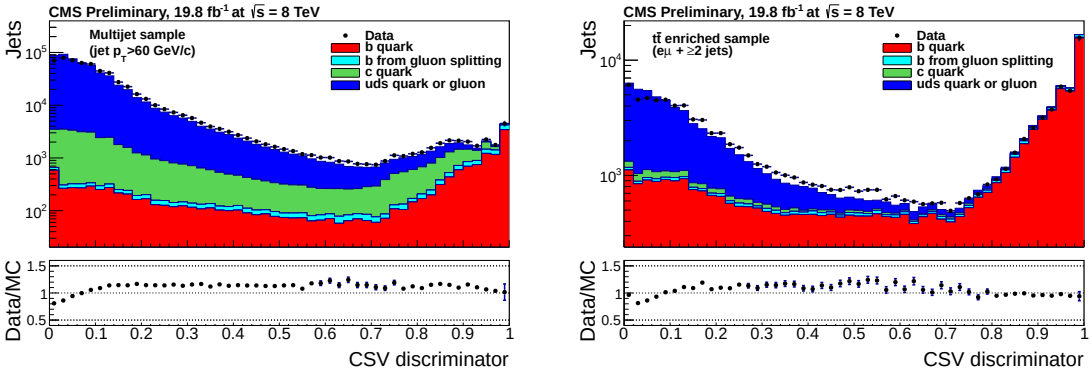

Figure 3.13: Distributions of the CSV discriminator in multijet (left) and $\mathrm{t\bar{t}}$ (right)-enriched samples at $\sqrt{s} = 8\,\mathrm{TeV}$ for data and simulation [187]. The simulated distributions are separately shown for b quarks, b quarks from gluon splitting, c quarks, and light partons. The description of the b-tagging-related quantities in simulation is in good agreement with data. The residual differences in the performance are corrected in physics analyses.

This definition was chosen to be in agreement with the requirements at trigger level in CMS.

Missing Transverse Energy

The missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ is the imbalance of the energy measurement in the plane perpendicular to the beam direction. Each particle that interacts with the detector deposits a signal corresponding to its energy. Under the assumption of no significant initial-state momentum orthogonal to the beam direction, the vectorial sum of the energy of all outgoing particles must vanish. Due to momentum conservation and assuming well-measured objects, a non-vanishing sum would be due to weakly interacting particles. These particles escape detection as, e.g. neutrinos or possible BSM weakly-interacting long-lived particles like neutral LSPs in SUSY models that conserve R-parity. Hence, high $E_{\mathrm{T}}^{\mathrm{miss}}$ is part of many BSM signatures and is one of the key observables to discriminate BSM signal from SM background.

Analogous to jets, $E_{\mathrm{T}}^{\mathrm{miss}}$ can be reconstructed from different input objects, e.g. only calorimeter information is used, optionally taking tracker information into account. In this thesis, PF candidates are used to obtain $E_{\mathrm{T}}^{\mathrm{miss}}$ where JEC are propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. The $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ vector used throughout this work is defined as the negative vectorial p_{T}-sum of all calibrated jets and all unclustered PF candidates

$$\vec{E}_{\mathrm{T}}^{\mathrm{miss}} = -\sum_{\mathrm{jet}} \vec{p}_{\mathrm{T,jet}}^{\,\mathrm{JEC}} - \sum_{i \in \mathrm{uncl.}} \vec{p}_{\mathrm{T},i} \; . \tag{3.8}$$

The missing transverse energy is defined as the absolute value $E_{\mathrm{T}}^{\mathrm{miss}} = |\vec{E}_{\mathrm{T}}^{\mathrm{miss}}|$. The performance of the different $E_{\mathrm{T}}^{\mathrm{miss}}$ algorithms used within CMS is documented in [143, 188].

4 CMS Tracker Alignment

The inner tracking system is the central component of the CMS detector. Excellent tracking performance is crucial for many physics analyses, e.g. the experimental analysis in this thesis relies on the identification of b-quark jets, and the b-tagging algorithms used within CMS utilise the information provided by the tracker [186,187]. Hence, the precise determination of the position and orientation of each tracker module, referred to as alignment, is essential. An accurate calibration and alignment of the tracker are required to measure the parameters of high-momentum tracks with high reconstruction efficiency [133], thus meeting the design goals of a transverse momentum resolution of 1.5 % and 10 % for muons with a momentum of 100 GeV and 1 TeV, respectively [133].

One of the most important inputs for track reconstruction is the *tracker geometry*, which is a parameterisation of the geometrical properties of the tracker modules. The tracker performance is limited by potential misalignment of the tracker geometry. However, the large number of tracker modules (see Section 3.2.1) with relative distances of up to $\approx 6\,\mathrm{m}$ are a challenge for the alignment procedure. Because of the limited accessibility of the tracker and the high precision required, the approach to this task is based on reconstructed tracks, thereby including also time-dependent effects. The alignment should reach a level of statistical accuracy well below the intrinsic silicon hit resolution of $10-30\,\mu\mathrm{m}$ [189, 190]. In addition to the statistical precision, systematic effects need to be considered. The determined tracker geometry could be systematically distorted, due to biases in the hit and track reconstruction, imprecise description of material effects or uncertainties in the magnetic field estimation, or by an insensitivity of the alignment procedure itself.

Statistical and systematic uncertainties on the tracker geometry need to be well under control to ensure the targeted physics performance, e.g. the b-tagging performance can be worsened by large misalignment [133, 186, 191].

This chapter contains, after a brief summary of the track reconstruction [136], an overview of the tracker alignment procedure within CMS followed by some details about the tracker alignment validation methods [137]. The central tool for the latter part has been largely rewritten in the course of the thesis, thus providing a better organised and more user-friendly interface to the individual validations.

4.1 Track Reconstruction

Since the alignment of the tracker and the reconstruction of charged particle trajectories, referred to as tracks, are intimately connected with each other, the details of the track reconstruction are briefly reviewed in this section.

Track reconstruction is an immensely challenging task due to the high track multiplicity of about 1000 at the design LHC luminosity [22]. High track-finding efficiency is demanded, while keeping the rate of falsely reconstructed tracks low. Additionally, the tracking software must be fast in order to be used also for the HLT (see Section 3.2.5), which has to cope with event rates around 100 kHz. For instance, the triggers used in the analysis presented in Chapter 5 use a lightweight PF algorithm, with tracks as input, to compute the hadronic activity H_{T} at trigger level.

The tracking procedure is performed in two successive steps, the local and global reconstruction. The former uses the detector readout information to reconstruct local hit candidates, i.e. the position of a charged particle's passage through a silicon module. The subsequent global reconstruction combines the hits to tracks, which are an input for the PF jets used in this work.

4.1.1 Local Reconstruction

In the first step of the reconstruction process, signals are digitised. If they are below specified thresholds, they are considered as noise, and are discarded before further processing. Neighbouring channels are clustered, according to certain signal-to-noise ratios for the individual channels and for the whole cluster. Using a cluster parameter estimator (CPE) in a subsequent step, the cluster positions and their uncertainties are calculated. These are defined in a local orthogonal coordinate system (u, v) in each sensor plane (see Fig. 4.1) and are, together with the charge and profile of the clusters, the output of the local reconstruction.

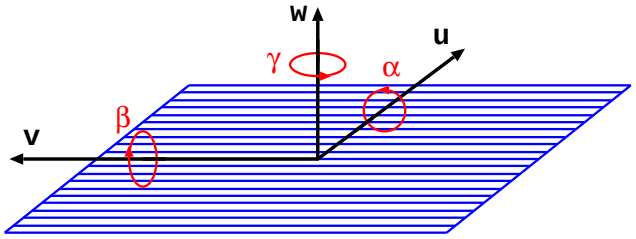

Figure 4.1: Sketch of a rectangular silicon strip module showing the axes of its local coordinate system, u, v, and w, and the respective local rotations α, β, γ. Taken from [137].

4.1.2 Global Reconstruction

The result of the local reconstruction is fed into the global reconstruction procedure. The translation of the local coordinate system of the hits into the global track coordinates takes into account deviations between the assumed and actual position, and surface deformations of the tracker modules as determined by the alignment procedure [137] described in Section 4.2. In addition, the status of the individual modules is considered, i.e. channels or complete modules that cannot provide a signal, or cause noise are deselected.

The global reconstruction is split into four consecutive stages: seed generation, track finding, track fitting, and track selection. This sequence is performed iteratively, referred to as *iterative tracking* [136], and is outlined in the following subsections. The iterations mainly differ in the configuration of the seed generation (see Table 4.1), and the final track selection.

Table 4.1: Configuration of the track seeding for each of the six iterative tracking steps. Listed are the seeding layers and the requirements on the minimum p_{T}, and the maximum transverse (d_0) and longitudinal (z_0) impact parameters with respect to the centre of the interaction region. The Gaussian standard deviation of the longitudinal profile of the interaction region is denoted as σ. The asterisk symbol indicates the impact parameter with respect to a pixel vertex, i.e. a vertex reconstructed prior to the seed generation with a very fast algorithm using only pixel detector information [136].

Iteration	Seeding Layers	p_{T} (GeV)	d_0 (cm)	$\lvert z_0 \rvert$
0	Pixel triplets	> 0.8	< 0.2	$< 3\,\sigma$
1	Mixed pairs with vertex	> 0.6	< 0.2	$< 0.2\,\mathrm{cm}^*$
2	Pixel triplets	> 0.075	< 0.2	$< 3.3\,\sigma$
3	Mixed triplets	> 0.35	< 1.2	$< 10\,\mathrm{cm}$
4	TIB 1+2 & TID/TEC ring 1+2	> 0.5	< 2.0	$< 10\,\mathrm{cm}$
5	TOB 1+2 & TEC ring 5	> 0.6	< 5.0	$< 30\,\mathrm{cm}$

The CMS tracking software is denoted as *Combinatorial Track Finder* (CTF), which is a modification of the combinatorial Kalman filter [192–194], thus an extension of the Kalman filter [195], that allows track finding and track fitting within the same framework.

The description below reflects the status of the software from May through August, 2011 [136]. Later versions use the same concept, but with different configurations of the iterations [196].

Seed Generation

The seed generation produces initial track candidates based on two or three hits. A seed is a first estimate of the trajectory parameters and the associated uncertainties. The trajectory of charged particles can be approximated by a helix parameterisation, due to the nearly uniform magnetic field in the tracker volume. Thus, one needs five parameters to define a trajectory. To extract these parameters, one needs three base points of the track, each providing three-dimensional (3-D) information, i.e. either three pixel hits, or two pixel hits and a constraint on the track origin based on the assumption that the particle originated in the pp interaction region.

Seeds are constructed in the innermost region because the channel occupancy of the inner pixel layers is due to the high granularity much lower than in the outer strip layers. In addition, the pixel layers provide 3-D measurements, resulting in more constraints and smaller uncertainties of the trajectory estimates. Finally, it has

been shown that seed generation in the inner tracking system results in higher track reconstruction efficiency [136].

The seed generation is configured by two parameter sets, *seeding layers* and *tracking regions*. The former are pairs or triplets of detector layers, that are searched for hits. The tracking regions are parts of the track-parameter space that limit the acceptable range for the minimum p_{T}, and the maximum transverse and longitudinal IP with respect to the assumed production point of the particle (see Table 4.1).

Track Finding

Track finding is a pattern-recognition task and constitutes the computationally most expensive part of track reconstruction. The CTF algorithm begins with the seeds generated in the previous step, and then propagates tracks by adding hits from successive layers, while updating the track parameters at each layer (see Fig. 4.2). The propagation window size depends on the uncertainties on the track parameters, which decrease with every added hit. If more than one hit is compatible, a candidate is built for each. Additionally, one candidate is always propagated without a hit in the layer, a *ghost hit*, because sometimes real hits cannot be reconstructed. Depending on the iteration, zero or one layers without a hit are allowed during track finding (see Table 4.2).

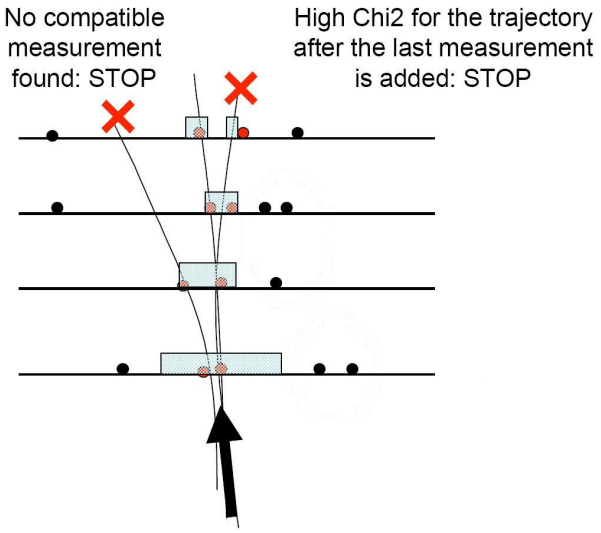

Figure 4.2: Track pattern recognition based on a Kalman filter [197]. The algorithm starts at the bottom layer and searches for compatible hits in the subsequent layers. The left path is aborted after two layers without compatible hits. The path ending at the right is rejected because of an incompatible last hit, while the central track candidate is kept.

This procedure is continued till the last layer and repeated from outside inwards, if a minimum number of hits has already been found (N_{rebuild}, see Table 4.2), to search for additional hits. In order to avoid a rapid increase in the number of candidates,

Table 4.2: Selection requirements applied to track candidates during the six track finding iterations: minimum p_{T}, minimum number of hits N_{hits}, and the maximum number of missing hits N_{lost}. The minimum number of hits required in the outward track building step to trigger the inward track building step N_{rebuild} is listed. However, candidates failing this requirement are not rejected. The later iterations have loose requirements on the impact parameters of the tracks. Therefore, the requirements on the number of hits are tightened to avoid random hits forming a track [136].

Iteration	p_{T} (GeV)	N_{hits}	N_{lost}	N_{rebuild}
0	0.3	3	1	5
1	0.3	3	1	5
2	0.1	3	1	5
3	0.1	4	0	5
4	0.1	7	0	5
5	0.1	7	0	4

at most five are kept for further propagation, based on their normalised χ^2. A bonus is given for each valid hit, and a penalty for each ghost hit.

Track Fitting

The track-finding procedure provides a collection of hits and a track-parameters estimate for each trajectory candidate. However, the full information is available only after the last hit in the trajectory (see Fig. 4.3), and the estimate can be biased by constraints applied in the seeding procedure. Therefore, the track is refitted using a Kalman filter and smoother [136].

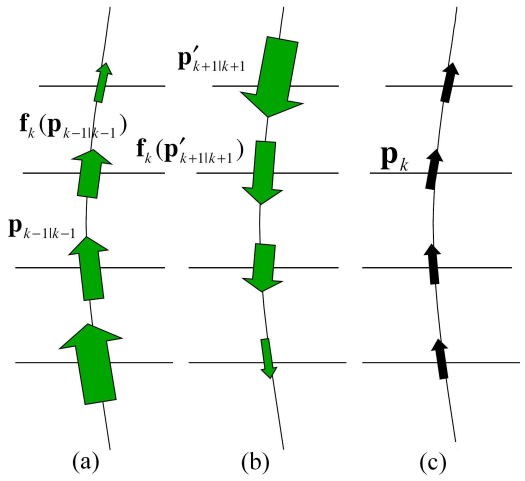

Figure 4.3: Track fitting [197]: (a) forward fitting, (b) backward fitting, and (c) trajectory smoothing. The arrow size illustrates the uncertainties on the track parameters on a given module.

The filter step starts at the innermost hit with a trajectory estimate obtained from a Kalman filter fit to the innermost hits of the track candidate. The fit then iterates through all hits, from inside outwards, while updating the trajectory estimate with each hit. In case of valid hits, the current track parameters are used to reevaluate the hit-position-uncertainty estimate, while for pixel hits also the position estimate itself is reevaluated.

After the first filter, the smoothing step follows. Hereby, the output of the first filter is used to initialise a second filter step from outside inwards. The final result is then computed from a weighted average of the outcome of the two filters (see Fig. 4.3), to obtain the optimal track parameters at any point. Best precision is achieved by using a *Runge–Kutta propagator* that takes into account material effects and inhomogeneities of the magnetic field. An accurate extrapolation of the trajectory is assured by a detailed magnetic field map, measured to a precision of $< 0.01\,\%$. The use of the Runge–Kutta propagator is most important for $|\eta| > 1$, where the field inhomogeneities are largest [136].

After filtering and smoothing, incorrectly associated hits, referred to as outliers, are searched and rejected based on their χ^2 compatibility with the corresponding track. This rejection is repeated until no hits are discarded anymore. If this procedure leads to two ghost hits in consecutive layers, the track is terminated. In case of tracks that finally arrive at less than three hits, these tracks are discarded completely.

The CMS software allows the repetition of the track fitting part, which is a common procedure in tracker alignment workflows, e.g. to validate alignment settings different from the ones used in the central reconstruction. A tool was developed in the course of this thesis to provide unified settings for this procedure, guaranteeing comparable results. It is part of a unified track selection and refitting sequence, referred to as *Unified Sequence* in the following.

Track Selection

In the interest of high efficiency, the track finding described above typically yields significantly many falsely reconstructed tracks. The rate of these tracks is reduced by applying certain quality criteria, which are tuned for each iteration of the tracking procedure. In general tracks are selected according to the number of hits, the χ^2/ndf of the final fit, and their compatibility with a primary vertex. The track selection applied after a possible refit in the tracker alignment procedures is also part of the Unified Sequence.

4.2 Tracker Alignment

For the track reconstruction described in Section 4.1, it is necessary to transform local hit-position measurements and their uncertainties into the global tracker-coordinate-system or, equivalently, the reverse. This task strongly depends on a precise knowledge of the tracker-module positions and orientations, sensor shapes, and assigned uncertainties. Prior to a description of the track-based internal alignment, that mainly

provides this information for the modules relative to each other, details about the determination of the global coordinates are given in the following section.

4.2.1 Global Position and Orientation of the Tracker

The absolute position and orientation of the tracker are obtained with survey measurements of the TOB, which is the largest subcomponent. Its shift and rotation around the beam axis are determined with respect to the beam axis. The other subcomponents are aligned relative to TOB using the internal alignment described in the following section.

Special attention has to be paid to the orientation of the tracker relative to the magnetic field, which is characterised by the tilt angles θ_x and θ_y. These describe rotations of the whole tracker volume around the x and y axes in the global CMS coordinate system. Uncorrected tracker tilts with respect to the magnetic field could bias the track parameter reconstruction and mass measurements of resonances inferred from their charged decay products. Therefore, the global tracker tilt angles have to be determined before the overall alignment corrections are obtained.

Ignoring deviations of the tilt angles from the ideal 0 mrad can result in wrong assumptions on the transverse magnetic field components, thus leading to a degradation of the observed track quality, which is estimated by the total χ^2, i.e. the χ^2 sum, of all fitted tracks divided by the number of tracks (see Fig. 4.4). The angles of optimal alignment are derived as the points with maximum overall track quality, i.e. minimum total χ^2.

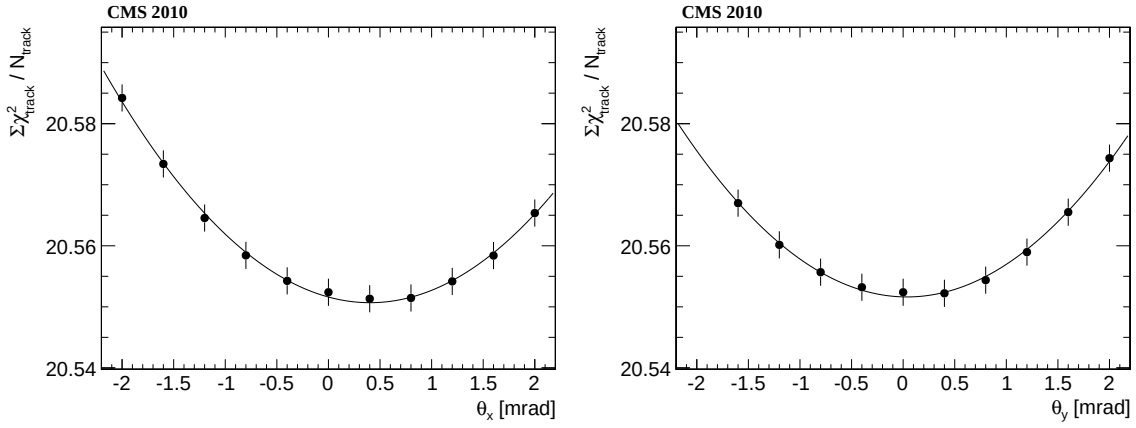

Figure 4.4: Dependence of the total χ^2 of the track fits, divided by the number of tracks N_{track}, on the assumed θ_x (left) and θ_y (right) tilt angles for $|\eta| < 2.5$ and $p_{\text{T}} > 1\,\text{GeV}$ [137].

Figure 4.5 shows the derived tilt angles for different pseudorapidity ranges. The right panel shows the results for simulated events without tracker misalignment, which are consistent with 0 mrad within the systematic uncertainties. The central tracker region within $|\eta| < 1.5$ exhibits smaller variations, that are well within a margin of

±0.1 mrad, which is used as rough estimate of the systematic uncertainty of the tilt angle determination.

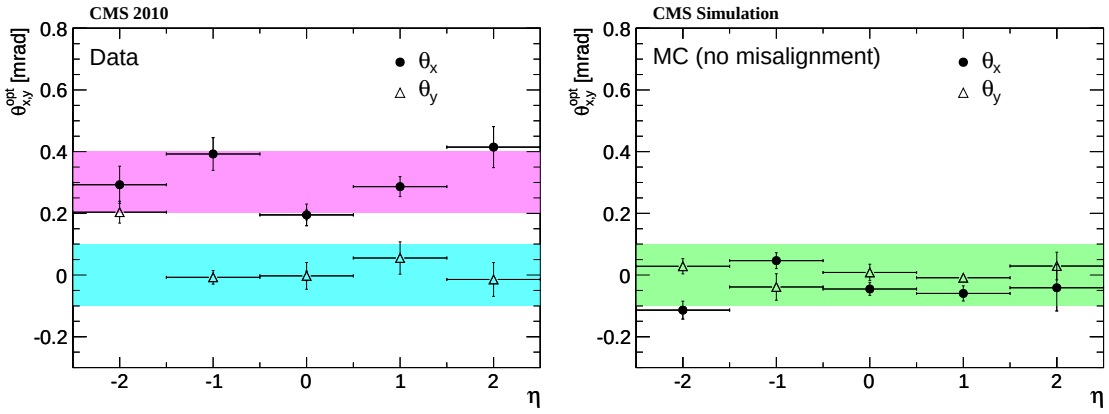

Figure 4.5: Tracker tilt angles θ_x (filled circles) and θ_y (hollow triangles) as a function of track pseudorapidity. The left panel shows the values measured in 2010 data and the right panel shows the values derived from simulated events. The shaded bands indicate the margins of ±0.1 mrad discussed in the text. The error bars are the root mean square of the distributions obtained when varying several parameters influencing the tilt angle determination [137].

4.2.2 Track-Based Alignment

If the tracker-module positions assumed in the track reconstruction differ from the true values, the distributions of the distances between hit position and the prediction from the track fit, referred to as track-hit residuals, are generally broadened. Hence, track-based alignment algorithms minimise the sum of squares of normalised residuals from many tracks. Under the assumption of independent hit position measurements m_{ij} with uncertainties σ_{ij}, the objective function to minimise reads

$$\chi^2(\mathbf{p},\mathbf{q}) = \sum_{j}^{\text{tracks}} \sum_{i}^{\text{measurements}} \left(\frac{m_{ij} - f_{ij}(\mathbf{p},\mathbf{q}_j)}{\sigma_{ij}} \right)^2 . \tag{4.1}$$

Here, f_{ij} is the track model prediction at the measurement position, which depends on the tracker geometry $\mathbf{p}$ and the track parameters $\mathbf{q}_j$. Design drawings, survey measurements, or previous alignment results can serve as starting geometry $\mathbf{p}_0$, that is used for an approximate track parameter determination $\mathbf{q}_{j0}$. The trajectory predictions f_{ij} can be linearised around these initial values because the alignment corrections are assumed to be small. A minimisation of (4.1) leads to a linear equation system involving the alignment parameters $\boldsymbol{\Delta}\mathbf{p}$ and corrections to the track parameters of all used n tracks $\boldsymbol{\Delta}\mathbf{q}^T = (\boldsymbol{\Delta}\mathbf{q}_1, \ldots, \boldsymbol{\Delta}\mathbf{q}_n)$. In case of non-small alignment corrections, the procedure has to be iterated.

A global fit approach [198] is used for the CMS tracker alignment, implemented in the MILLEPEDE II program [199]. It takes advantage of properties of the above-

mentioned linear equation system, that allow a reduction of the large system to a smaller one for the alignment parameters only. For 10^7 tracks with 20 parameters on average and 10^5 alignment parameters, the number of elements in the coefficient matrix is reduced by a factor larger than 4×10^6, without losing information for the alignment parameter determination [137].

In the following subsections, the track and alignment parameterisation is briefly described, followed by an introduction to the concept of hierarchical and differential alignment, a discussion of *weak modes*, and closing with a summary of the alignment strategy.

Track and Alignment Parameterisation

To describe a track trajectory in a homogeneous magnetic field, five parameters are needed when neglecting material effects like multiple scattering. However, the interaction with the material is significant in the CMS tracker. To thoroughly treat multiple scattering one can increase the number of track parameters to $n_{\text{par}} = 5+2n_{\text{scat}}$, where the additional parameters can be two deflection angles for each of the n_{scat} thin scatterers passed by the particles. Thick scatterers can be approximated with two thin scatterers. This complete parameterisation often leads to $n_{\text{par}} > 50$ for cosmic ray tracks [137], which are especially many parameters because cosmic particles do not originate in the tracker centre but traverse the whole detector.

In the general case, the computational effort to derive the inverse coefficient matrix of the above mentioned linear equation system scales with n_{par}^3, thus leading to a significant amount of computing time. The *general broken lines* (GBL) track refit [200–202] avoids the n_{par}^3 scaling by means of a custom track parameterisation. The implementation of this approach in MILLEPEDE II saves a factor of 6.5 in CPU time for isolated muon tracks and 8.4 for cosmic ray tracks compared to solving a linear equation system by matrix inversion [137].

The CMS tracker modules can be approximated as flat planes, i.e. deviations from the assumed tracker geometry can be parameterised with up to three shifts (u, v, w) and three rotations (α, β, γ) in the local coordinate system (see Fig. 4.1). But tracks with large incident angles, relative to the module normal, are sensitive to deviations from the flat-module approximation. Such deviations can emerge from possible curvatures of the sensors, or relative misalignment of the sensors in two-sensor modules. They are taken into account by extending the number of alignment parameters to up to nine degrees of freedom per sensor instead of six per module. Modified Legendre polynomials up to the second order are used to parameterise the sensor shapes.

Uncorrelated measurements in u and v direction are provided by the pixel modules, whereas the strip modules in TIB and TOB are only sensitive in u direction because the strips are parallel to the v axis. However, the strips in TID and TEC are not parallel, thus the reconstructed hits provide two-dimensional information. But the measurements in u and v are highly correlated. Therefore, their covariance matrix is diagonalised, and the less precisely measured direction, after diagonalisation, is not used in the alignment procedure [137].

Hierarchical and Differential Alignment

The CMS tracker has a hierarchical structure (see Fig. 4.6). Translations and rotations of whole substructures are treated with six additional alignment parameters each. The advantage of these large-substructure parameters is twofold. If the number of tracks is insufficient for the determination of the alignment parameters at module level, one can restrict the procedure to the much smaller set of these substructure parameters. Furthermore, they can be used in a hierarchical alignment approach together with the parameters of the sensors. Thus, coherent displacements of large structures along the non-sensitive coordinate v of the strip modules can be incorporated.

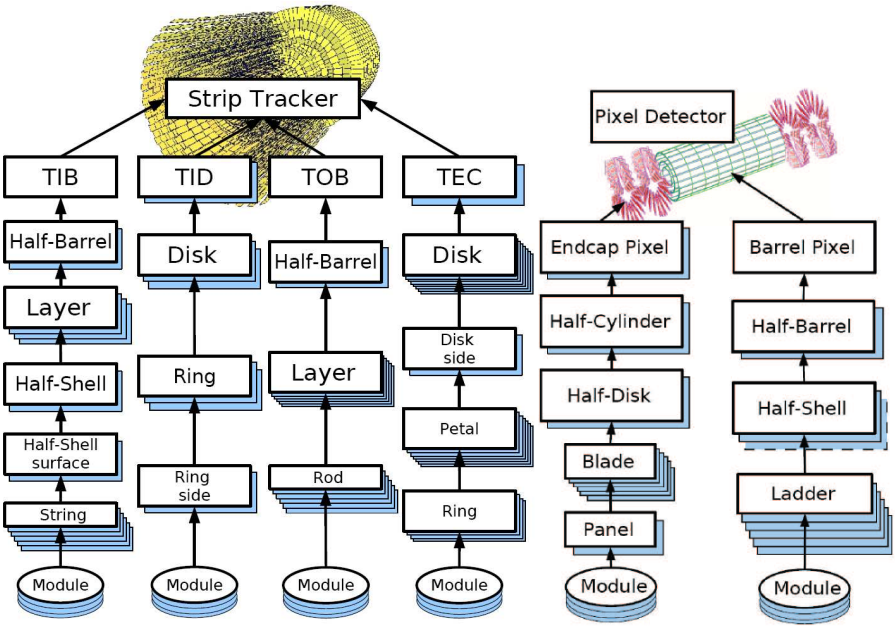

Figure 4.6: Hierarchical structure of the strip tracker (left) [203] and the pixel tracker (right) [191]. The hierarchy is driven by the mechanical mounting structure and the alignment objects relevant for track reconstruction.

The hierarchical approach results in redundant degrees of freedom because large-structure movements can be either expressed by the corresponding parameters or by the parameters of their components. These degrees of freedom are removed by linear equality constraints, that are implemented in MILLEPEDE II by extending the original linear equations systems using Lagrangian multipliers. This concept is also used to constrain the undefined overall shifts and rotations of the whole tracker [137].

The alignment procedure can treat time-dependent and time-independent parameters simultaneously by means of the *differential alignment*. For each *interval of validity* (IOV) a different parameter is used, thus allowing the use of the full statistical power of the dataset to determine the time-independent parameters, while taking into account the time dependence of the others. This approach can be joined with the above described hierarchical alignment, where large structures are IOV dependent, but the sensors remain stable relative to their large structure.

Weak Modes

Linear combinations of the alignment parameters that change at most slightly the track-hit residuals, thus the total χ^2 in (4.1), are called weak modes. These may arise if coherent changes of the alignment parameters $\mathbf{\Delta p}$ can be counterbalanced by changes of the track parameters $\mathbf{\Delta q}$. For instance, an overall shift of the tracker would be compensated by variations of the track impact parameters. This can be avoided by fixing the overall shift by means of the constraints mentioned above. However, other weak modes exist, that influence especially the p_{T} of the tracks, and contribute significantly to the systematic uncertainties of the track fitting procedure [137].

The type of possible weak modes depends on the tracker geometry and granularity, the topology of the tracks used for alignment, and on the alignment and track parameterisation. Weak modes in alignment approaches using only tracks traversing the beam line can be classified in cylindrical coordinates, i.e. by Δr, Δz, and $\Delta\phi$ [204]. To keep these modes under control, additional information needs to be included in (4.1). For instance, different track topologies and physics constraints can be incorporated, e.g. cosmic ray tracks breaking the cylindrical symmetry, or knowledge about the resonance whose charged daughter particles produce tracks. The decay channel $\mathrm{Z}^0 \rightarrow \mu^+\mu^-$ is of particular importance because muons are precisely measured in CMS and the reconstructed Z^0 mass is sensitive to the *twist* weak mode (see Fig. 4.7), i.e. a scenario in which the tracker modules are coherently shifted along ϕ [137].

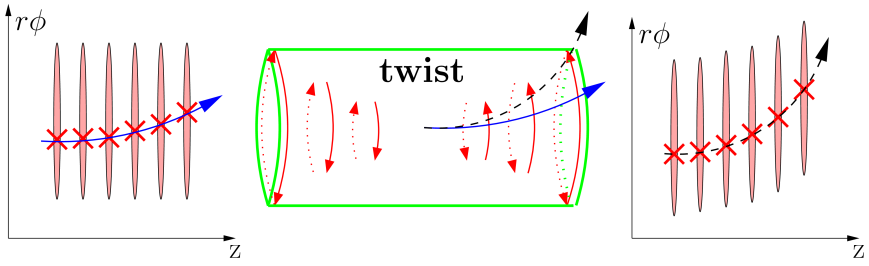

Figure 4.7: Illustration of the twist weak mode. The tracker modules are coherently shifted along ϕ. Taken from [205].

Strategy of the Internal Tracker Alignment

The strategy to obtain the alignment parameters is exemplified in the following using the procedure of the 2011 tracker alignment [137]. In general, the tracking detector has not undergone significant movements throughout this data-taking period. The large structures have been monitored using a system of infrared lasers [206], and statistical studies of primary vertex residuals (see Section 4.3.2). The statistical alignment precision has been validated [207, 208] and shows no need for time-dependent module-level parameters.

Two different tracker readout modes are applied during data taking. One is the *peak mode*, which reads the full signal shape. This mode is used for parts of the

cosmic data taking, but it is too slow for collision events with high interaction rates. Therefore, the faster *deconvolution mode* is used, which measures the signal turn-on at three points and extrapolates the signal shape. However, the disadvantage of this method is that not all charge carriers in a module reach the readout channels, resulting in a thinner effective active area of the silicon sensors. This effect is taken into account in the calibration by means of *back-plane corrections* but which have no significant influence [137]. The strong magnetic field overlaid to the electric field in the sensors shifts the drift direction of the charge carriers by a certain angle, the *Lorentz angle*. Associated calibration parameters are not used if the Lorentz drift is stable in time because this can be compensated by the alignment corrections.

Typically, the data sample used for alignment is split to first run a full-scale alignment utilising the different track topologies, as well as mass and vertex information. Afterwards, the large structures are monitored to detect relevant movements of the tracker components.

Four types of tracks are used in the alignment procedure, each reconstructed with the algorithms described in Section 4.1, and providing specific information about the tracker geometry. These are minimum bias samples, isolated muon samples, di-muon samples, and cosmic-ray-track samples. For all data sets, basic quality criteria are applied, for which the above-mentioned Unified Sequence could be used.

Events with isolated muons are selected by requiring global muons with certain kinematic requirements. These events are mainly populated with muons from leptonic W-boson decays and amount to 15 million tracks used for alignment.

Minimum bias events are collected with a combination of triggers, based on, e.g. pick-up signals marking the intersection of two proton bunches, signals from the beam scintillator counters [22], or minimum requirements on hit and track multiplicity.

Di-muon events originating from Z^0-boson decays are selected online with any trigger requesting two muons. Offline these muons are required to be global muons and oppositely charged. To obtain a pure sample of Z^0-boson candidates, the invariant di-muon mass is restricted to be within $85.8 - 95.8\,\text{GeV}$.

Tracks induced by cosmic rays are recorded both in peak and deconvolution mode. Peak mode data are taken before the LHC started operations and in beam-free periods between LHC fills. Furthermore, deconvolution mode data are taken both during and between LHC fills, using a special cosmic ray trigger.

The concept of hierarchical and differential alignment is applied, and IOVs are defined using the above-mentioned procedure based on primary vertex residuals. Overall, more than 200 000 alignment parameters are determined simultaneously, taking into account more than 100 constraints.

4.3 Tracker Alignment Validation

The above described alignment method is a very complex procedure, thus the validation of the alignment is a crucial part of the whole workflow. To fulfil this task, several tools have been developed and most of them can be steered with the uniform control interface of the *All-in-One Validation Tool*, which is largely rewritten and

extended in the course of this thesis. It allows the easy usage of the underlying tools without detailed knowledge of the algorithms, thus facilitating weekly validations by alternating users and at the same time yielding consistent results.

The new features include the validation of arbitrary calibration and alignment conditions and the configuration of the di-muon validation tool, which previously required expert knowledge to be run in stand-alone mode. Also tedious preparation steps are now automated, thus further increasing user-friendliness.

The first section deals with the validations integrated in the All-in-One Validation Tool, followed by the tools that exist to date only as stand-alone versions.

4.3.1 All-in-One Validation Tool

The All-in-One Validation Tool is commonly used in recent and past alignment campaigns to provide information about the status of the tracker, the effect of different alignment settings, or for weekly monitoring of the track-hit residuals. The integrated tools, which can be configured with a unified interface, are briefly described in the following.

Track-Based Offline Validation

In this method, track-hit residuals are determined, which are obtained without the probed hit to remove biases due to correlations between hit and track. The tool is used to measure the quality of alignment and related effects by studying the track-hit-residual distributions of individual modules and larger structures, as well as global track variables. During data taking it is used on a weekly basis to spot possible misalignment effects, but also in larger alignment campaigns to validate the alignment accuracy as illustrated in Fig. 4.8.

Di-Muon Validation

Possible biases in the reconstruction of resonances due to misalignment in the tracking detector can be studied with $Z^0 \to \mu^+\mu^-$ or $\Upsilon(1S) \to \mu^+\mu^-$ events. Di-Muon events are selected and the invariant mass of the di-muon system is fitted with a Voigtian or Gaussian function for Z^0 or $\Upsilon(1S)$ candidates, respectively. An exponential function is used to model the background. Especially $Z^0 \to \mu^+\mu^-$ events provide a handle on possible twists in the tracker geometry as shown in Fig. 4.9. This tool previously required expert knowledge, but has been integrated in the All-in-One Validation Tool, and therefore benefits from the extended options to validate various alignment and calibration conditions.

Cosmic Track Split Validation

The idea of this method is, to split cosmic tracks at their point of closest approach to the beam line, thus creating two individual track candidates. Subsequently the track parameters are compared to detect systematic misalignment. The method is e.g. sensitive to an off-centring of the barrel layers and endcap rings (*sagitta* misalignment).

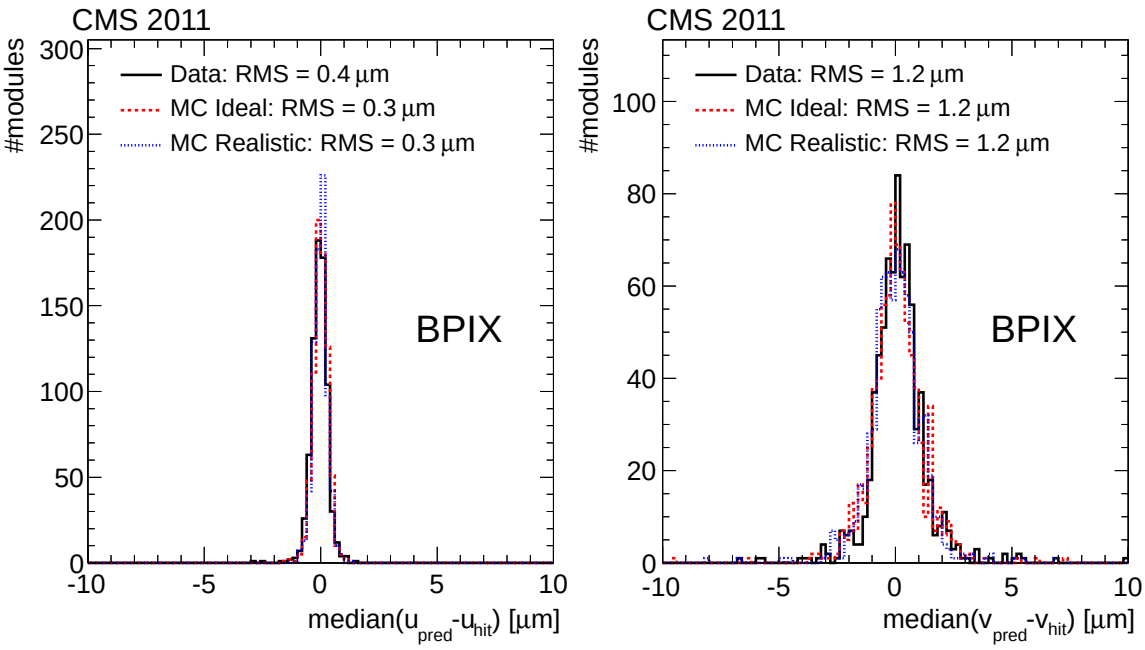

Figure 4.8: Distributions of the medians of the track-hit residuals (DMR) for the pixel tracker barrel in u (left) and v (right) coordinates. Shown are the distributions after alignment with 2011 data (solid), in simulation without misalignment (dashed) and with realistic misalignment (dotted). The obtained widths of the distributions are the same within the sensitivity limit of the DMR method. Thus, the alignment procedure can correct misalignment in data and simulation [137].

Geometry Comparison

This validation is special with respect to the other presented methods because it does not require data but only alignment parameters $\mathbf{\Delta p}$ of two tracker geometries. The tools compares these two tracker geometries at any hierarchy level. Global displacements can be eliminated to find relative displacements of any structure with respect to any other higher-level structure. It is used, e.g. to validate the positive effect of the alignment procedure with respect to an initial alignment setup, or with respect to the ideal tracker geometry. The successful realignment of misaligned geometries can be monitored with this tool, thus providing a possibility to judge, how many iterations of the alignment procedure are necessary to recover from misalignment.

4.3.2 Primary Vertex Validation

The Primary Vertex Validation is intended to detect biases in the pixel geometry description. It studies the distances between tracks and the associated primary vertex reconstructed without the respective track, referred to as unbiased track-vertex residuals. Systematic misalignment would bias the distributions of the residuals depending on the pseudorapidity and azimuthal angle of the tracks. During data taking, it is run on a daily basis, thus providing relatively fast feedback on possible move-

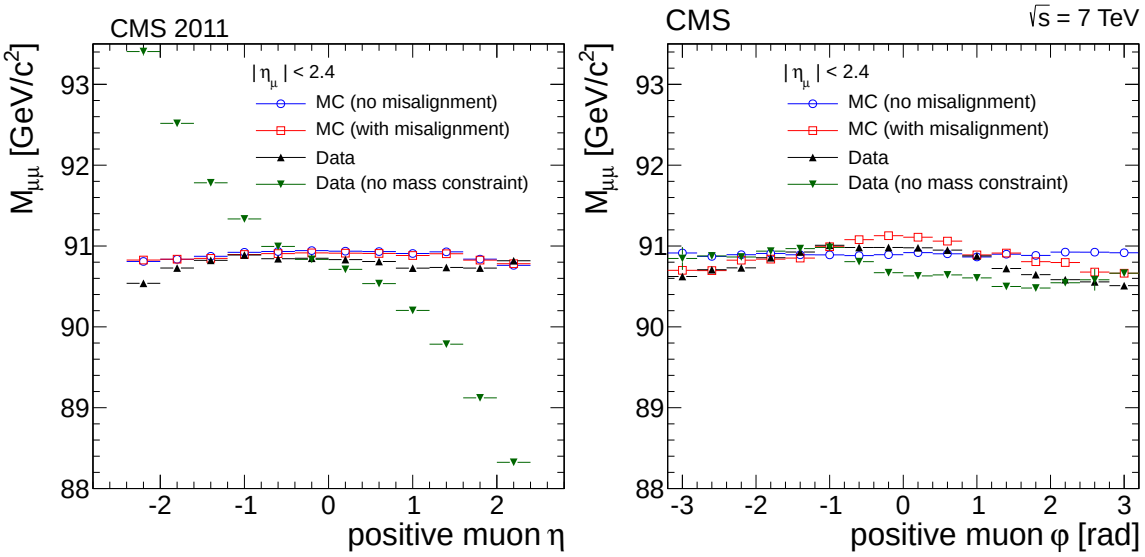

Figure 4.9: Invariant mass of $Z^0 \to \mu^+\mu^-$ candidates with respect to the pseudorapidity (left) and the azimuthal angle (right) of the positively charged muon. The shown distributions are from aligned data (black upward-pointing triangles), and from simulation without misalignment (blue circles) and with misalignment (red squares). The data distribution is also shown for a tracker geometry obtained without using Z^0-boson-mass information (green downward-pointing triangles). The clear trend in the latter distributions is attributed to a twist weak mode (see Section 4.2.2). This bias in the Z^0-mass reconstruction is removed by including the additional information of the Z^0 resonance [137].

ments of the large substructures of the tracker. In the alignment of 2011 it is used to determine the IOVs used for the alignment parameters, hence it is an important ingredient of all tracker alignment campaigns.

4.3.3 E/p Validation

Calorimetric information provides an additional measure of systematic effects in the track momentum determination. The value of this additional knowledge lies in the alternative information with respect to the $Z^0 \to \mu^+\mu^-$ decays, which are already used in the alignment procedure itself.

The tool measures the difference in the azimuthal angle of two tracks with the same transverse momentum and polar angle. Weak modes, that alter the azimuthal angle of the modules, would have an opposite effect on the track curvature for positively and negatively charged particle trajectories. Thus, the tool is able to detect twists in the tracker alignment.

5 Search for a Light NMSSM Higgs Boson

On the 4th of July, 2012 the discovery of a Higgs boson with a mass around 125 GeV has been announced [5,6] after decades of intense research induced by the prediction of a scalar elementary particle [54] within the BEH mechanism [54–56]. The properties of the newfound particle are within the experimental accuracy in agreement with the expectations from the SM [7–9], but it might well be that it corresponds to a state of an extended Higgs sector. Such a scenario is motivated by the known shortcomings of the SM outlined in Chapter 1, which suggest an extension of the theoretical description. Supersymmetry (SUSY), as a promising candidate, predicts in all its variants an extended Higgs sector.

A well motivated non-minimal implementation of SUSY is the NMSSM, featuring a singlet extension as described in Section 1.4.1. This results in extended Higgs and neutralino sectors with respect to the MSSM, giving rise in particular to a richer Higgs phenomenology. Interestingly, scenarios with a light Higgs boson in the mass range even below m_{Z^0} are not excluded, and occur very generically within the NMSSM (see Section 1.4.2). Such a light Higgs boson couples only weakly to gauge bosons because of a large singlet component, thereby evading constraints from previous Higgs boson searches in this low mass range.

Finding a light Higgs boson in addition to the one at 125 GeV is certainly a good way to show that the found Higgs boson is not the Higgs boson of the SM. This chapter describes the search for an additional Higgs boson, that is lighter than the currently known one. The analysis strategy is outlined, including online and offline event selection, background estimation, signal modelling, and the determination of systematic uncertainties. The search results and their interpretation are described in Chapter 6.

The analysis is published in [209] and figures from this document or from the corresponding public website [210] are indicated as such. All other figures are labelled as “Private Work”.

5.1 Signal Event Topology and Analysis Strategy

As discussed in Sections 1.4.2 and 1.5, previous searches do not exclude a light NMSSM Higgs boson with reduced couplings to gauge bosons and SM fermions. As a consequence of the reduced couplings, such a Higgs boson is not produced in standard channels (see Fig. 1.5) with discernible strength. Therefore, a search within cascade decays of supersymmetric particles is performed as suggested by [99]. The topology of such an event is illustrated in Fig. 1.7 and exhibits the following features

to distinguish them from events originating from SM processes:

- two non-b-quark jets with high p_{T} from the decays of the strongly interacting SUSY particles into the much lighter neutralinos/charginos,
- two b-quark jets from the h_1 decay,
- large $E_{\mathrm{T}}^{\mathrm{miss}}$ due to the two LSPs at the end of each cascade-decay chain.

The analysis strategy is to select events with at least two b-tagged jets according to the criteria described in Section 5.3, and to compute the invariant mass m_{bb} of the two selected jets. The contributions expected from SM processes, considered as background in the signal region, are derived from simulation, or by using a data-driven method in case of the QCD multijet background. The m_{bb} distributions estimated for the different background sources are then linearly combined together with the signal distribution, which is derived from simulation for various mass hypotheses, and fitted to the data. The contribution from h_1 would then manifest itself as a peak at the mass of h_1.

In order to avoid a possible bias in the results, a blinding policy is adopted during the development phase of the analysis. The real data m_{bb} distribution is never looked at in the signal region for invariant masses below 150 GeV. Instead, background-enriched control regions with very small expected signal contributions are used to verify the modelling of the background.

5.1.1 Contributions from Supersymmetric Cascades

In addition to the h_1 resonance, there are further contributions from the SUSY cascade in the NMSSM scenarios considered for this analysis (see Section 1.4.2). All of them are depicted in Fig. 5.1 and comprise the following sources of b-quark pairs: The top left diagram in Fig. 5.1 illustrates the production of NMSSM Higgs bosons, i.e. h_1, h_2, and a_1. Most similar to the Higgs boson production is the creation of a Z^0 boson which decays into a pair of b quarks as depicted in Fig. 5.1 (top right).

However, when selecting two b-tagged jets there are also non-resonant contributions as illustrated in Fig. 5.1 (bottom left), where two top-quarks are produced in a gluino decay, which subsequently decay into $\mathrm{W}^{\pm}$ bosons and two b quarks. Another source of non-resonant SUSY events in agreement with the signal topology is shown in Fig. 5.1 (bottom right) where the decay products of two different sources would be combined, e.g. events where the selected jets originate from Higgs boson decays, but not from the same one.

5.1.2 Background from Standard Model Processes

The SM processes, that have similarities with the signal topology described in this section are, ordered by decreasing size of their contribution, $\mathrm{t\bar{t}}$ production in association with jets, QCD multijet, $\mathrm{W} \to \ell\nu$ associated with jets, and $\mathrm{Z}^0 \to \nu\bar{\nu}$ associated with jets. Sizeable $E_{\mathrm{T}}^{\mathrm{miss}}$ in $\mathrm{t\bar{t}}$ events is caused by leptonic decays of the W bosons

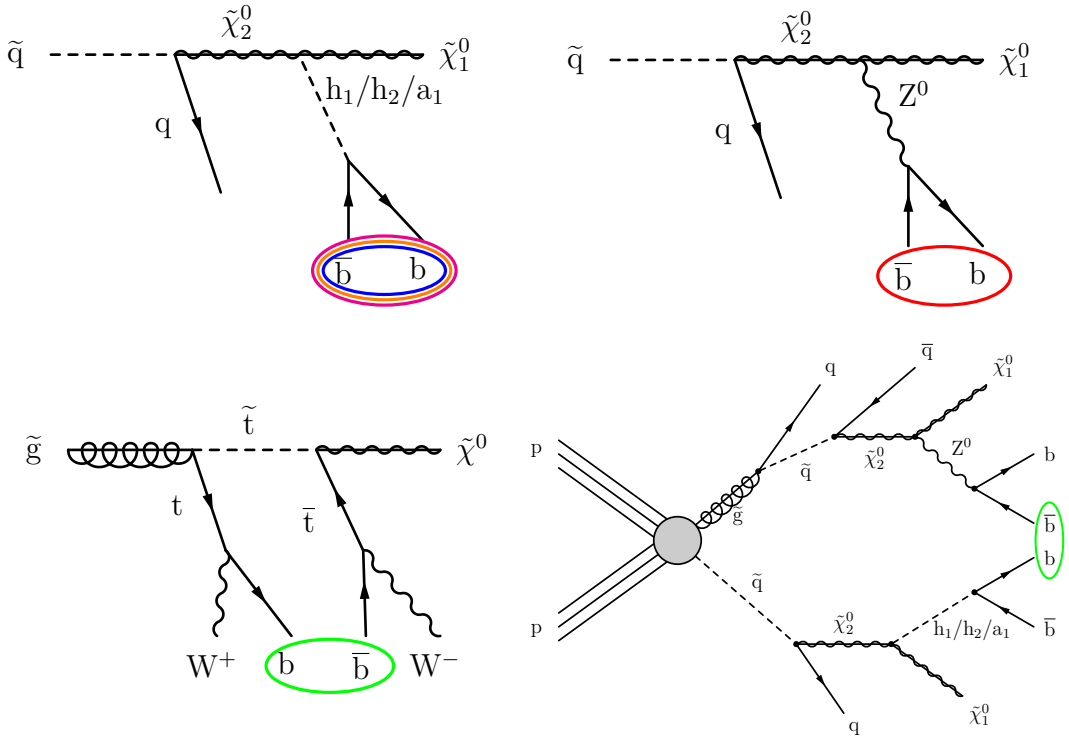

Figure 5.1: Example Feynman diagrams illustrating the NMSSM contributions from Higgs resonances (top left), the Z^0 resonance (top right), and non-resonant contributions (bottom). The ellipses indicate the selected pair of b quarks.

produced in a t-quark decay. Therefore, di-leptonic and semi-leptonic decays of the $t\bar{t}$ system, i.e. with two or one lepton in the final state, are the main contribution due to neutrinos produced in association with the leptons. In principle, QCD multijet events exhibit at most small E_T^{miss}. However, mismeasurement of jet energies can result in sizeable missing energy. In these events the $\vec{E}_T^{miss}$ vector is typically aligned with one of the two most energetic jets which is used to estimate this contribution as explained in detail in Section 5.4.2. For the remaining two processes, the E_T^{miss} is caused by the undetected neutrinos, but their overall contribution is highly reduced by the requirement of two b-tagged jets.

Further SM backgrounds are tested, as described in Section 5.4, but found to be negligible. Thus, they are not included in the final result.

5.2 Data and Simulated Samples

The data used in this analysis correspond to an integrated luminosity of $19.7\,fb^{-1}$, collected by the CMS experiment at a centre-of-mass energy of 8 TeV during the 2012 data-taking period.

The signal process is characterised by the associated production of many jets (see Fig. 1.7), which facilitates the online selection with a trigger that uses the H_T

variable. Such a trigger is used to collect the data for this analysis. Only the events certified by CMS to have good quality (see Section 3.2.6) are used for the analysis. For the trigger efficiency determination a control trigger on isolated muons is used.

The NMSSM signal samples listed in Table 5.1 are generated in the two scenarios described in Section 1.4.2. The event generator PYTHIA and the full detector simulation are used for the data sets of the modified P4 scenario, whereas MADGRAPH and the fast simulation are used for the decoupled squarks scenario. The NMSSM-specific input is created with NMSSMTOOLS [115–121] and inserted in the event generation process using the SLHA interface [125, 126].

Table 5.1: Signal samples used in the analysis. For each sample the respective NMSSM scenario and h_1 mass hypothesis are listed together with the number of generated events. The samples in the modified P4 scenario are generated using the full detector simulation, whereas the sample in the decoupled squarks scenario uses the fast simulation (see Section 3.3.3). In addition, the latter sample contains events generated for various settings of the NMSSM parameters m_{h_1}, M_1, M_2, and M_3, as listed in Table 1.10.

Scenario	m_{h_1}	# events
Modified P4 scenario	30 GeV	125 012
	35 GeV	125 012
	40 GeV	490 400
	45 GeV	484 530
	50 GeV	466 170
	55 GeV	495 550
	60 GeV	498 086
	65 GeV	995 032
	70 GeV	921 064
	75 GeV	993 007
	80 GeV	978 616
	85 GeV	987 182
	90 GeV	250 025
	95 GeV	249 209
	100 GeV	249 509
Decoupled squarks scenario	60 – 95 GeV	9 058 666

The SM background samples are listed in Table 5.2. The cross sections are calculated at next-to-leading order (NLO) or next-to-NLO (NNLO) according to the best precision available. The QCD multijet cross sections correspond to LO, but are not used for the extraction of the final result.

Table 5.2: Simulated samples of the SM backgrounds considered in this analysis. For each sample, the event generator, the cross section, and the integrated luminosity corresponding to the number of generated events in this sample are listed. Parton showering and hadronisation are done with PYTHIA.

Sample	Event Generator	σ (pb)	$\int \mathcal{L}$ (fb^{-1})
Hadronic $t\bar{t}$ + jets	MADGRAPH	102.9	303
Semi-leptonic $t\bar{t}$ + jets	MADGRAPH +TAUOLA	98.6	253
Di-leptonic $t\bar{t}$ + jets	MADGRAPH +TAUOLA	23.6	508
Multijet QCD, $H_T = 100-250$ GeV	MADGRAPH	1.06×10^7	4.47×10^{-3}
Multijet QCD, $H_T = 250-500$ GeV	MADGRAPH	2.85×10^5	9.50×10^{-2}
Multijet QCD, $H_T = 500-1000$ GeV	MADGRAPH	8710	3.51
Multijet QCD, $H_T > 1000$ GeV	MADGRAPH	212	65.3
$W \to \ell\nu$ + 1 jet	MADGRAPH	6440	3.59
$W \to \ell\nu$ + 2 jets	MADGRAPH	2090	16.3
$W \to \ell\nu$ + 3 jets	MADGRAPH	619	25.1
$W \to \ell\nu$ + 4 jets	MADGRAPH	255	52.5
$Z^0 \to \nu\bar{\nu}$, $H_T = 50-100$ GeV	MADGRAPH	381	52.5
$Z^0 \to \nu\bar{\nu}$, $H_T = 100-200$ GeV	MADGRAPH	160	34.8
$Z^0 \to \nu\bar{\nu}$, $H_T = 200-400$ GeV	MADGRAPH	41.5	113
$Z^0 \to \nu\bar{\nu}$, $H_T > 400$ GeV	MADGRAPH	5.27	775
Single t production (tW channel)	POWHEG +TAUOLA	11.8	42.3
Single $\bar{t}$ production (tW channel)	POWHEG +TAUOLA	11.8	41.9
Single t production (t channel)	POWHEG +TAUOLA	55.5	67.7
Single $\bar{t}$ production (t channel)	POWHEG +TAUOLA	30.0	65.5
Single t production (s channel)	POWHEG +TAUOLA	3.89	66.8
Single $\bar{t}$ production (s channel)	POWHEG +TAUOLA	1.76	79.5
WZ^0	PYTHIA +TAUOLA	32.3	309
Z^0Z^0	PYTHIA +TAUOLA	7.6	1290
$Z^0/\gamma^* \to \ell\ell$, $M_{Z^0/\gamma^*} = 10-50$ GeV	MADGRAPH	907	7.86
$Z^0/\gamma^* \to \ell\ell$, $M_{Z^0/\gamma^*} > 50$ GeV	MADGRAPH	3510	8.68

5.3 Event Selection

The search for a light Higgs boson is performed on the data set taken in 2012 by CMS. Certain criteria are applied during data taking and in the offline analysis to select data enriched with events exhibiting the signal topology, while reducing the fraction of background-like events. This section describes the trigger used for this task, the event-wise cleaning and correction procedures, and the offline selection.

5.3.1 Online Event Selection

The signal events from SUSY cascades have a large hadronic activity due to the very energetic jets from the squark decays and the two b jets from the h_1 decay. This facilitates triggering using HLT paths that use H_T as defined in Section 3.4.3 and calculated with Particle Flow (PF) objects at trigger level. In the second half of 2012 an improved version is used to better mitigate PU effects. Both versions compute the hadronic activity H_T and select events if H_T is greater than a threshold of 650 GeV.

Because of the limited processing time per event during the online selection, the PF algorithm at trigger level is simplified with respect to the offline version. Therefore, the efficiencies of the online and offline selection using the H_T variable differ from each other. This difference is expressed in terms of trigger efficiencies $\epsilon_{\text{Trigger}}$ as a function of H_T computed with the offline PF algorithm. The trigger efficiency is calculated with the *tag-and-probe* method, i.e. an independent sample is selected by means of a *tag* object. In this analysis an uncorrelated trigger on isolated muons is used to compute the efficiency of the *probe* object H_T,

$$\epsilon_{\text{Trigger}} = \frac{N_{\text{tag+probe}}}{N_{\text{tag}}} \ . \tag{5.1}$$

Here, N_{tag} is the number of events selected by the trigger on isolated muons, and $N_{\text{tag+probe}}$ is the number of events that are also selected by the H_T trigger. This quantity is computed in bins of different offline H_T, and the results are visualised in Fig. 5.2. In addition to the trigger selection, an E_T^{miss} requirement of 200 GeV is applied, which is also used later in the offline analysis.

The offline H_T in Fig. 5.2 is defined as the scalar sum of all PF jets passing $p_T > 40$ GeV and $|\eta| < 3.0$, which is the same selection as on trigger level but using jets obtained with the simplified PF algorithm. The efficiencies in Fig. 5.2 are fitted to the data using the function

$$\epsilon_{\text{Trigger}}(H_T) = \frac{N}{2}\left[\text{erf}(s \cdot H_T - t) + 1\right] \ , \tag{5.2}$$

where the fitted parameters are the normalisation N, the position of the turn-on t, and the slope of the turn-on s. The Gauss error function erf is defined as

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} \mathrm{d}t \ . \tag{5.3}$$

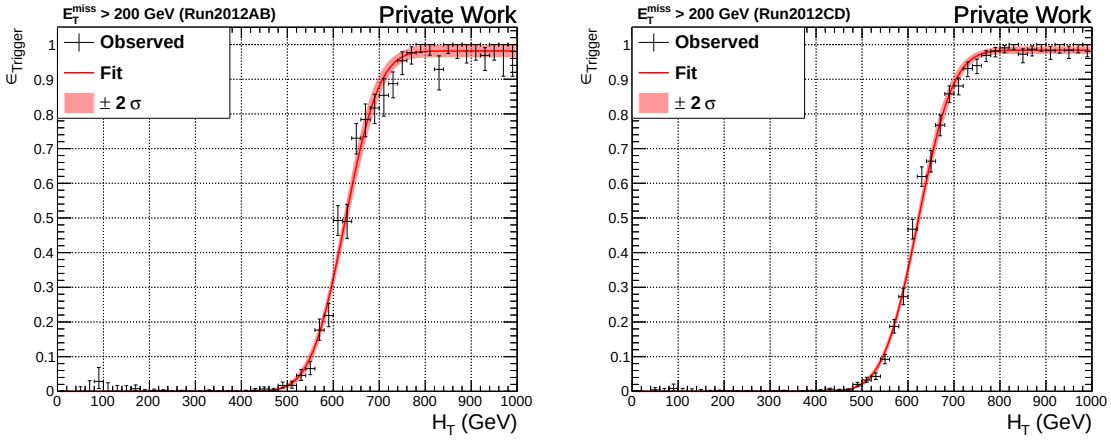

Figure 5.2: Efficiencies as a function of offline H_T for the H_T-based triggers used in the first part of the 2012 data-taking period (left) and in the second part of 2012 (right). The efficiencies are measured in a data set collected with a trigger on isolated muons and requiring $E_T^{miss} > 200\,GeV$.

The fit results are listed in Table 5.3, and from Fig. 5.2 one can conclude that the triggers are fully efficient for an offline H_T of 750 GeV. Thus, an offline H_T of at least 750 GeV is required to select events in data and simulation, where no trigger selection is applied, with the same efficiency. Omitting the online selection in the simulation is necessary because the data sets are simulated prior to data taking, thus the final trigger menu could not be included. The uncertainty on the measured efficiency, indicated by the red band enclosing the fitted curve in Fig. 5.2, is propagated to the final result as uncertainty on the trigger weight in simulated events.

Table 5.3: Results of the fit of the parameterisation of $\epsilon_{Trigger}(H_T)$ as specified in (5.2) to the trigger efficiencies in Fig. 5.2 for the two different trigger versions used in 2012.

Data-Taking Period	N	t	s
First part of 2012	$(98.2 \pm 1.0)\,\%$	$(626 \pm 3)\,GeV$	$(1.20 \pm 0.06) \times 10^{-2}$
Second part of 2012	$(98.4 \pm 0.4)\,\%$	$(622 \pm 2)\,GeV$	$(1.18 \pm 0.03) \times 10^{-2}$

Other trigger paths are studied, in particular triggers based on b tagging. However, the respective triggers available in the CMS online menu test only the four leading jets, which are often not the ones from the h_1 decay as can be seen from the p_T rank of the jets selected in the offline analysis in Fig. 5.3 (top). They are thus poorly efficient for this analysis and are not considered in the following.

5.3.2 Event Cleaning and Data-to-Simulation Corrections

Real data events are affected by detector noise or mismeasurements. These effects can lead to spurious E_T^{miss} such that an event with no genuine E_T^{miss} is selected in searches

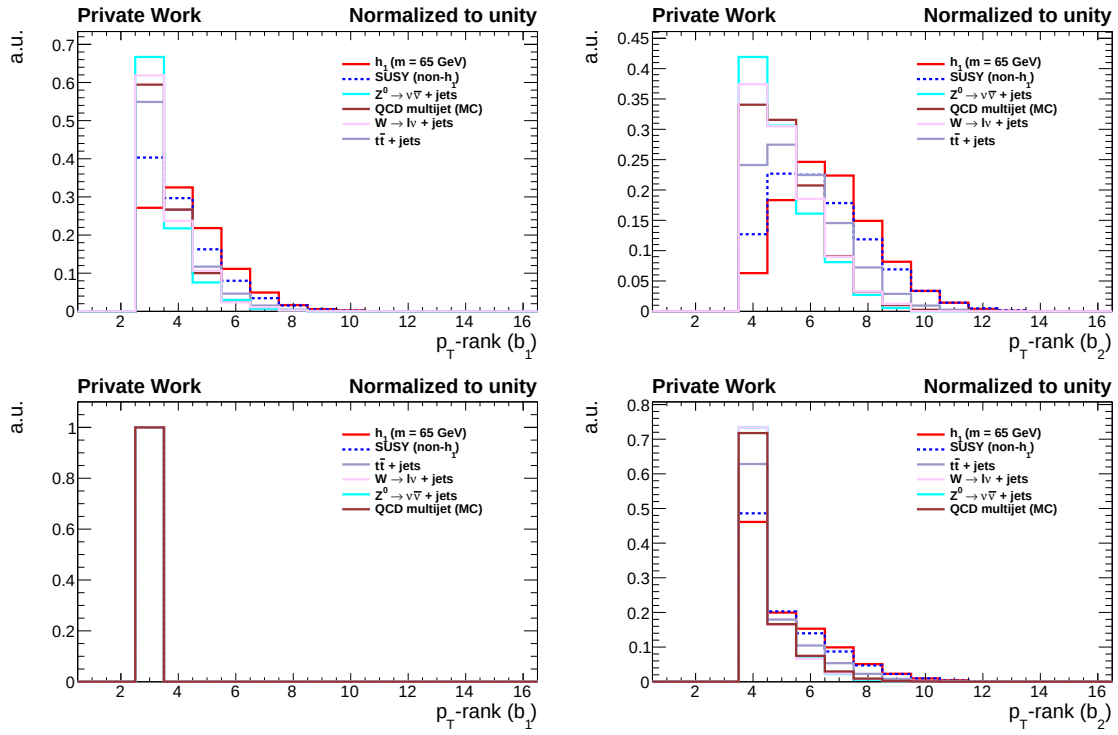

Figure 5.3: Rank in p_T of the harder (left) and the softer selected jet (right) for the various contributions. The distributions are shown after a selection with at least two b-tagged jets (top) and after a selection with exactly one b-tagged jet (bottom). In the latter case the second jet is chosen to be the third non-b-tagged jet, but is otherwise treated as the usual second b-tagged jet b_2. The single-b selection is used in the QCD multijet background determination. Due to a b veto on the two leading jets, the entries start only at 3.

with signatures as the one described in Section 5.1, which contain weakly interacting particles in the final states. Thus, a removal of false-$E_\mathrm{T}^\mathrm{miss}$ events is necessary to mitigate these effects. In addition to the filtering of these pathological events, event-based corrections have to be applied to simulation. Those are needed because certain simulation input is only available after data taking, or because the performance of identification algorithms differs between data and simulation. These event filters and corrections are described in this section.

Removal of Spurious $\boldsymbol{E}_\mathbf{T}^\mathbf{miss}$ Events

To discard events with large false $E_\mathrm{T}^\mathrm{miss}$, a variety of event filters is developed by the CMS collaboration [143, 188, 211, 212]. The cleaning requirements and algorithms used in this analysis are described in the following.

To ensure the presence of a pp collision, all events must contain a reconstructed primary vertex according to the criteria listed in Section 3.4.3. Events from beam-background processes, e.g. interactions of the protons with residual gas molecules, are rejected by requiring at least 25 % tracks of an event to be of *high purity*, i.e. these

tracks must fulfil certain quality criteria. Beam-background processes with a muon are identified with a dedicated filter using information from the CSC subdetector (see Section 3.2.4).

Events affected by instrumental noise are rejected by the standard noise filters described in the following. Instrumental noise in the HPDs (see Section 3.2.3) and their readout boxes causes anomalous signals in the HCAL. Such events are identified and discarded based on timing and pulse-shape information. Approximately 1 % of the ECAL crystals suffer from electronics noise and are ignored in the readout. However, a significant amount of energy may be lost leading to high spurious $E_{\mathrm{T}}^{\mathrm{miss}}$. Those events are vetoed based on the distance between the masked cells and jets together with the energy surrounding the masked ECAL cells. Another noise filter removes such events if the trigger primitive E_{T} exceeds a certain threshold.

Events have been observed with large ECAL and HCAL deposits, but with small track multiplicities. In those events, the tracking algorithm fails for some of its iterations because of too many seed clusters. Such events are removed by requiring the scalar p_{T}-sum of the tracks associated with the primary vertex to be greater than 10 % of the scalar p_{T}-sum of all jets within the tracker acceptance.

The filters described above are all applied on real data. The noise filters are also applied on simulated events, though the effect is minor.

Pileup Reweighting

As stated in Section 3.4.3, on average 21 PU interactions are observed per bunch crossing during the 2012 data-taking period. In simulated data, the hard-scattering process is overlaid with minimum bias interactions to mimic PU conditions as present in data. However, since the MC samples were generated in advance of the data taking, there is only a rough agreement between data and simulation with respect to PU conditions. Therefore, the simulation samples have to be reweighed such that the distribution of the number of PU interactions matches the distribution observed in data [213].

The PU modelling is validated in the signal region, but without any $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement. Figure 5.4 compares the predicted distributions of reconstructed primary vertices with the one measured in data, which corresponds to the number of PU interactions multiplied with the primary-vertex-reconstruction efficiency. Good agreement is observed between the data and simulated events after reweighting.

b-Tagging-Efficiency Corrections

The performance of the b-tagging algorithms differs between real data and simulated events (see Fig. 3.13), and between full simulation and fast simulation of the CMS detector response. In order to account for this deviation, a reweighting procedure based on b-tagging-efficiency scale-factors is applied. These scale factors are defined as the ratio of the b-tagging efficiency in data and simulation, and full and fast simulation, respectively.

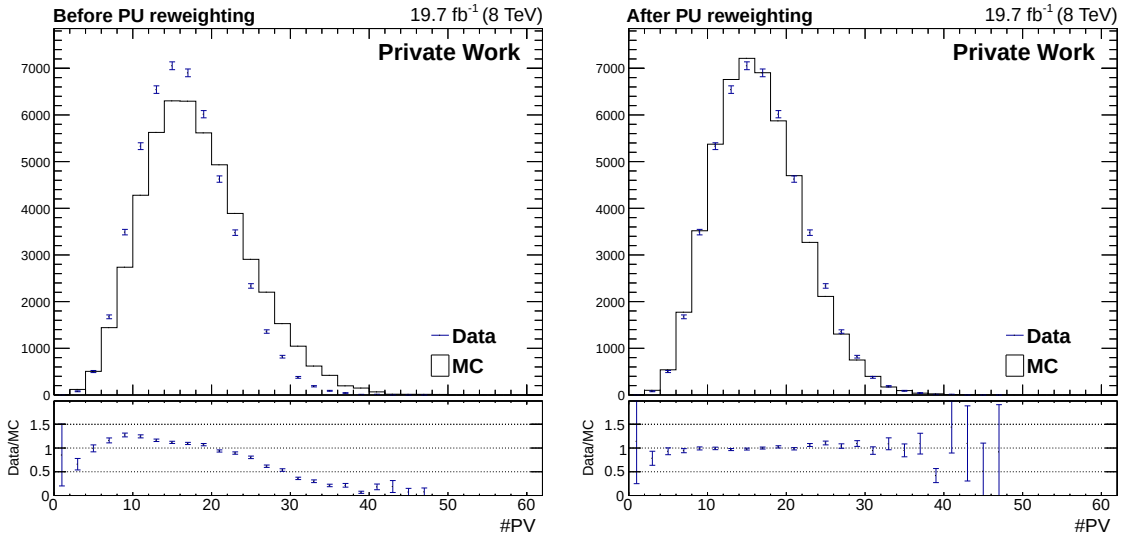

Figure 5.4: Number of reconstructed primary vertices before (left) and after (right) PU reweighting. In order to compare only the shape, the simulation is normalised to the number of data events taken in 2012. The full offline selection (see Section 5.3.3) is applied, but the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement is removed to increase the statistical precision in this validation of the reweighting procedure.

Reweighting of Full Simulation to Data A reweighting is applied to all simulated samples (see Tables 5.1, 5.2), based on scale factors derived from different topologies, e.g. $\mathrm{t\bar{t}}$ and QCD multijet events [214]. The scale factors for this procedure depend on p_{T}, η, and the true flavour of the simulated jet. The scale factors for heavy-flavour jets (b, c) are the same, but with twice the uncertainty for c-quark jets. All other types of jets (u, d, s, g) use common "mis-tag" scale factors:

$$\mathrm{SF}(\epsilon_{\mathrm{offl}}^{\text{b-tag}}) = 0.95 \pm 0.01 \,, \tag{5.4a}$$

$$\mathrm{SF}(\epsilon_{\mathrm{offl}}^{\text{mis-tag}}) = 1.2 \pm 0.2 \,, \tag{5.4b}$$

where the numbers given are typical values for the medium working point of the CSV (CSVM) tagger [187].

Reweighting of Full Simulation to Fast Simulation CMS provides p_{T}-, η-, and flavour-dependent correction factors, which are defined as the ratio of full simulation over fast simulation b-tagging efficiencies [215]. These correction factors are derived individually for b-quark, c-quark, and "light"-flavour (u, d, s, g) jets.

Overall b-tagging efficiencies in the fast simulation are higher, since various detector effects are not fully incorporated. The correction factors are applied to the fast simulation signal sample in the decoupled squarks scenario (see Table 5.1) analogous to the procedure for the data-to-simulation reweighting in order to match the performance in fast simulation to full simulation. Typical values for these corrections

are [216]

$$\mathrm{CF}(\epsilon^{\mathrm{b}}) = 0.992 \pm 0.013 \,, \tag{5.5}$$

$$\mathrm{CF}(\epsilon^{\mathrm{c}}) = 0.991 \pm 0.022 \,, \tag{5.6}$$

$$\mathrm{CF}(\epsilon^{\mathrm{udsg}}) = 1.094 \pm 0.116 \,, \tag{5.7}$$

where the jet p_{T} is within 80 – 100 GeV and the light-flavour jets are central with $|\eta| < 1.2$.

5.3.3 Offline Event Selection

As discussed in Section 5.3.1, the large hadronic activity in signal events as described in Section 5.1, motivates the use of the variable H_{T} to enrich events with the desired topology. To ensure a fully efficient trigger, the offline selection of this analysis requires $H_{\mathrm{T}} > 750\,\mathrm{GeV}$, which is the beginning of the trigger-efficiency plateau in Fig. 5.2. The medium working point of the CSV b-tagging algorithm (CSVM) is used in the analysis. Each event must have two leading non-b-tagged jets with $p_{\mathrm{T}}^{\mathrm{Lead}} > 250\,\mathrm{GeV}$ and $p_{\mathrm{T}}^{\mathrm{SubLead}} > 100\,\mathrm{GeV}$, respectively. Moreover, two b-tagged jets are required, which can be of any rank in p_{T}, except the first and the second rank. Additional jets are not vetoed. If not stated otherwise, all jets have to fulfil $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.4$.

If more than two jets are identified as b-quark jets, the pair with the smallest $\Delta R(\mathrm{b},\mathrm{b})$ is chosen, where $\min \Delta R(\mathrm{b},\mathrm{b}) < 1.5$ is required. In addition, the event must fulfil $E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$, and the azimuthal angle between $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ and each of the leading jets has to be $\Delta\phi(j_{1,2}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}}) > 0.5$ in order to suppress QCD multijet contributions due to mismeasurement of the jet energy.

The invariant mass of the two selected b-tagged jets is then computed, and used to extract the final results.

The event selection variables are shown in Appendix D.1, and a comparison of data to the predicted distributions shows that they are well described.

Optimisation of Selection Criteria

The selection criteria are the result of a detailed optimisation technique targeting the highest sensitivity reflected in the most stringent expected upper limit in the case of non-observation. The procedure employs the signal samples of the modified P4 scenario listed in Table 5.1.

The expected limits for the optimisation studies are derived for three different mass points ($m_{\mathrm{h}_1} = 40$, 65 and 80 GeV), representative for the entire probed h_1 mass range 30 – 100 GeV. The procedure explained in Section 5.6 is used, but with a subset of the systematic uncertainties enumerated in Section 5.7. The detailed results of these studies are summarised in Appendix D.2 and lead to the above described selection criteria.

Selection Efficiency for Data and Signal Events

The values of the selection efficiency are listed in detail for data in Table 5.4. The largest effect on the rate have the requirements on the multiplicity of b-tagged jets and on $E_{\mathrm{T}}^{\mathrm{miss}}$.

Table 5.4: Number of events observed in data after each selection step. The efficiency of each selection step relative to the previous step is also listed. In the first line, "all" means the events which pass the trigger selection.

	events	efficiency
all	43111487	1.000
$H_{\mathrm{T}} > 750\,\mathrm{GeV}$	26773080	0.621
#non-b-tagged jets ≥ 2	26392661	0.986
#jets ≥ 4	17776856	0.674
#b-tagged jets ≥ 2	705797	0.040
b veto for leading two jets	230730	0.327
$p_{\mathrm{T}}^{\mathrm{Lead}} > 250\,\mathrm{GeV}$, $p_{\mathrm{T}}^{\mathrm{SubLead}} > 100\,\mathrm{GeV}$	196059	0.850
$\Delta\phi(j_1, \vec{E}_{\mathrm{T}}^{\mathrm{miss}}) > 0.5$	160955	0.821
$\min \Delta R(\mathrm{b}_1, \mathrm{b}_2) < 1.5$	71538	0.444
$\Delta\phi(j_2, \vec{E}_{\mathrm{T}}^{\mathrm{miss}}) > 0.5$	50933	0.712
$E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$	148	0.003
total		0.000003

For selected h_1-mass points in the modified P4 scenario the expected event yields and efficiencies are listed for events with an h_1 boson decaying into a pair of b-quarks ($\mathrm{h}_1 \to \mathrm{b\bar{b}}$) in Table 5.5 and for all NMSSM contributions containing a b-quark pair (NMSSM $\to \mathrm{b\bar{b}}$) in Table 5.6.

The total signal selection efficiencies for the processes $\mathrm{h}_1 \to \mathrm{b\bar{b}}$ and NMSSM $\to \mathrm{b\bar{b}}$ are defined as

$$\epsilon(\mathrm{h}_1 \to \mathrm{b\bar{b}}) = \frac{\#\text{ events where selected b-jets originate from } \mathrm{h}_1}{\#\text{ NMSSM events with } \mathrm{h}_1 \to \mathrm{b\bar{b}}}\,, \tag{5.8a}$$

$$\epsilon(\mathrm{NMSSM} \to \mathrm{b\bar{b}}) = \frac{\#\text{ selected NMSSM events with} \geq 2\text{ b-quarks}}{\#\text{ NMSSM events with} \geq 2\text{ b-quarks}}\,. \tag{5.8b}$$

These efficiencies are shown in Fig. 5.5 as a function of the h_1-boson mass.

In case of the modified P4 scenario, the efficiency $\epsilon(\mathrm{h}_1 \to \mathrm{b\bar{b}})$ exhibits a mass dependence (see Table 5.7), driven by the efficiency of the matching between the selected jets and an $\mathrm{h}_1 \to \mathrm{b\bar{b}}$ decay (see Fig. 5.5 (left), Table 5.5). The mass dependence of the matching is caused by the usually low p_{T} of the second jet matched to an $\mathrm{h}_1 \to \mathrm{b\bar{b}}$ decay (see Fig. 5.6, top right). The general $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ requirement removes a sizeable fraction of events because the second b quark does not pass this threshold.

The mass dependence of $\epsilon(\mathrm{NMSSM} \to \mathrm{b\bar{b}})$ is reduced with respect to $\epsilon(\mathrm{h}_1 \to \mathrm{b\bar{b}})$

Table 5.5: Expected $h_1 \to b\bar{b}$ event yields for three different mass points in the modified P4 scenario after each selection step. The yields are normalised to the integrated luminosity of $19.7\,\text{fb}^{-1}$ assuming the cross sections listed in Table 1.7. The efficiency of each selection step relative to the previous step is also listed as well as the h_1 matching efficiency, i.e. the efficiency of requiring that both selected b-tagged jets originate from the same h_1. In the first line, "all" means the events which pass the trigger selection.

	$m_{h_1} = 40\,\text{GeV}$		$m_{h_1} = 65\,\text{GeV}$		$m_{h_1} = 85\,\text{GeV}$	
	events	eff.	events	eff.	events	eff.
all	1096.7	1.000	554.0	1.000	234.9	1.000
$H_T > 750\,\text{GeV}$	1071.6	0.977	541.6	0.978	229.0	0.975
#non-b-tagged jets ≥ 2	1068.4	0.997	540.2	0.997	228.5	0.998
#jets ≥ 4	1058.3	0.991	535.6	0.992	227.0	0.993
#b-tagged jets ≥ 2	507.4	0.479	253.2	0.473	114.1	0.503
b veto for leading two jets	289.7	0.571	157.7	0.623	72.0	0.631
$p_T^{Lead} > 250\,\text{GeV}$, $p_T^{SubLead} > 100\,\text{GeV}$	263.1	0.908	143.4	0.909	65.0	0.903
$\Delta\phi(j_1, \vec{E}_T^{miss}) > 0.5$	252.2	0.959	138.6	0.967	62.1	0.955
$\min \Delta R(b_1, b_2) < 1.5$	171.3	0.679	92.5	0.667	38.2	0.615
$\Delta\phi(j_2, \vec{E}_T^{miss}) > 0.5$	149.3	0.871	81.5	0.881	33.4	0.874
$E_T^{miss} > 200\,\text{GeV}$	103.9	0.696	58.2	0.714	22.7	0.681
total (excl. h_1 matching)		0.095		0.105		0.097
h_1 matching	27.0	0.260	22.7	0.391	12.0	0.529
total (incl. h_1 matching)		0.025		0.041		0.051

because the selected jets are not required to originate from the same particle, but can be any b-tagged jets (see Fig. 5.5 (left), Table 5.5). This is confirmed by the blue dashed line in Fig. 5.6 (top right) showing a more energetic spectrum than for the jets that require an h_1 matching (red solid line).

In general, the requirement of two b-tagged jets has the lowest efficiency. The reason for this is not only the b-tagging efficiency, but mainly the p_T spectrum of the softer b-quark. As explained above, this effect is less strong if the origin of the b-tagged jet is neglected, i.e. if no h_1 matching is required. In addition, it is shown in Fig. D.5 that the loose working point of the CSV discriminator does not improve the sensitivity of the analysis because of the significantly increasing background.

An additional mass-dependent trend in the efficiencies is driven by the decreasing efficiency of the b veto on the two leading jets (see Tables 5.5 and 5.6). Below 65 GeV the average multiplicity of b quarks per event increases because of the decay chain from an h_2 boson into a pair of h_1 bosons, each decaying into a pair of b quarks. This h_2-decay mode is dominant for low h_1-masses ($m_{h_1} < m_{h_2}/2$), thus leading to a higher b-quark multiplicity and a higher probability of a b-quark jet among the

Table 5.6: Expected NMSSM $\rightarrow b\bar{b}$ event yields for three different mass points in the modified P4 scenario after each selection step. The yields are normalised to the integrated luminosity of $19.7\,\mathrm{fb}^{-1}$ assuming the cross sections listed in Table 1.7. The efficiency of each selection step relative to the previous step is also listed. In the first line, "all" means the events which pass the trigger selection.

	$m_{h_1} = 40\,\mathrm{GeV}$		$m_{h_1} = 65\,\mathrm{GeV}$		$m_{h_1} = 85\,\mathrm{GeV}$	
	events	eff.	events	eff.	events	eff.
all	1540	1.000	1363	1.000	1198	1.000
$H_T > 750\,\mathrm{GeV}$	1500	0.974	1328	0.974	1167	0.974
#non-b-tagged jets ≥ 2	1496	0.997	1325	0.998	1164	0.998
#jets ≥ 4	1481	0.990	1314	0.992	1154	0.991
#b-tagged jets ≥ 2	674	0.455	585	0.445	508	0.440
b veto for leading two jets	385	0.571	350	0.598	303	0.595
$p_T^{Lead} > 250\,\mathrm{GeV}$, $p_T^{SubLead} > 100\,\mathrm{GeV}$	349	0.905	316	0.903	272	0.899
$\Delta\phi(j_1, \vec{E}_T^{miss}) > 0.5$	334	0.959	304	0.962	260	0.958
$\min \Delta R(b_1, b_2) < 1.5$	216	0.645	185	0.611	152	0.585
$\Delta\phi(j_2, \vec{E}_T^{miss}) > 0.5$	188	0.871	162	0.876	133	0.869
$E_T^{miss} > 200\,\mathrm{GeV}$	132	0.701	115	0.708	93	0.704
total		0.085		0.084		0.078

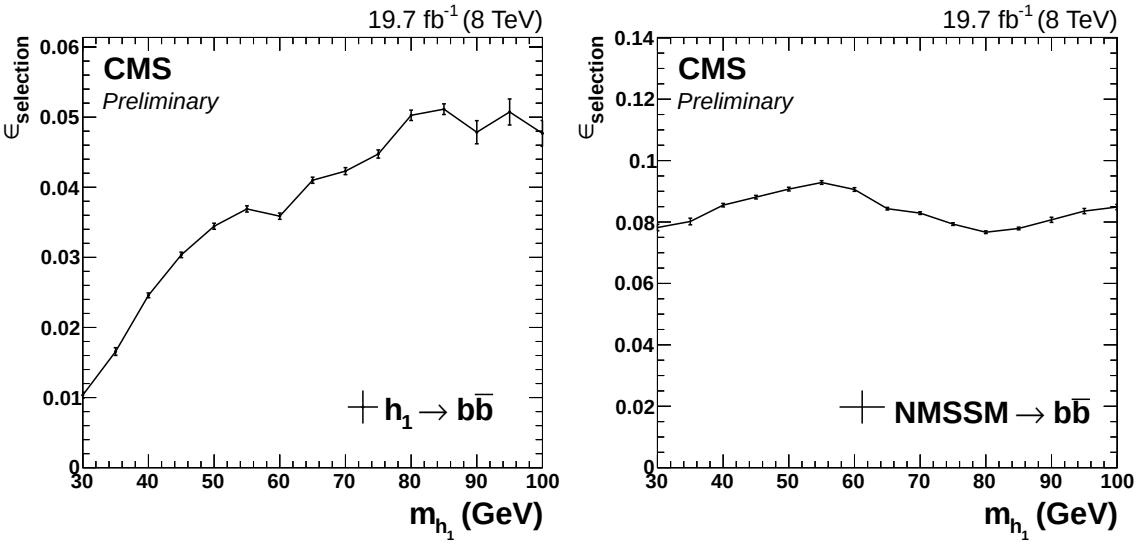

Figure 5.5: The left panel shows the selection efficiency of the $h_1 \rightarrow b\bar{b}$ events including h_1 matching efficiency, i.e. both selected jets stem from the same h_1. The right panel shows the selection efficiency of all NMSSM $\rightarrow b\bar{b}$ events.

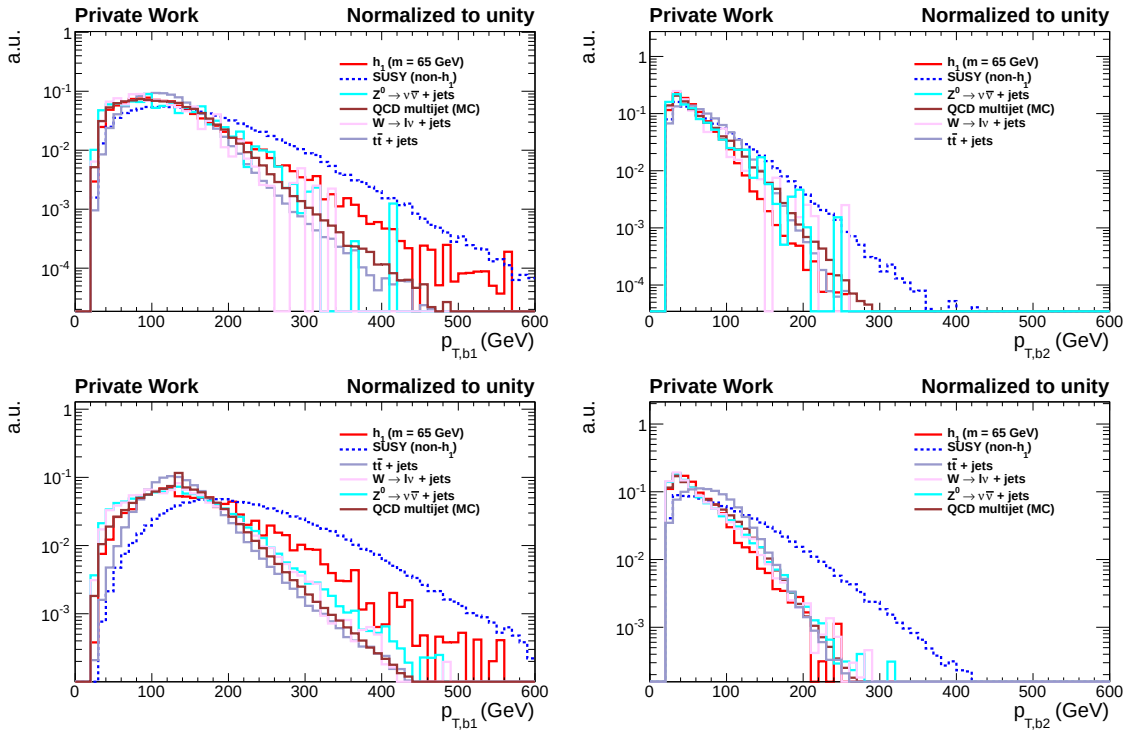

Figure 5.6: p_{T} spectra of the harder selected jet (left) and the softer selected jet (right) for the various contributions. The distributions are shown after a selection with at least two b-tagged jets (top) and after a selection with exactly one b-tagged jet (bottom). In the latter case the second jet is chosen to be the third non-b-tagged jet, but is otherwise treated as the usual second b-tagged jet b_2. The single-b selection is used in the QCD multijet background determination. The p_{T} spectrum of the softer jet is falling, starting from the general jet-p_{T} requirement of 25 GeV. This means for the signal that one b can be lost because it does not pass the general p_{T} requirement.

leading two jets.

The drop at $m_{\mathrm{h}_1} = 60\,\mathrm{GeV}$ in Fig. 5.5 (left) is a joint effect of individual steps in the event selection, most notably the increase in b-quark-jet multiplicity upon the opening of the $\mathrm{h}_2 \to \mathrm{h}_1\mathrm{h}_1$ channel, which influences the effect of the b veto on the two leading jets as described above.

The selection efficiencies in the decoupled squarks scenario are summarised in Fig. 5.7 and a detailed table of selection efficiencies can be found in Table 5.7 for a few representative parameter points.

The efficiencies in both scenarios can be approximately compared using the right plot in Fig. 5.5 and the $M_3 = 1000\,\mathrm{GeV}$-row of the central right plot in Fig. 5.7 because the gaugino mass parameters in the latter plot are almost the same as in the modified P4 scenario (see Table 1.6). For the overlapping h_1 mass range the total efficiencies are lower in the decoupled squarks scenario because of the different typical event topology with higher relative non-resonant contributions in this scenario, especially from decay chains as shown in Fig. 5.1 (bottom left). The b-quark jets in such events

Table 5.7: Expected NMSSM $\to \mathrm{b\bar{b}}$ event yields in the decoupled squarks scenario for four different NMSSM parameter points after each selection step. The yields are normalised to the integrated luminosity of $19.7\,\mathrm{fb}^{-1}$ assuming the cross sections listed in Table 1.11. The efficiency of each selection step relative to the previous step is also listed. The values of the gaugino mass parameters are $M_1 = 350\,\mathrm{GeV}$ and $M_2 = 350\,\mathrm{GeV}$, and the mass of h_1 is $m_{\mathrm{h}_1} = 80\,\mathrm{GeV}$. In the first line, "all" means the events which pass the trigger selection.

	$M_3 = 0.8\,\mathrm{TeV}$		$M_3 = 1\,\mathrm{TeV}$		$M_3 = 1.2\,\mathrm{TeV}$		$M_3 = 1.4\,\mathrm{TeV}$	
	events	eff.	events	eff.	events	eff.	events	eff.
all	306.75	1.000	49.33	1.000	8.90	1.000	1.79	1.000
$H_\mathrm{T} > 750\,\mathrm{GeV}$	260.34	0.849	47.30	0.959	8.80	0.988	1.78	0.996
#non-b-tagged jets ≥ 2	259.82	0.998	47.18	0.997	8.77	0.997	1.78	0.997
#jets ≥ 4	259.73	1.000	47.16	0.999	8.77	1.000	1.78	1.000
#b-tagged jets ≥ 2	179.38	0.691	33.22	0.705	6.16	0.702	1.23	0.692
b veto for leading two jets	66.53	0.371	12.00	0.361	2.24	0.364	0.46	0.377
$p_\mathrm{T}^\mathrm{Lead} > 250\,\mathrm{GeV}$, $p_\mathrm{T}^\mathrm{SubLead} > 100\,\mathrm{GeV}$	42.13	0.633	9.46	0.788	1.96	0.875	0.42	0.897
$\Delta\phi(j_1, \vec{E}_\mathrm{T}^\mathrm{miss}) > 0.5$	39.67	0.942	8.97	0.948	1.84	0.941	0.39	0.946
$\min \Delta R(\mathrm{b}_1, \mathrm{b}_2) < 1.5$	19.01	0.479	4.43	0.494	0.91	0.493	0.20	0.503
$\Delta\phi(j_2, \vec{E}_\mathrm{T}^\mathrm{miss}) > 0.5$	15.62	0.821	3.74	0.844	0.77	0.846	0.17	0.838
$E_\mathrm{T}^\mathrm{miss} > 200\,\mathrm{GeV}$	7.16	0.459	2.08	0.555	0.49	0.641	0.12	0.716
total		0.023		0.042		0.055		0.066

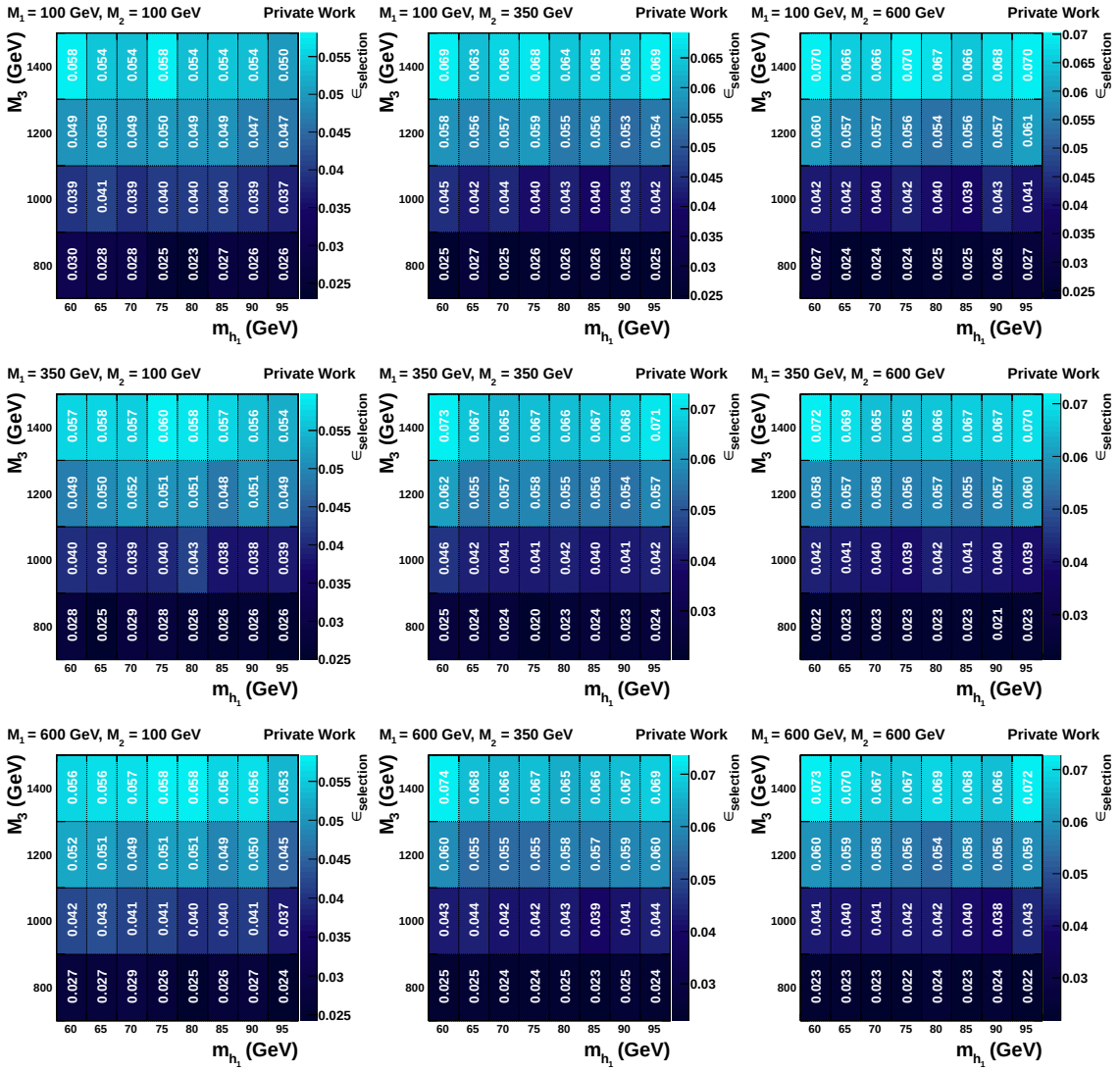

Figure 5.7: NMSSM $\rightarrow b\bar{b}$ selection-efficiency maps for the various combination of M_1 and M_2. In general the selection efficiency increases for increasing M_3.

have higher p_T values than b-quark jets of the resonant contributions in Fig. 5.1 (top), thus leading to a lower efficiency of the b veto for the two leading jets with respect to the modified P4 scenario (see Tables 5.6 and 5.7).

5.4 Background Estimation

The main backgrounds populating the signal region are from $t\bar{t}$ and QCD multijet processes. The former is estimated using simulation, and for the latter a dedicated data driven method is used as is described in Section 5.4.2. Other, minor backgrounds like $Z^0 \rightarrow \nu\bar{\nu}$ are purely estimated from simulation.

5.4.1 $t\bar{t}$ Background

The $t\bar{t}$ background is predicted using simulation and is validated in a dedicated control sample. The signal region is defined without any requirements on the lepton multiplicity in the event because a lepton veto reduces approximately one third of the expected signal contribution and requiring at least one lepton would decrease the sensitivity even more. Thus, the main background from $t\bar{t}$ processes with high $E_{\mathrm{T}}^{\mathrm{miss}}$ is dominantly from di-leptonic and semi-leptonic $t\bar{t}$, where each top quark decays into a b quark and a W boson. In case both W bosons decay into hadrons no genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ is present in those events. If at least one of the W boson decays into a lepton and a neutrino, the event may have sizeable $E_{\mathrm{T}}^{\mathrm{miss}}$ and hence can pass all signal selection requirements. The efficiencies of each selection step in simulated $t\bar{t}$ events are listed in Table 5.8. This background is estimated using simulation and normalised according to theory calculations [161]. Various control regions are devised to validate this approach. Based on the results in these regions, which are shown in the following subsections, an uncertainty of 20 % is assigned to the $t\bar{t}$ normalisation.

Table 5.8: Event yields of $t\bar{t}$ events in association with jets after each selection step. The yields are normalised to the integrated luminosity of 19.7 fb^{-1} assuming the cross sections listed in Table 5.2. The efficiency of each selection step relative to the previous step is also listed. In the first line, "all" means the events which pass the trigger selection.

	events	efficiency
all	225614.83	1.000
$H_{\mathrm{T}} > 750\,\mathrm{GeV}$	98866.23	0.438
#non-b-tagged jets ≥ 2	97581.03	0.987
#jets ≥ 4	94909.93	0.973
#b-tagged jets ≥ 2	40343.02	0.425
b veto for leading two jets	14861.42	0.368
$p_{\mathrm{T}}^{\mathrm{Lead}} > 250\,\mathrm{GeV}$, $p_{\mathrm{T}}^{\mathrm{SubLead}} > 100\,\mathrm{GeV}$	9428.45	0.634
$\Delta\phi(j_1, \vec{E}_{\mathrm{T}}^{\mathrm{miss}}) > 0.5$	8055.92	0.854
$\min \Delta R(\mathrm{b}_1, \mathrm{b}_2) < 1.5$	2386.85	0.296
$\Delta\phi(j_2, \vec{E}_{\mathrm{T}}^{\mathrm{miss}}) > 0.5$	1795.16	0.752
$E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$	82.50	0.046
total		0.000366

Leptonic Control Region with Low $E_{\mathrm{T}}^{\mathrm{miss}}$

The $t\bar{t}$ background prediction is first validated in a low $E_{\mathrm{T}}^{\mathrm{miss}}$ region. In order to avoid QCD-multijet event-contamination, one isolated lepton is required as well. This suppresses most of the QCD multijet events and thus a clean $t\bar{t}$ sample is selected. The kinematic distributions can be seen in Fig. 5.8, and the invariant mass in two variants of this control region is shown in Fig. 5.9. A good agreement is observed between

data and simulation, therefore the theoretical cross section is used to normalise the background. To cover the remaining differences a systematic uncertainty of 20 % on the normalisation of this background is applied, which is compatible with the uncertainty on the measured value [217]. The signal region requires high $E_{\mathrm{T}}^{\mathrm{miss}}$ and a good description of this variable is crucial. To investigate the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ behaviour dedicated control regions are defined, as described in the following sections.

M_{T}-Based Leptonic Control Region with High $E_{\mathrm{T}}^{\mathrm{miss}}$

Using simulation to estimate the $\mathrm{t\bar{t}}$ contribution requires a good description of $E_{\mathrm{T}}^{\mathrm{miss}}$ especially in the tail of the distribution because an $E_{\mathrm{T}}^{\mathrm{miss}}$ of 200 GeV is required by the event selection (see Section 5.3.3). Therefore, a $\mathrm{t\bar{t}}$-enriched control region with low signal contamination is devised where the following requirements differ with respect to the signal region selection: Exactly one lepton is required and the transverse mass,

$$M_{\mathrm{T}} = 2p_{\mathrm{T}}^{\mathrm{lepton}} E_{\mathrm{T}}^{\mathrm{miss}} \left\{1 - \cos\left[\Delta\phi(\mathrm{lepton}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})\right]\right\} , \tag{5.9}$$

between $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ and the four-vector of the lepton has to be less than 100 GeV. The requirement on the lepton multiplicity removes QCD-multijet contributions and M_{T} is used to remove signal contributions while keeping $\mathrm{t\bar{t}}$ contributions. In case of $\mathrm{t\bar{t}}$ events the lepton originates from a W-boson decay, and $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ is attributed to the associated neutrino, i.e. without detector effects M_{T} would not exceed m_{W}.

Further loose selection criteria are chosen, that avoid signal contamination, but ensure that the dominant $\mathrm{t\bar{t}}$ sub-contribution is well contained in this control region: $H_{\mathrm{T}} \in [750\,\mathrm{GeV}, 1500\,\mathrm{GeV}]$, exactly two b-tagged jets, less than seven jets in total, and no requirements on $p_{\mathrm{T}}^{\mathrm{Lead/SubLead}}$ and $\Delta\phi(j_{1/2}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})$. The $\min \Delta R(\mathrm{b}_1, \mathrm{b}_2)$ criterion is inverted to further reduce signal contamination, i.e. $\min \Delta R > 1.5$,

Two $E_{\mathrm{T}}^{\mathrm{miss}}$ selections are investigated in this control region: The sideband with $E_{\mathrm{T}}^{\mathrm{miss}} > 150\,\mathrm{GeV}$ (see Fig. 5.10, left) has higher statistics and shows that the shape is well described for high $E_{\mathrm{T}}^{\mathrm{miss}}$. However, the less populated $E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$ sideband (see Fig. 5.10, right) is closer to the signal region and can still be used to show that the normalisation is correct within an uncertainty of 20 % which is assigned to this background in the final fit (see also Table 5.9).

Table 5.9: Event yields in the M_{T}-based $\mathrm{t\bar{t}}$ control region with $E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$.

Data events	51
SM events	56.6
h_1 events	0.00−0.08
non-h_1 SUSY events	1.2−1.6
(h_1 + non-h_1)/SM	0.021−0.024

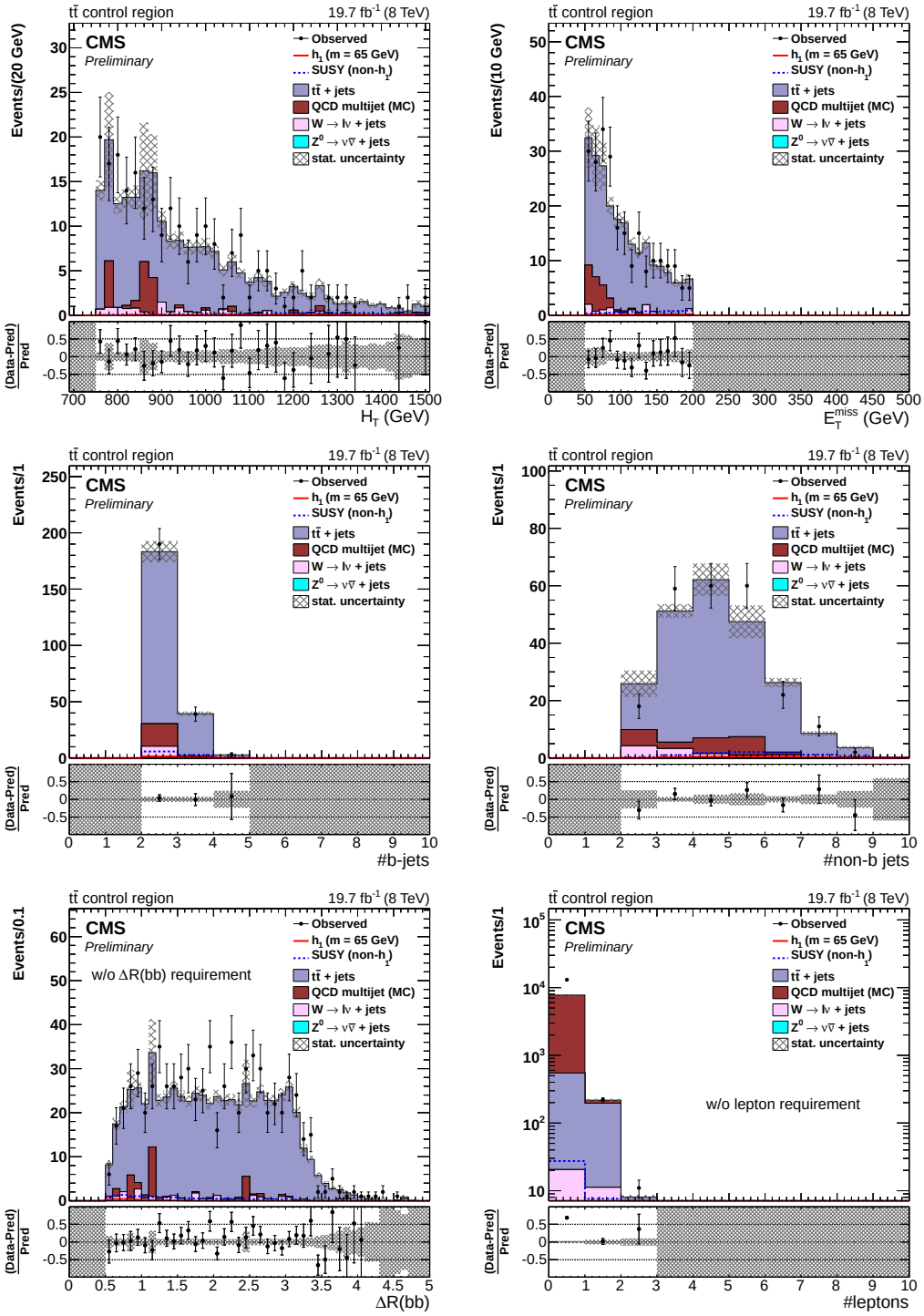

Figure 5.8: Control distributions in the $\mathrm{t\bar{t}}$ control region with $E_{\mathrm{T}}^{\mathrm{miss}} \in [50, 200]\,\mathrm{GeV}$, and #(isolated leptons) ≥ 1. The lower left plot has no $\Delta R(\mathrm{b}, \mathrm{b})$ requirement and the lower right plot has no requirement on the lepton multiplicity. Good agreement is observed. Note, that the point at #leptons = 0 in the lower right plot is not part of the $\mathrm{t\bar{t}}$ control region and the deviation observed there is attributed to the unreliable QCD-multijet simulation.

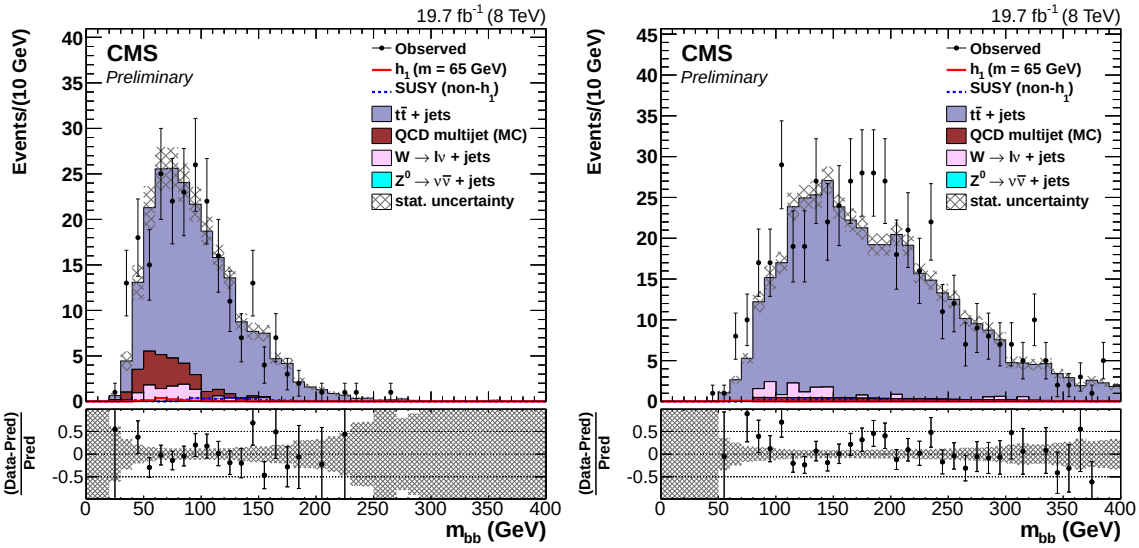

Figure 5.9: Invariant mass of the two selected b-tagged jets in the $\mathrm{t\bar{t}}$ control regions with $E_{\mathrm{T}}^{\mathrm{miss}} \in [50, 200]\,\mathrm{GeV}$, #(isolated leptons) ≥ 1, and with $\min \Delta R(\mathrm{b}, \mathrm{b}) < 1.5$ (left) and with $\min \Delta R(\mathrm{b}, \mathrm{b}) > 1.5$ (right), respectively.

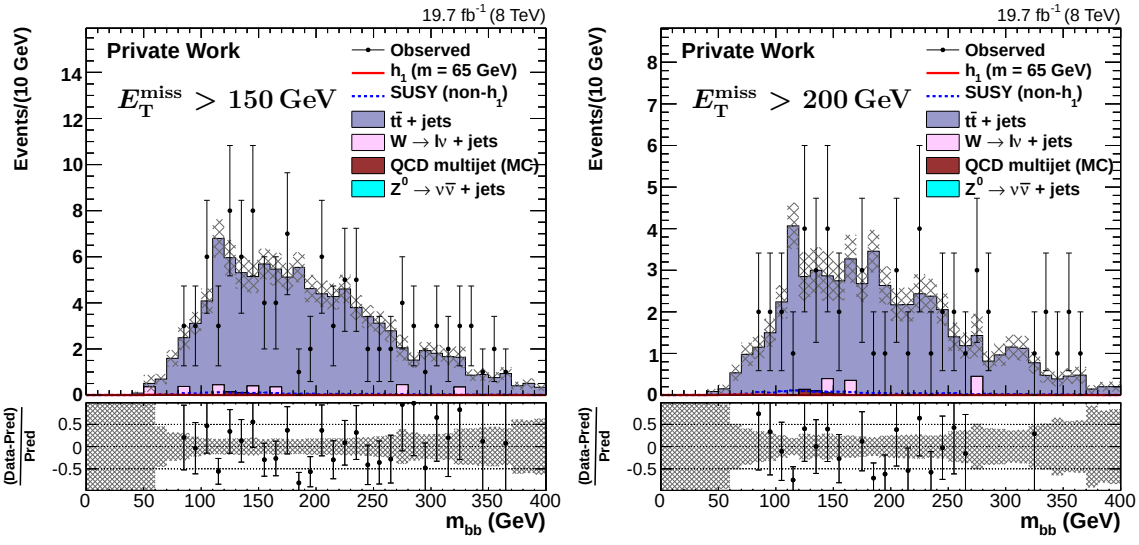

Figure 5.10: Invariant mass of the two selected b-jets in the $\mathrm{t\bar{t}}$ control regions based on M_{T} between the one selected lepton and $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ for $E_{\mathrm{T}}^{\mathrm{miss}} > 150\,\mathrm{GeV}$ (left) and $E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$ (right).

Single-b Control Region with High $E_{\mathrm{T}}^{\mathrm{miss}}$

To exclude possible bias on $E_{\mathrm{T}}^{\mathrm{miss}}$ due to the M_{T} requirement in the previous control region, an additional control region with high $E_{\mathrm{T}}^{\mathrm{miss}}$ is employed, which differs from the signal region by selecting exactly one b-tagged jet and inverting the ΔR requirement between the two selected jets. The non-b-tagged jet with the highest CSV value is selected as the second jet to compute the invariant mass. The QCD multijet contribution in this region is non-negligible and it is estimated from simulation and scaled with the data-to-simulation ratio obtained in the signal region using the data-driven method of Section 5.4.2. The uncertainty on the QCD multijet normalisation is taken into account. The results in this control region (see Fig. 5.11) support the correct modelling of the $\mathrm{t\bar{t}}$ background in the presence of high $E_{\mathrm{T}}^{\mathrm{miss}}$.

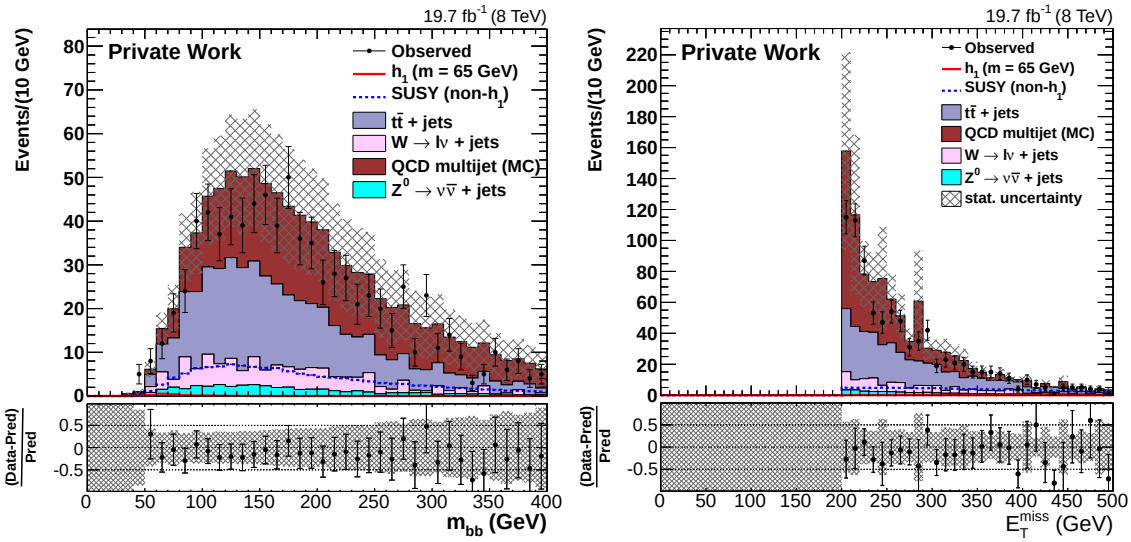

Figure 5.11: Invariant mass of the two selected b-tagged jets (left) and $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution (right) in the single-b control region.

5.4.2 QCD Multijet Background

The vast majority of the QCD multijet events are suppressed by the high $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement. The remaining events can have genuine $E_{\mathrm{T}}^{\mathrm{miss}}$, for example from semi-leptonic bottom hadron decays within a jet, where the neutrinos escape detection. But the dominant amount of the events that pass the selection are spurious $E_{\mathrm{T}}^{\mathrm{miss}}$ contributions which can arise from mismeasurements of jets. In such events the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ vector is aligned with one of the leading jets.

As illustrated in Fig. 5.12 QCD multijet events with large $E_{\mathrm{T}}^{\mathrm{miss}}$ can be highly suppressed by selecting events, where the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ vector is not close to the leading or subleading jet ($\Delta\phi(j_1/j_2, \vec{E}_{\mathrm{T}}^{\mathrm{miss}}) > 0.5$) as in Fig. 5.13 (left).

The remaining contribution is estimated using the inverted $\Delta\phi(j_2, \vec{E}_{\mathrm{T}}^{\mathrm{miss}}) < 0.5$ requirement (see Fig. 5.13, right) based on initial work by [218]. This $\Delta\phi$ sideband

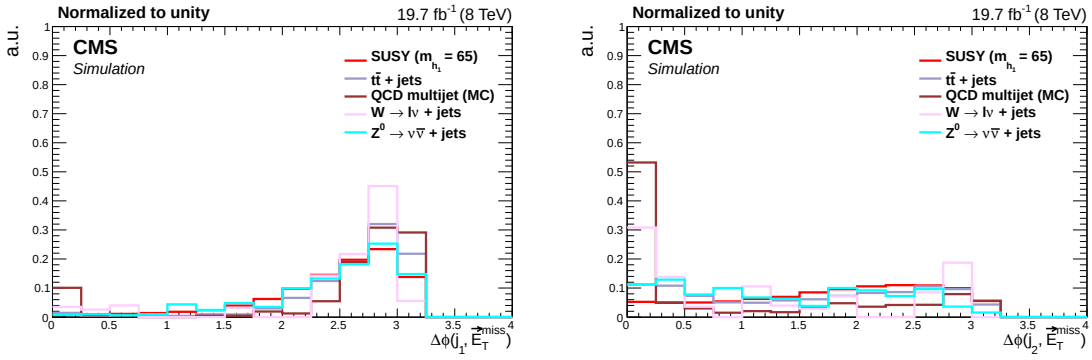

Figure 5.12: Distributions of the azimuthal angle between $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ and the leading (left) and subleading (right) jet for various MC samples after all selection steps, except for the requirements on the azimuthal angles shown in the figures. One clearly sees a concentration at low values of $\Delta\phi$ for the QCD contribution.

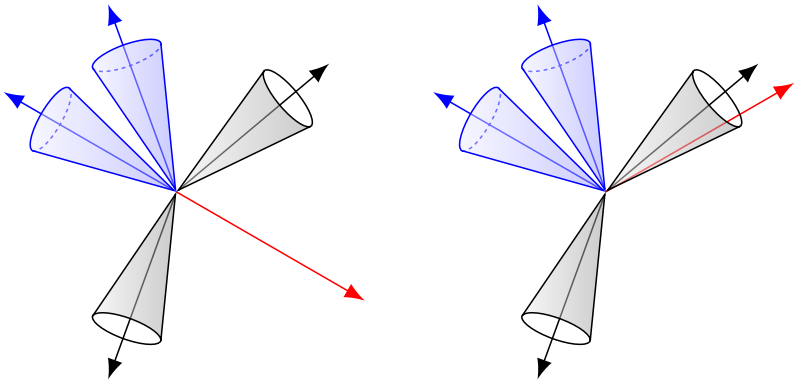

Figure 5.13: $\Delta\phi$ signal region (left) and sideband (right). The latter is QCD enriched and used to estimate the QCD contribution in the $\Delta\phi$ signal region.

is QCD dominated and contributions from $\mathrm{t\bar{t}}$ and minor electroweak processes are subtracted using simulation. To extrapolate to the $\Delta\phi$ signal region the overall normalisation and the four-momenta of the selected b-tagged jets have to be corrected.

The general procedure is closely related to an ABCD-method as illustrated in Fig. 5.14. This technique is commonly used in high energy physics to predict background contributions in the signal region by means of two uncorrelated variables. Under this assumption the signal region (D) is predicted from sideband (A) using transfer functions from sideband (A) to (B), and from sideband (A) to (C).

The two variables used for the implementation of the method in this analysis are the multiplicity of b-tagged jets n_{b} and the aforementioned azimuthal angle between $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ and the subleading jet $\Delta\phi(j_2, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})$. However, the assumption of uncorrelated variables is to some degree violated, thus a variant of the ABCD-method is

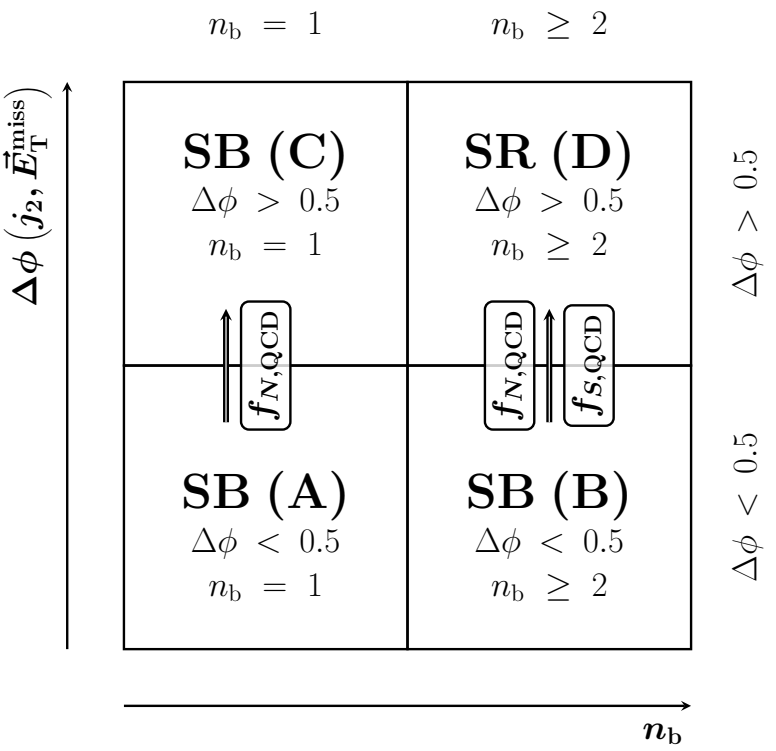

Figure 5.14: Diagram to illustrate the ABCD-like structure of the QCD-multijet-background prediction method. All selection steps (see Section 5.3.3) are applied, except for the variables on the x and y axis of the above diagram, n_b and $\Delta\phi(j_2, \vec{E}_\mathrm{T}^\mathrm{miss})$. SR (D) corresponds to the left diagram and SB (B) corresponds to the right diagram in Fig. 5.13. SB (A) and SB (C) are used to extract the normalisation factor, which is applied together with a shape correction in the transition from SB (B) to SR (D).

constructed. The general procedure is summarised with the following equation

$$\begin{aligned} m(\mathrm{b}_1,\mathrm{b}_2)^{\text{estimated QCD}}_{\Delta\phi>0.5} = (m(\mathrm{b}_1,\mathrm{b}_2)^{\text{Data}} - m(\mathrm{b}_1,\mathrm{b}_2)^{\text{MC}}_{\text{non-QCD}})_{\Delta\phi<0.5} \\ \times f_{S,\mathrm{QCD}}(E_\mathrm{T}^\mathrm{miss}, f_{\mathrm{b}_i,\mathrm{QCD}}) \times f_{N,\mathrm{QCD}}(E_\mathrm{T}^\mathrm{miss}) \; . \end{aligned} \tag{5.10}$$

where $m(\mathrm{b}_1,\mathrm{b}_2)$ is the invariant mass of the two selected b-tagged jets, $f_{N,\mathrm{QCD}}$ is a normalisation factor (see Fig. 5.14), and $f_{S,\mathrm{QCD}}$ is a shape correction factor as will be discussed in the following subsections. The selections applied in the four regions in Fig. 5.14 are as described in Section 5.3.3, except for the two variables on the x and y axis of the diagram in Fig. 5.14, n_b and $\Delta\phi(j_2, \vec{E}_\mathrm{T}^\mathrm{miss})$.

Normalisation

In order to obtain the QCD multijet prediction in the signal region with two or more b-tagged jets ($n_\mathrm{b} \geq 2$) from the $\Delta\phi$ sideband, the normalisation factor $f_{N,\mathrm{QCD}}$ needs to be applied. This factor is defined as

$$f_{N,\mathrm{QCD}}(E_{\mathrm{T}}^{\mathrm{miss}}, n_{\mathrm{b}}) = \frac{\#\text{Events with } E_{\mathrm{T}}^{\mathrm{miss}},\ n_{\mathrm{b}} \text{ b-tagged jets, } \Delta\phi > 0.5}{\#\text{Events with } E_{\mathrm{T}}^{\mathrm{miss}},\ n_{\mathrm{b}} \text{ b-tagged jets, } \Delta\phi < 0.5}\ , \tag{5.11}$$

and is $E_{\mathrm{T}}^{\mathrm{miss}}$-dependent as can be seen in Fig. 5.15. This ratio is approximately independent of the number of b-tagged jets n_{b},

$$f_{N,\mathrm{QCD}}(E_{\mathrm{T}}^{\mathrm{miss}}, n_{\mathrm{b}} = 1) \approx f_{N,\mathrm{QCD}}(E_{\mathrm{T}}^{\mathrm{miss}}, n_{\mathrm{b}} \geq 2)\ , \tag{5.12}$$

therefore the factor derived for $n_{\mathrm{b}} = 1$ (single-b) is also used for $n_{\mathrm{b}} \geq 2$ (double-b). The single-b data-set is selected by requiring at least four jets with the same p_{T} thresholds as in the signal region, but with exactly one b-tagged jet. The jet with the third-highest p_{T} is chosen for the construction of the invariant mass because the softer selected jet in simulated QCD multijet events has in the majority of the events the highest possible rank in p_{T} (see Fig. 5.3, top right). The selected non-b-tagged jet has to fulfil the same requirements as the second selected jet in the signal region.

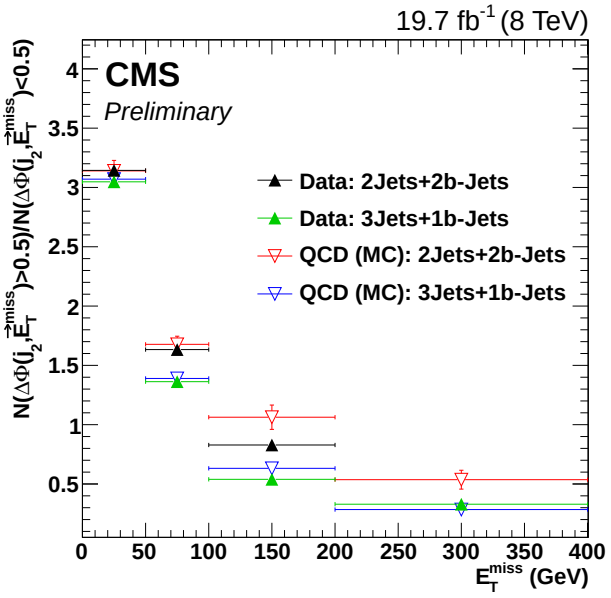

Figure 5.15: QCD-normalisation factors measured in an orthogonal data set requiring exactly one b-tagged jet (green filled upward triangle: data, blue hollow downward triangle: QCD simulation) as a function of $E_{\mathrm{T}}^{\mathrm{miss}}$. For comparison the measured correction factors of the data sets with at least two b-tagged jets (black filled upward triangle: data, red hollow downward triangle: QCD simulation) are shown in the low $E_{\mathrm{T}}^{\mathrm{miss}}$ region for comparison.

For a better statistical precision, a full ABCD-method starting from sideband (A) (see Figs. 5.14, E.1) and applying the ratio with $(n_{\mathrm{b}} \geq 2)/(n_{\mathrm{b}} = 1)$ and $\Delta\phi < 0.5$ would be desirable. However, the shape of the m_{bb} distribution with $n_{\mathrm{b}} \geq 2$ differs from the distribution with $n_{\mathrm{b}} = 1$ because in the latter case one of the two selected

jets is the third non-b-tagged jet. This jet is typically more energetic then the corresponding jet in the double-b sample (see Fig. 5.6, right) because in case of $n_{\mathrm{b}} \geq 2$, the selected jets can be of any rank (see Fig. 5.3, top) except the highest and second-highest rank (b veto, see Section 5.3.3), but for the single-b selection the rank of the selected jets is biased towards higher ranks.

In order to avoid this shape variation, sideband (B) is chosen as starting point and the normalisation factor obtained in the single-b sample is applied. However, Fig. 5.15 shows a systematic deviation between the single-b and double-b distributions both in data and QCD simulation. Therefore, the ratio between the double-b and single-b distributions is computed and is used as a correction to the single-b normalisation factor (see Fig. 5.16).

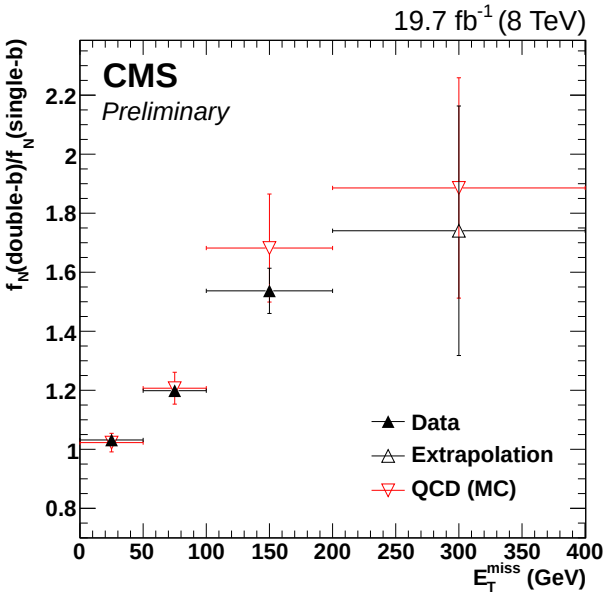

Figure 5.16: QCD-normalisation correction measured in data (black filled upward triangle) and QCD simulation (red hollow downward triangle). The extrapolated value for $E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$ in data is plotted with a different marker (black hollow upward triangle).

Since the systematic difference between double-b and single-b is confirmed in QCD multijet simulation also for $E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$, the normalisation correction factor used for data is obtained as an extrapolation of the difference in the highest $E_{\mathrm{T}}^{\mathrm{miss}}$ sideband to the $E_{\mathrm{T}}^{\mathrm{miss}}$ signal region. The uncertainty of the overall normalisation factor is estimated using the double-b–single-b difference in QCD simulation for $E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$. Therefore, the final normalisation factor is

$$f_{N,\mathrm{QCD}} = 0.57 \pm 0.25\,, \tag{5.13}$$

where the uncertainty of 43.9 % on this factor is used as systematic uncertainty on the QCD-multijet normalisation.

Shape Correction

As pointed out in the last subsection the shape of the m_{bb} distribution in the sidebands (A) and (B) (see Figs. 5.14) differs from each other, but also the transition from sideband (B) to signal region (D) causes a shape variation. This effect is corrected for using $f_{S,\mathrm{QCD}}$, which is a function of additional jet energy scale factors $f_{\mathrm{b}_i,\mathrm{QCD}}$ derived for each of the selected jets.

This approach is motivated by the effect of an inverted $\Delta\phi(j_2, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})$ requirement on the kinematics of the two b-tagged jets which can be seen in Fig. 5.17. Requiring the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ vector to be close to that of the subleading jet results in softer p_{T} spectra of the b-tagged jets because this enriches dijet-like events in the selected sample with sizeable jet-energy mismeasurement leading to less energetic jets for the lower p_{T} ranks, i.e. the selected b-tagged jets (b veto for the leading two jets).

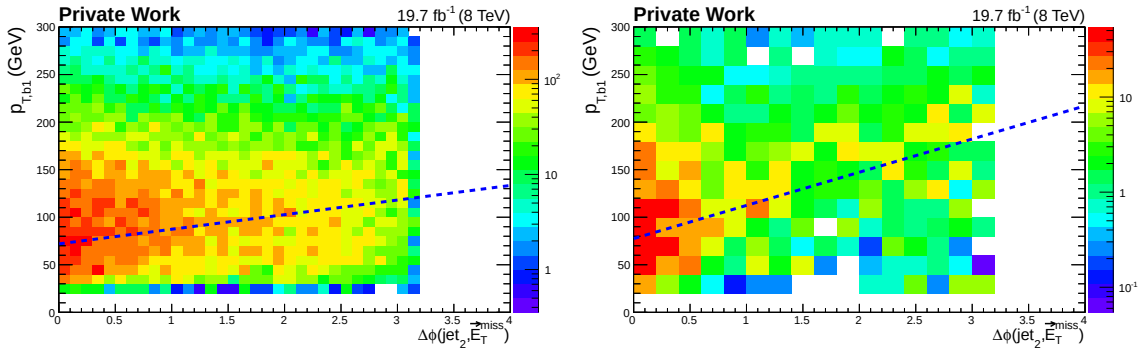

Figure 5.17: Correlation between the b-tagged jet-p_{T} and $\Delta\phi(j_2, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})$. The p_{T}-spectrum is harder in the $\Delta\phi(j_2, \vec{E}_{\mathrm{T}}^{\mathrm{miss}}) > 0.5$ signal region with respect to the $\Delta\phi(j_2, \vec{E}_{\mathrm{T}}^{\mathrm{miss}}) < 0.5$ sideband. The effect is illustrated by the blue dashed line crossing the maximum in each $\Delta\phi(j_2, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})$-slice and is more prominent if requiring $E_{\mathrm{T}}^{\mathrm{miss}} > 100\,\mathrm{GeV}$ (right) than without $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement (left).

The above-mentioned effect on the b-jet p_{T} is addressed with additional jet-energy-scale factors applied to the four-vectors of each of the selected b-tagged jets. These are defined as the ratio of the mean of the p_{T} distribution in the $\Delta\phi$ signal region and the mean of its p_{T} distribution in the $\Delta\phi$ sideband,

$$f_{\mathrm{b}_i,\mathrm{QCD}}(E_{\mathrm{T}}^{\mathrm{miss}}) = \frac{\overline{p_{T,\mathrm{b}_i}}(\#\text{Events with } E_{\mathrm{T}}^{\mathrm{miss}},\, n_{\mathrm{b}} \geq 2,\, \Delta\phi > 0.5)}{\overline{p_{T,\mathrm{b}_i}}(\#\text{Events with } E_{\mathrm{T}}^{\mathrm{miss}},\, n_{\mathrm{b}} \geq 2,\, \Delta\phi < 0.5)}\,. \tag{5.14}$$

The effect on the b-jet p_{T} increases when requiring higher $E_{\mathrm{T}}^{\mathrm{miss}}$ in the event (see Fig. 5.17, right) because this enriches the selected sample with events exhibiting genuine $E_{\mathrm{T}}^{\mathrm{miss}}$, i.e. less dijet-like events. Therefore, the shape correction factors are $E_{\mathrm{T}}^{\mathrm{miss}}$ dependent (see Fig. 5.18).

These factors are measured in the same $E_{\mathrm{T}}^{\mathrm{miss}}$ bins as the normalisation factor. The correction factors for $E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$ are extrapolated from the highest $E_{\mathrm{T}}^{\mathrm{miss}}$

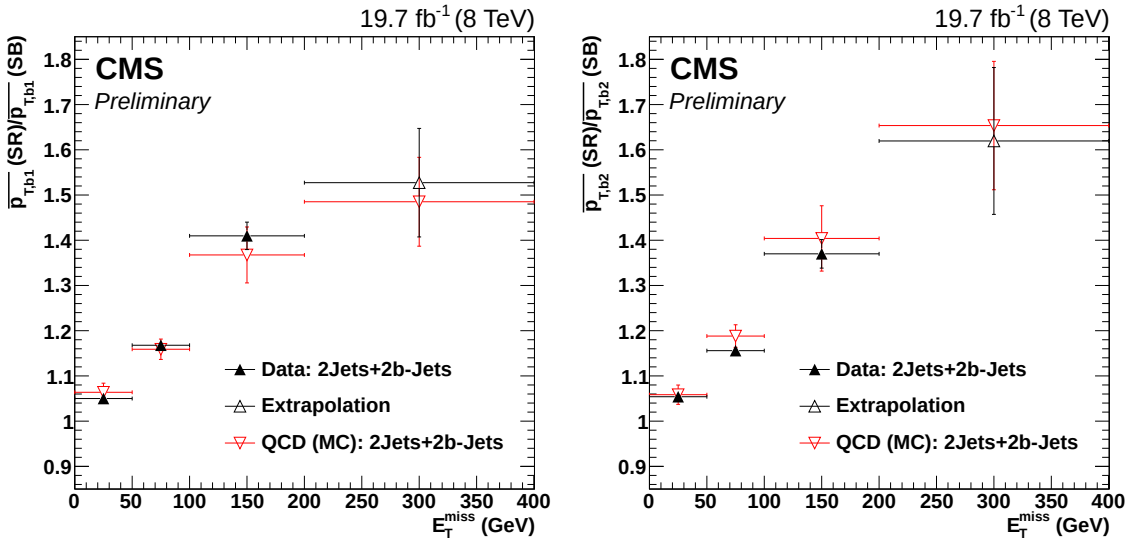

Figure 5.18: QCD-shape-correction factors derived from the ratio of the jet-p_{T} mean in the $\Delta\phi$ signal region and in the $\Delta\phi$ sideband. The correction factor in the $E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$ signal region is derived by extrapolating the data/MC difference from the highest $E_{\mathrm{T}}^{\mathrm{miss}}$ sideband ($100\,\mathrm{GeV} < E_{\mathrm{T}}^{\mathrm{miss}} < 200\,\mathrm{GeV}$).

sideband,

$$f^{\mathrm{Data}}_{\mathrm{b}_i,\mathrm{QCD}}(E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}) = \Big(f^{\mathrm{Data}}_{\mathrm{b}_i,\mathrm{QCD}}(100\,\mathrm{GeV} > E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}) - f^{\mathrm{QCD,MC}}_{\mathrm{b}_i,\mathrm{QCD}}(100\,\mathrm{GeV} > E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV})\Big) + f^{\mathrm{QCD,MC}}_{\mathrm{b}_i,\mathrm{QCD}}(E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV})\,, \tag{5.15}$$

$$f^{\mathrm{Data}}_{\mathrm{b}_1,\mathrm{QCD}}(E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}) = 1.53 \pm 0.18\,, \tag{5.16}$$

$$f^{\mathrm{Data}}_{\mathrm{b}_2,\mathrm{QCD}}(E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}) = 1.62 \pm 0.26\,, \tag{5.17}$$

where the uncertainties are derived by propagating the statistical uncertainties of the quantities in (5.15) and account to $12 - 16\,\%$. The impact of this uncertainty on the shape is shown in Fig. 5.19.

Validation

The method is validated in events with $E_{\mathrm{T}}^{\mathrm{miss}} < 200\,\mathrm{GeV}$; namely the measured normalisation factor $f_{N,\mathrm{QCD}}$, and the shape correction factors $f_{\mathrm{b}_i,\mathrm{QCD}}$ are applied in three bins of the low-$E_{\mathrm{T}}^{\mathrm{miss}}$ sideband ($E_{\mathrm{T}}^{\mathrm{miss}} = [0-50\,\mathrm{GeV}, 50-100\,\mathrm{GeV}, 100-200\,\mathrm{GeV}]$, see Fig. 5.20). In contrast to the region, where the factors are measured, the validation regions have a veto on isolated leptons to create a QCD-enriched region, where shape differences due to a mismodelling of the QCD-multijet background would have a strong impact. Figure 5.20 (right) shows that the m_{bb} shape is well predicted by the method described in Section 5.4.2, which corrects the four-vectors of the selected jets.

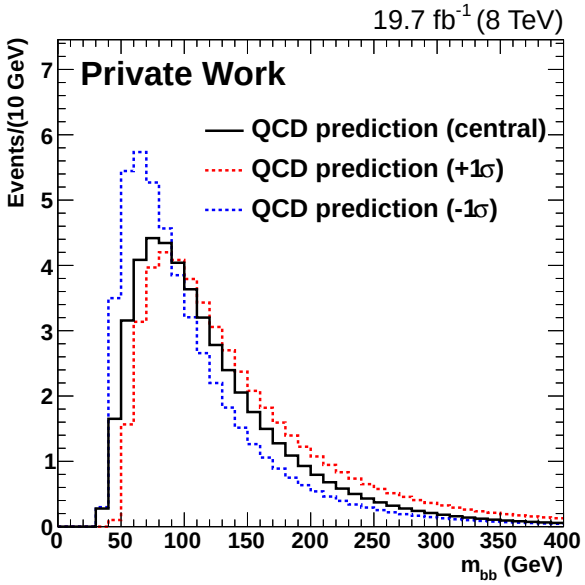

Figure 5.19: QCD-multijet prediction and its shapes obtained by varying the shape correction factors $f_{\mathrm{b}_i,\mathrm{QCD}}$ within their uncertainties.

It works for all $E_{\mathrm{T}}^{\mathrm{miss}}$ bins of the validation region. A comparison of the right-hand side with the left-hand side of Fig. 5.20, where no shape correction is applied, shows clearly that the agreement of data and predicted distributions is improved by the shape correction. Hence, this method can be safely used also for the $E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$ bin.

Parameterisation

Due to the low statistics in the $\Delta\phi$ sideband (sideband (B) in Fig. 5.14) the QCD-background prediction for the signal region as described above suffers from bin-by-bin fluctuations. In order to prevent such statistical fluctuations from mimicking a peak, a parameterisation of the QCD background is performed.

The QCD multijet shape can be expected to have a maximum on the low mass side, a tail on the high mass side, but no resonant structures. Various parameterisations are considered; they must meet these basic requirements, and must be able to fit the QCD multijet spectrum in the signal region and the QCD control regions. The studied parameterisations are the Crystal Ball function [219–221], Bernstein polynomials, and the Novosibirsk function [222] (see also Appendix E.2.1). The result of the fits of these functions to the QCD prediction is shown in Fig. 5.21.

Although the Crystal Ball, the Bernstein polynomials, and the Novosibirsk function perform equally well for the signal region prediction, the latter function is chosen because it is also found to be applicable for parameterising the QCD-multijet-m_{bb} distribution in the validation regions (see Section 5.4.2) with higher statistics (see Appendix E.2.2).

In case of the Bernstein polynomials, it would be possible to raise the order of the polynomials in order to fit to the QCD control region predictions. However, this does

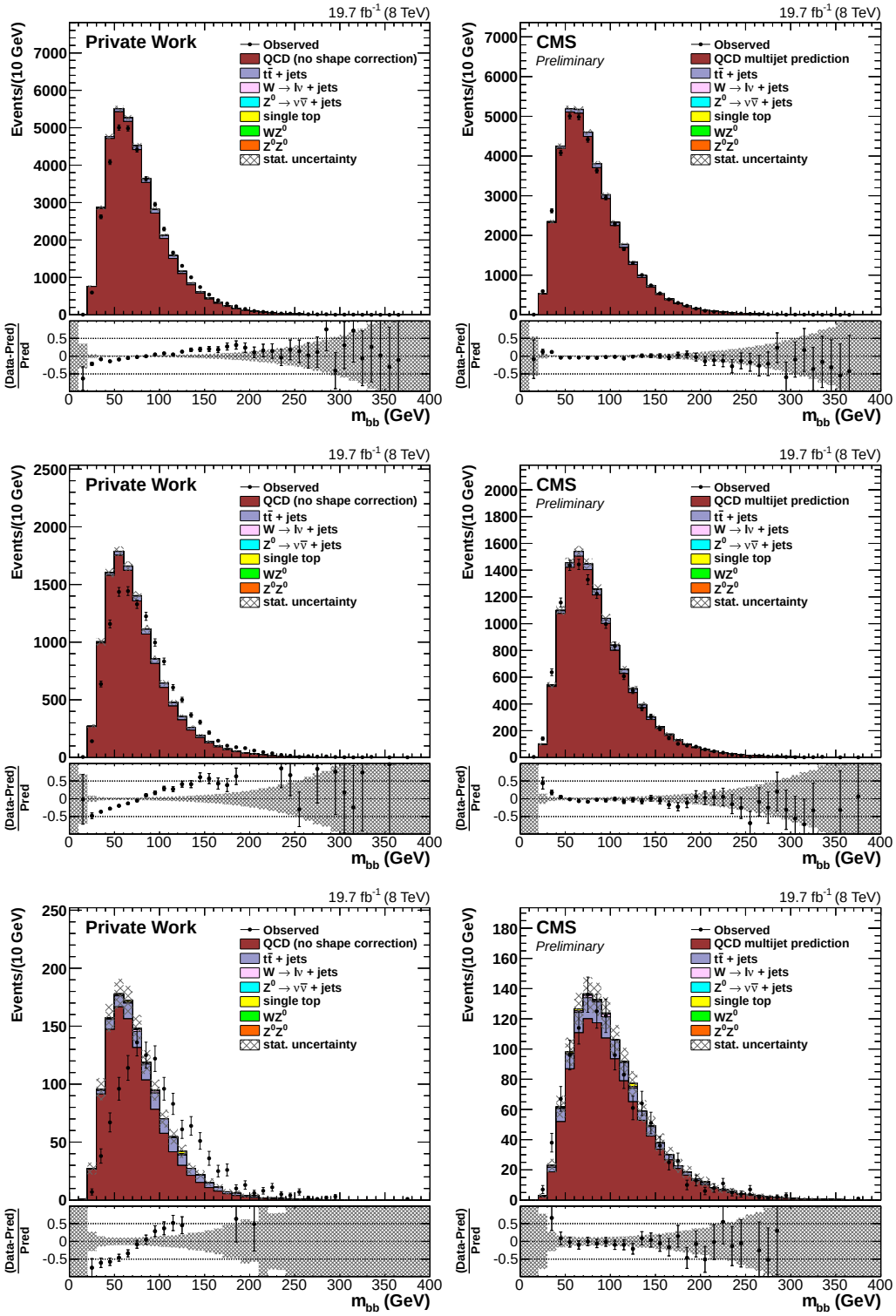

Figure 5.20: Validation of the QCD background estimation method without (left) and with (right) shape correction in control regions with a veto on isolated leptons and in bins of $E_{\mathrm{T}}^{\mathrm{miss}}$: $E_{\mathrm{T}}^{\mathrm{miss}} \in [0, 50]\,\mathrm{GeV}$ (top), $E_{\mathrm{T}}^{\mathrm{miss}} \in [50, 100]\,\mathrm{GeV}$ (centre), $E_{\mathrm{T}}^{\mathrm{miss}} \in [100, 200]\,\mathrm{GeV}$ (bottom).

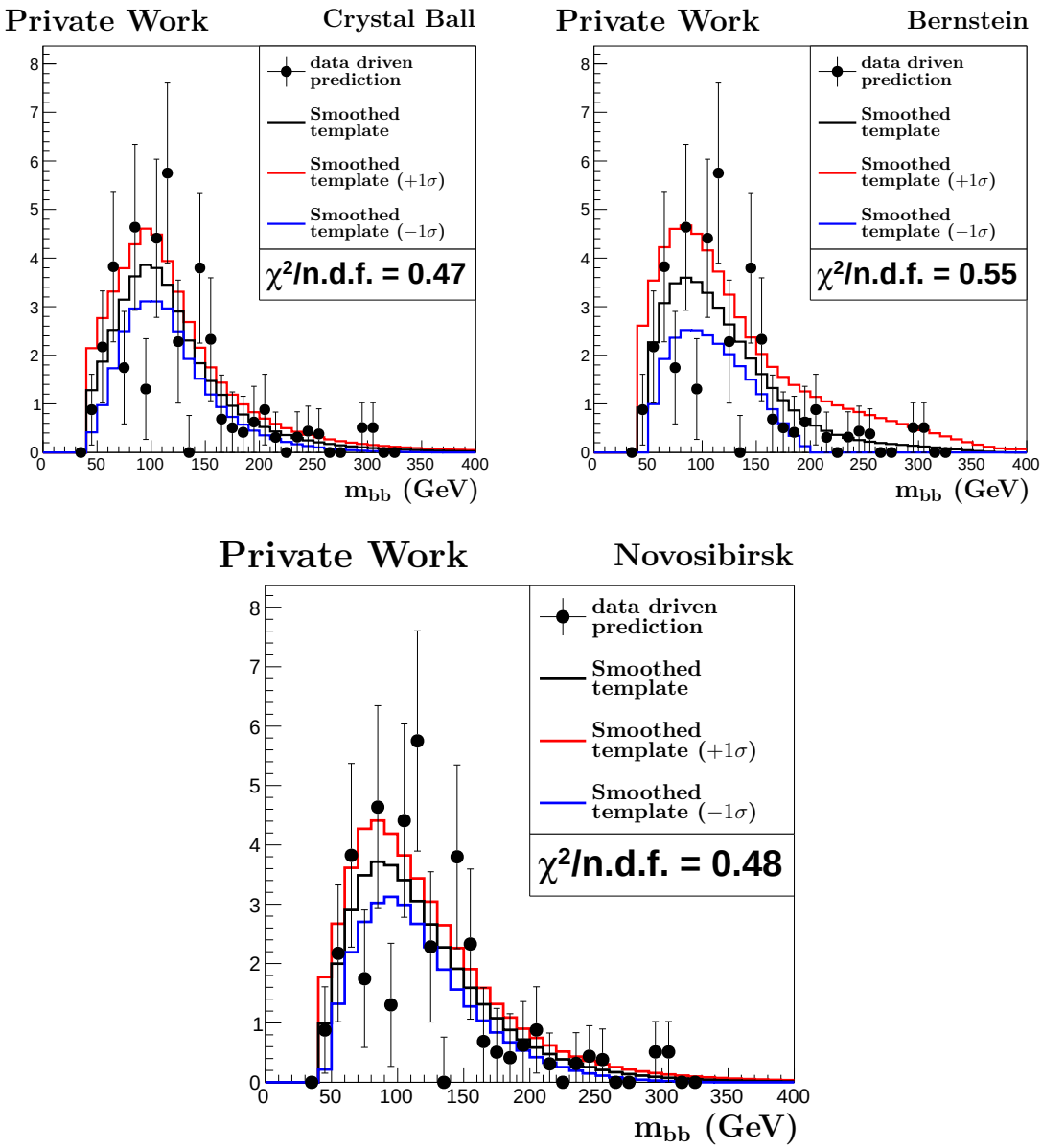

Figure 5.21: Fit results of the QCD signal region prediction: Crystal Ball (top left), Bernstein (8th order, top right) and Novosibirsk (bottom). The best fit quality is achieved for the Crystal Ball and the Novosibirsk function. However, only the Novosibirsk function is able to describe the QCD multijet prediction in the low-$E_{\mathrm{T}}^{\mathrm{miss}}$ sidebands with much higher event rates (see Appendix E.2.2).

not work anymore in the signal region because the polynomial would follow every fluctuation of the QCD prediction, contradicting the idea of the parameterisation which aims for a smoother shape.

The uncertainty of the fit to the QCD prediction for the signal region is treated as additional shape uncertainty using the red and blue histograms in Fig. 5.21 (bottom).

5.4.3 Other Backgrounds

Further background processes can contribute to the signal region due to high $E_{\mathrm{T}}^{\mathrm{miss}}$ in their event signature. These are $\mathrm{Z}^0 \to \nu\bar{\nu}$ and $\mathrm{W} \to \ell\nu$ events in association with jets (see Tables 5.10, 5.11). Since the expected event yields for these processes are well below the $\mathrm{t\bar{t}}$ and QCD-multijet predictions, their impact on the final result is minor. Therefore, simulation is used to estimate these backgrounds and the normalisation is taken from theory calculations [161].

Table 5.10: Event yields of $\mathrm{Z}^0 \to \nu\bar{\nu}$ events in association with jets after each selection step. The yields are normalised to the integrated luminosity of $19.7\,\mathrm{fb}^{-1}$ assuming the cross sections listed in Table 5.2. The efficiency of each selection step relative to the previous step is also listed. In the first line, "all" means the events which pass the trigger selection.

	events	efficiency
all	14797.22	1.000
$H_{\mathrm{T}} > 750\,\mathrm{GeV}$	8515.27	0.575
#non-b-tagged jets ≥ 2	8350.30	0.981
#jets ≥ 4	4344.13	0.520
#b-tagged jets ≥ 2	99.41	0.023
b veto for leading two jets	30.21	0.304
$p_{\mathrm{T}}^{\mathrm{Lead}} > 250\,\mathrm{GeV}$, $p_{\mathrm{T}}^{\mathrm{SubLead}} > 100\,\mathrm{GeV}$	25.90	0.857
$\Delta\phi(j_1, \vec{E}_{\mathrm{T}}^{\mathrm{miss}}) > 0.5$	23.57	0.910
$\min \Delta R(\mathrm{b}_1, \mathrm{b}_2) < 1.5$	11.41	0.484
$\Delta\phi(j_2, \vec{E}_{\mathrm{T}}^{\mathrm{miss}}) > 0.5$	8.41	0.737
$E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$	2.31	0.275
total		0.000156

Single Top, WZ^0 and $\mathrm{Z}^0\mathrm{Z}^0$ Backgrounds

The contributions from single top, WZ^0 and $\mathrm{Z}^0\mathrm{Z}^0$ are investigated. The former may have true $E_{\mathrm{T}}^{\mathrm{miss}}$, one true b quark, and possibly a mis-tagged one. The latter two may have true $E_{\mathrm{T}}^{\mathrm{miss}}$ and two true b quarks from the Z^0 decay. However, their contributions are found to be negligible as is summarised in Appendix F (Tables F.1, F.2, and F.3). Hence, they are not considered in the final background estimate.

Table 5.11: Event yields of $W \to \ell\nu$ events in association with jets after each selection step. The yields are normalised to the integrated luminosity of $19.7\,\mathrm{fb}^{-1}$ assuming the cross sections listed in Table 5.2. The efficiency of each selection step relative to the previous step is also listed. In the first line, "all" means the events which pass the trigger selection.

	events	efficiency
all	120404.41	1.000
$H_\mathrm{T} > 750\,\mathrm{GeV}$	44002.01	0.365
#non-b-tagged jets ≥ 2	43318.77	0.984
#jets ≥ 4	24327.63	0.562
#b-tagged jets ≥ 2	459.34	0.019
b veto for leading two jets	167.35	0.364
$p_\mathrm{T}^\mathrm{Lead} > 250\,\mathrm{GeV}$, $p_\mathrm{T}^\mathrm{SubLead} > 100\,\mathrm{GeV}$	138.04	0.825
$\Delta\phi(j_1, \vec{E}_\mathrm{T}^\mathrm{miss}) > 0.5$	119.95	0.869
$\min \Delta R(\mathrm{b}_1, \mathrm{b}_2) < 1.5$	60.20	0.502
$\Delta\phi(j_2, \vec{E}_\mathrm{T}^\mathrm{miss}) > 0.5$	43.81	0.728
$E_\mathrm{T}^\mathrm{miss} > 200\,\mathrm{GeV}$	4.81	0.110
total		0.000040

$Z^0 \to b\bar{b}$ Background

Contributions from this background would arise from fake-$E_\mathrm{T}^\mathrm{miss}$ events. No simulated $Z^0 \to b\bar{b}$ sample has been available, therefore simulated Drell–Yan samples (see Table 5.2) are used, where the Z^0 decays only into leptons. The four-vectors of these leptons are interpreted as b-tagged jets that always pass the b-tagging requirements. The cross section of the sample is scaled by the ratio of branching fractions for the Z^0 to decay to b quarks and the Z^0 to decay to leptons.

Less than four events pass the event selection in this test, thus yielding a very small conservative upper bound (see Appendix F, Table F.4). In addition, this contribution would be anyway included by the method used to derive the QCD-background prediction, which targets false-$E_\mathrm{T}^\mathrm{miss}$ contributions.

5.5 Signal Modelling

The simulated NMSSM data sets listed in Table 5.1 contain basically two types of events. One type is defined as events where at least one h_1 is produced and decays into a pair of b-quarks, and the remainder, which is referred to as non-h_1 SUSY contributions. The latter consists of four components: $Z^0 \to b\bar{b}$, as depicted in the upper right of Fig. 1.7 and in Fig. 5.1 (top right), $h_2 \to b\bar{b}$ (see Fig. 5.1, top left), $a_1 \to b\bar{b}$ (see Fig. 5.1, top left), and non-resonant contributions (see Fig. 5.1, bottom).

A summary of the h_1 contribution and non-h_1 SUSY contributions in the signal region is shown in Fig. 5.22. The signal templates are extracted from the samples

listed in Table 5.1 for h_1 masses from 30 GeV to 100 GeV in the modified P4 scenario, and from 60 GeV to 95 GeV in the decoupled squarks scenario. In case of the decoupled squarks scenario a template is extracted for each parameter point from the sample listed at the bottom of Table 5.1.

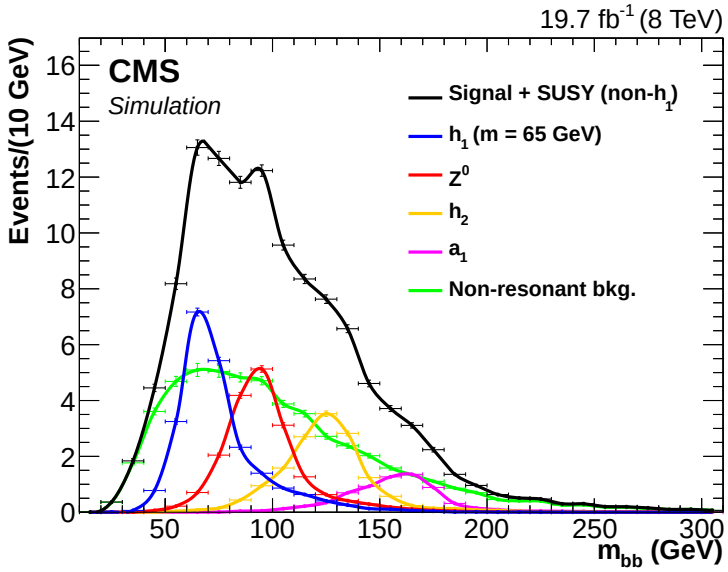

Figure 5.22: The h_1 contributions, labelled as "Signal", and the non-h_1 SUSY prediction of the modified NMSSM P4 scenario with $m_{h_1} = 65\,\text{GeV}$ and normalised to the integrated luminosity of $19.7\,\text{fb}^{-1}$ collected in 2012 by CMS at a centre-of-mass energy of 8 TeV. The subcontributions use the same colour coding as the ellipses in Fig. 5.1.

5.6 Statistical Inference on Light Higgs Boson Hypotheses

In this section the statistical methods are summarised, which are used to extract information about certain signal hypotheses from the data, i.e. the determination of signal process cross sections and upper limits. The procedure is implemented in a toolkit provided by the CMS collaboration and follows the recommendations of the LHC Higgs Combination Group [223]. The modified frequentist criterion CL_s [224, 225] is adopted using the RooStats software package [226]. The methods are based on likelihood functions constructed from the binned m_{bb} distributions taking into account systematic uncertainties via nuisance parameters. For each bin i of the model to be tested, i.e. *background-only* or *signal+background*, the predicted event yield ν_i is parameterised as

$$\nu_i = \mu s_i(\theta) + b_i(\theta)\ . \tag{5.18}$$

Here, s_i and b_i represent for each bin i the signal prediction and the background prediction, respectively. The parameter μ is called *signal strength modifier* and is the *parameter of interest* in this model. In addition to this parameter, signal and background predictions depend on θ, i.e. the vector of *nuisance parameters* which are used to account for systematic uncertainties in the final fit. These are themselves constrained by prior measurements or estimation procedures (see Section 5.7). This knowledge is cast into $p(\tilde{\theta}|\theta)$, the probability to yield $\tilde{\theta}$ given the true value θ. This function is of the form of a log-normal function for normalisation uncertainties

$$p(x|\hat{x},\kappa) = \frac{1}{\sqrt{2\pi}\ln(\kappa)} \exp\left(\frac{\ln(x/\hat{x})^2}{2\ln(\kappa)^2}\right)\frac{1}{x}\ , \tag{5.19}$$

with $\kappa = \mathrm{e}^{\sigma}$, where σ corresponds to one standard deviation. For uncertainties that alter the shape of the m_{bb} distribution in a correlated way, each bin gets assigned a normal distribution with mean 0 and unit width. The likelihood $\mathcal{L}$ is then constructed as the product of Poisson probabilities for each bin and the constraints p for the nuisance parameters,

$$\mathcal{L}(\mathrm{data}|\mu,\theta) = \prod_i \frac{[\mu s_i(\theta) + b_i(\theta)]^{n_i}}{n_i!}\mathrm{e}^{\mu s_i(\theta)+b_i(\theta)} \times p(\tilde{\theta}|\theta)\ . \tag{5.20}$$

Here, "data" represents either the observed event yields in each bin or pseudo-data, i.e. generated data points. The likelihood serves the measurement of the signal cross section by means of a *maximum likelihood* fit. In searches for new physics, this function is used to either quantify an excess or to derive upper limits on μ as will be briefly outlined in the following.

5.6.1 Profile Likelihood Ratio

Profile likelihood ratios are used as test statistic to quantify an excess with respect to the background expectation or to compute exclusion limits of the signal. For the limit calculation this quantity is defined as

$$\tilde{q}_\mu = -2\ln\frac{\mathcal{L}(\mathrm{data}|\mu,\hat{\theta}_\mu)}{\mathcal{L}(\mathrm{data}|\hat{\mu},\hat{\theta})}\ , \qquad \text{with a constraint } 0 \leq \hat{\mu} \leq \mu\ , \tag{5.21}$$

where $\hat{\theta}_\mu$ are the nuisance parameter values that maximise $\mathcal{L}$ given the data and a fixed signal strength μ, and the estimators $\hat{\mu}$ and $\hat{\theta}$ globally maximise the likelihood. The lower bound on $\hat{\mu}$ protects against downward fluctuations of the background, whereas the upper bound imposes one-sided limits, i.e. upward fluctuations are not considered as evidence against the signal hypothesis with strength μ.

5.6.2 Expected and Observed Upper Limits

In order to extract upper limits on the signal strength, which can be translated into exclusion limits within a model parameter space, the probability density functions (pdf) $f(\tilde{q}_\mu|\mu,\hat{\theta}_\mu^{\text{obs}})$ and $f(\tilde{q}_\mu|0,\hat{\theta}_0^{\text{obs}})$ have to be computed, i.e. assuming a signal with strength μ in the signal+background hypothesis and $\mu = 0$ in the background-only hypothesis, respectively. These distributions are obtained from pseudo-datasets, that are generated for each value of μ. From the pdfs one can construct two p-values,

$$p_s(\mu) = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{\text{obs}}|\text{signal+background}) = \int_{\tilde{q}_\mu^{\text{obs}}}^{\infty} f(\tilde{q}_\mu|\mu,\hat{\theta}_\mu^{\text{obs}})\mathrm{d}\tilde{q}_\mu \,, \tag{5.22a}$$

$$1 - p_b(\mu) = P(\tilde{q}_\mu \geq \tilde{q}_0^{\text{obs}}|\text{background-only}) \quad = \int_{\tilde{q}_0^{\text{obs}}}^{\infty} f(\tilde{q}_\mu|0,\hat{\theta}_0^{\text{obs}})\mathrm{d}\tilde{q}_\mu \,, \tag{5.22b}$$

which define the probabilities to obtain a value of $\tilde{q}_\mu$ greater or equal to the observed value $\tilde{q}_\mu^{\text{obs}}$. The ratio of (5.22a) and (5.22b),

$$\text{CL}_\text{s}(\mu) = \frac{p_s(\mu)}{1 - p_b(\mu)} \,, \tag{5.23}$$

is used to exclude a given signal strength μ at $1 - \text{CL}_\text{s}(\mu)$ confidence level. The confidence level in this analysis is chosen to be 95 %, and the signal strength $\mu^{95\,\%}$ is defined such that $1 - \text{CL}_\text{s}(\mu^{95\,\%}) = 0.95$.

The whole procedure is executed to calculate an upper limit using the observed data. But one can also generate a set of pseudo-data to obtain limits based on the expected background taking into account statistical fluctuations, i.e. calculating $\pm1\sigma$ and $\pm2\sigma$ uncertainty bands. However, this procedure is very computing intensive, thus an *asymptotic approximation* [227] of the test statistic pdfs in the large data sample limit is used. The pdfs then read

$$\begin{aligned} f(\tilde{q}_\mu|\mu') =& \frac{1}{\sqrt{8\pi}}\frac{1}{\sqrt{\tilde{q}_\mu}}\exp\left[-\frac{1}{2}\left(\sqrt{\tilde{q}_\mu} + \frac{\mu-\mu'}{\sigma}\right)^2\right] \\ &+ \begin{cases} \frac{1}{\sqrt{8\pi}}\frac{1}{\sqrt{\tilde{q}_\mu}}\exp\left[-\frac{1}{2}\left(\sqrt{\tilde{q}_\mu} - \frac{\mu-\mu'}{\sigma}\right)^2\right] \,, & \tilde{q}_\mu \geq \mu^2/\sigma^2 \\ \frac{1}{\sqrt{8\pi}}\frac{1}{2\mu/\sigma}\exp\left[-\frac{1}{2}\frac{\tilde{q}_\mu - \frac{\mu^2-2\mu\mu'}{\sigma^2}}{\mu^2/\sigma^2}\right] \,, & \tilde{q}_\mu \geq \mu^2/\sigma^2 \,, \end{cases} \end{aligned} \tag{5.24}$$

where $\sigma = \mu/\tilde{q}_{\mu,\text{A}}$. The index A indicates that the test statistic is evaluated on an *Asimov data set*, which is defined as the expected signal and background contributions for nominal nuisance parameter values [227]. The median and the uncertainty bands of the test statistic pdf can then be directly computed from its cumulative function without generating pseudo-data.

5.7 Systematic Uncertainties

One key aspect in the extraction of cross sections or exclusion limits is the proper estimation of systematic uncertainties attributed to the signal and background predictions. Systematic uncertainties on the QCD-multijet background method as described in detail in section 5.4.2 include the uncertainties on the normalisation, on the shape correction, and on the shape parameterisation. The remaining systematics on the background normalisation are taken from measurements [217, 228, 229] and account to 20 % for each contribution considered in the final fit ($t\bar{t}$, $W \to \ell\nu$, $Z^0 \to \nu\bar{\nu}$).

Further systematic uncertainties that are included will be described in the following. An overview of the considered systematic uncertainties can be found in Table 5.12.

Table 5.12: Considered sources and types of systematic uncertainties. The affected templates are check-marked. For rate-only uncertainties, the effect on the rate is listed. In case of shape-altering uncertainties, migration effects are taken into account, i.e. events which do not pass the selection or migrate out of the fit range also change the normalisation. Even if the effect is minor for some of the uncertainties, it is not neglected.

Systematics source	Signal	Background	type	rate effect
Normalisation of $t\bar{t}$	☐	☑	rate	20 %
Normalisation of $W \to \ell\nu$	☐	☑	rate	20 %
Normalisation of $Z^0 \to \nu\bar{\nu}$	☐	☑	rate	20 %
Normalisation of QCD	☐	☑	rate	43.9 %
Shape correction QCD	☐	☑	shape + rate	
QCD shape parameterisation	☐	☑	shape + rate	
MC statistics $t\bar{t}$	☐	☑	shape + rate	
MC statistics $W \to \ell\nu$	☐	☑	shape + rate	
MC statistics $Z^0 \to \nu\bar{\nu}$	☐	☑	shape + rate	
Signal statistics	☑	☐	shape + rate	
Luminosity	☑	☑	rate	2.6 %
Trigger	☑	☑	shape + rate	
Pileup	☑	☑	shape + rate	
PDF uncertainty	☑	☐	shape + rate	
Offline b-tag (b c)	☑	☑	shape + rate	
Offline b-tag (u d s g)	☑	☑	shape + rate	
JES	☑	☑	shape + rate	
JER	☑	☑	shape + rate	
Electron energy scale	☑	☑	shape + rate	
Muon energy scale	☑	☑	shape + rate	
Tau energy scale	☑	☑	shape + rate	
Unclustered energy	☑	☑	shape + rate	

5.7.1 Uncertainty Estimation Procedures

The systematic uncertainties considered in the analysis are estimated according to CMS-internal prescriptions or external recommendations, as is indicated in the respective places. The individual procedures are briefly described in this subsection.

Jet- and $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$-Related Uncertainties

Jet-related uncertainties are closely connected to $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainties because they have a strong impact on the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. Therefore, these uncertainties are jointly discussed in the following.

Jet Energy Scale The jet energy scale (JES) is varied by $2-5\,\%$ according to the p_{T}- and η-dependent JES uncertainty estimates, provided centrally by the CMS collaboration. This variation is also propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation using the $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainty tool provided by CMS [230].

Jet Energy Resolution The jet energy resolution (JER) is varied within its estimated η-dependent uncertainty . This variation is also propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation using the above-mentioned uncertainty tool.

Electron/Photon Energy Scale The uncertainty on the energy scale of electrons and photons affects indirectly the final result because these physics objects are an input of the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. The uncertainty on the electron- and photon-energy scale ($0.6\,\%$ in ECAL barrel and $1.5\,\%$ in ECAL endcap) is then propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation with the $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainty tool.

Muon Energy Scale Also muons are not used directly to extract the signal, but their impact on $E_{\mathrm{T}}^{\mathrm{miss}}$ has to be included. Therefore, an uncertainty of $0.2\,\%$, as recommended by the CMS collaboration, is used and propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation.

Tau Energy Scale As is the case for the other leptons, the uncertainty on the energy scale of tau leptons affects the result by its effect on $E_{\mathrm{T}}^{\mathrm{miss}}$. An uncertainty estimate of $3\,\%$ for the tau energy scale is provided by the CMS collaboration and propagated to an uncertainty on $E_{\mathrm{T}}^{\mathrm{miss}}$.

Unclustered Energy Varying the unclustered energy by $\pm 10\,\%$ [188, 230] results in a systematic uncertainty on $E_{\mathrm{T}}^{\mathrm{miss}}$ which is obtained using the above-mentioned uncertainty tool.

b-Tagging Efficiency Scale Factors

The b-tagging scale factors used to correct the b-tagging efficiency in simulation are varied within their uncertainties calculated with the functions provided centrally by CMS [231, 232], which are obtained from b-tagging efficiency measurements [187].

The b-tagging correction factors used to match the full and fast simulation performance are varied within their uncertainties obtained with the payloads provided by CMS [233–236].

PDF Uncertainties

The envelope of the m_{bb} distribution is obtained using the PDF sets provided by CTEQ6.6 [237], MSTW2008 [238–240], and NNPDF [241] following the PDF4LHC recommendation [242].

Luminosity Measurement Uncertainty

The normalisation of all simulation-based backgrounds is varied within the measured luminosity uncertainty of 2.6 % [243] for the full 2012 data set.

Pileup-Related Uncertainty

The systematic uncertainty due to the PU modelling is taken into account by varying the minimum bias cross section used for the reweighting by ±5 % as recommended in [244].

Trigger Uncertainty

The uncertainties of the fitted trigger efficiencies (see Section 5.3.1) are used as systematic uncertainty.

Statistical Uncertainties of Signal and Background Simulation

The shape altering effects of statistical fluctuation in the m_{bb} distribution of the signal and background predictions are taken into account as bin-by-bin uncertainties.

5.7.2 Impact of Individual Systematic Uncertainties

The impact of the systematic uncertainties on the result of the analysis is quantified by their relative impact on the expected $\mathrm{h}_1 \to \mathrm{b\bar{b}}$ cross section limit which is calculated as described in Section 5.6. For each uncertainty, the limit is derived with and without the respective nuisance parameter. The relative variation

$$\Delta l_{\mathrm{exp}} = \frac{l_{\mathrm{exp}}(\text{w/o nuisance}) - l_{\mathrm{exp}}(\text{w/ nuisance})}{l_{\mathrm{exp}}(\text{w/ nuisance})} \tag{5.25}$$

of the expected cross section limit obtained for the individual uncertainties is listed in Table 5.13 for different mass hypotheses. No dominant individual systematic uncertainty can be identified. Only the influence of the b-tagging scale factors, the $t\bar{t}$ normalisation uncertainty, and the QCD-related uncertainties have an impact above the 1 %-threshold. In general, the analysis sensitivity is limited by the statistical precision in the signal region.

Table 5.13: Relative variation of the expected cross section limit in % obtained when omitting the nuisance parameter associated to the individual systematic uncertainty.

Systematics source	$m_{h_1} = 40\,\text{GeV}$	$m_{h_1} = 65\,\text{GeV}$	$m_{h_1} = 85\,\text{GeV}$
Normalisation of $t\bar{t}$	0.969 %	2.298 %	1.774 %
Normalisation of $W \to \ell\nu$	0.000 %	0.000 %	0.000 %
Normalisation of $Z^0 \to \nu\bar{\nu}$	0.000 %	0.000 %	0.000 %
Normalisation of QCD	0.030 %	0.787 %	2.466 %
Shape correction QCD	2.211 %	5.485 %	2.105 %
QCD shape parameterisation	1.211 %	1.143 %	0.105 %
MC statistics $t\bar{t}$	1.333 %	1.193 %	1.414 %
MC statistics $W \to \ell\nu$	0.212 %	0.356 %	0.150 %
MC statistics $Z^0 \to \nu\bar{\nu}$	0.030 %	0.000 %	0.000 %
Signal statistics	0.030 %	0.025 %	0.060 %
Luminosity	0.363 %	0.673 %	0.571 %
Trigger	0.061 %	0.102 %	0.060 %
Pileup	0.000 %	0.178 %	0.120 %
PDF uncertainty	0.197 %	0.203 %	0.241 %
Offline b-tag (b c)	0.757 %	1.549 %	0.707 %
Offline b-tag (u d s g)	0.000 %	0.000 %	0.045 %
JES	0.061 %	1.727 %	1.323 %
JER	0.061 %	0.102 %	0.000 %
Electron energy scale	0.000 %	0.000 %	0.000 %
Muon energy scale	0.000 %	0.000 %	0.000 %
Tau energy scale	0.273 %	0.914 %	0.361 %
Unclustered energy	0.000 %	0.000 %	0.045 %

6 Results and Interpretation

After presenting the analysis details in the previous chapter, the results of the search for light Higgs bosons are provided. The expected event yields in the signal region from the modified P4 scenario are listed in Table 6.1 together with the contributions from the SM backgrounds and the number of observed data events. A comparison of expected background yields and observed data shows good agreement. The final analysis takes full advantage of the signal- and background-shape information.

Table 6.1: Expected event yields for the various signal ($m_{\mathrm{h}_1} = 65\,\mathrm{GeV}$) and background contributions and the rate observed in data.

Contribution	Rate
$\mathrm{h}_1 \to \mathrm{b\bar{b}}$	22.7 ± 3.1
SUSY (non-h_1)	92.4 ± 13.6
$\mathrm{t\bar{t}}$	82.5 ± 16.5
QCD	56.4 ± 24.8
$\mathrm{W} \to \ell\nu$	4.8 ± 1.0
$\mathrm{Z}^0 \to \nu\bar{\nu}$	2.3 ± 0.5
total expected SM background	146 ± 42.8
data	148 ± 12.2

The m_{bb} data histogram is fitted by a model combining the background templates and the signal template. A maximum likelihood fit is performed taking shape uncertainties (see Section 5.7) into account via nuisance parameters (see Section 5.6). Uncertainties on the normalisations of the background templates are incorporated, while the signal normalisation is floating freely. The non-h_1 contributions from the SUSY cascades are treated depending on the interpretation approach as is discussed below.

First, a background-only fit is performed to the observed m_{bb} distribution. The fit result in Fig. 6.1 shows good agreement between data and the SM hypothesis, corresponding to $\chi^2/\mathrm{ndf} = 1.02$ and a fit probability of 43 %. The signal contributions expected from the modified P4 scenario are superimposed for $m_{\mathrm{h}_1} = 65\,\mathrm{GeV}$, illustrating the high sensitivity of the analysis to this scenario.

In this chapter these results are interpreted with two signal definitions as shown in Fig. 6.2. First, a search for single resonant structures on top of the SM background is performed, treating only the $\mathrm{h}_1 \to \mathrm{b\bar{b}}$ contributions as signal (see Fig. 6.2, left). The

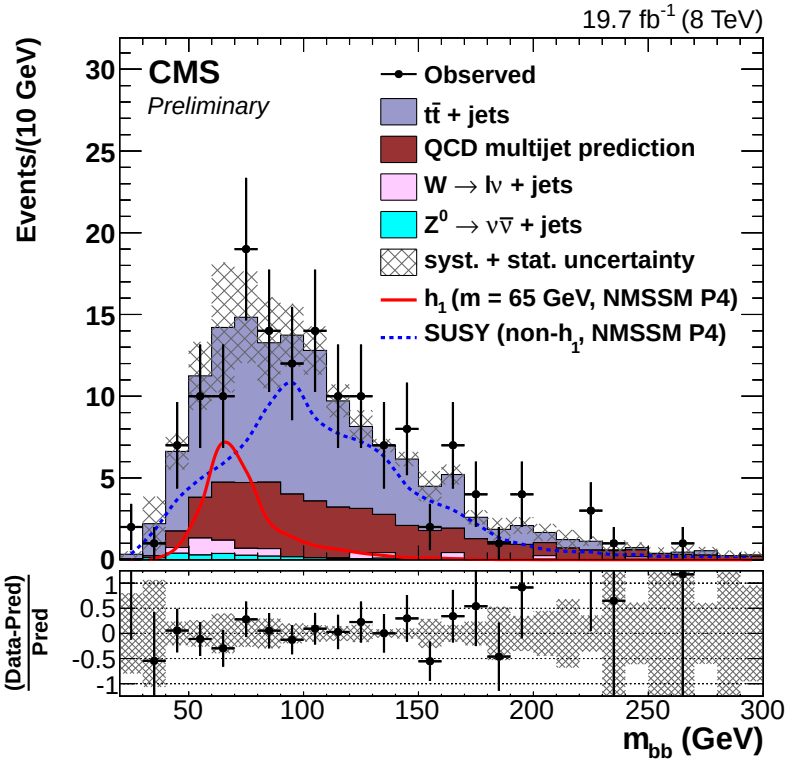

Figure 6.1: Background-only fit result of the invariant mass of the selected pair of b-tagged jets, where the SM background contributions are stacked. The h_1 and non-h_1 contributions are overlaid to illustrate the sensitivity of the analysis to the modified NMSSM P4 scenario.

second approach includes also the non-h_1 contributions from SUSY cascades in the signal definition, to derive model-dependent results within certain NMSSM scenarios.

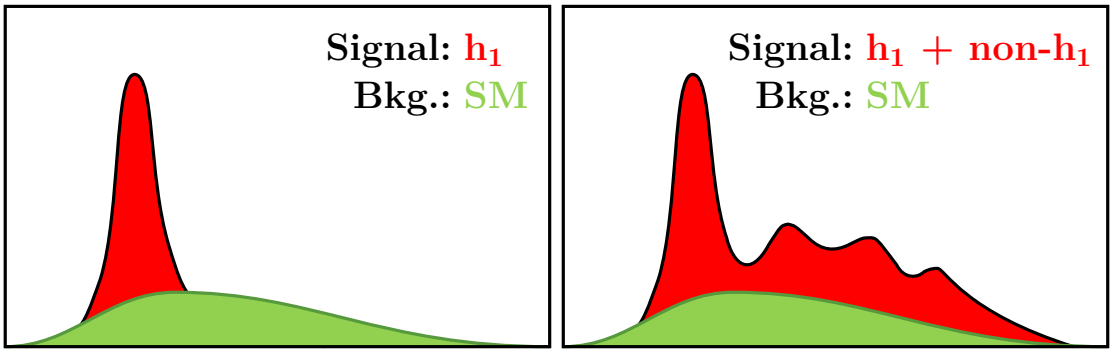

Figure 6.2: The different signal definitions used to interpret the observations. The signal in the left plot is defined as $h_1 \to b\bar{b}$ ignoring the non-h_1 contributions. Therefore, it serves as a minimal-model-dependent approach to obtain limits on $h_1 \to b\bar{b}$. The right plot uses the full NMSSM spectrum as predicted by the scenario under investigation.

6.1 Search for a Single Resonant Structure

The background-only fit result in Fig. 6.1 shows no indication of a broad excess. In order to quantify this observation, the data are probed for a single resonance, under the assumption that such a signal does not contaminate the control regions used to derive the background predictions. Thus, before the spectrum is analysed with model-specific approaches, a measurement of the cross section of h_1-only production times the branching fraction into b quarks is performed in the mass range of 30 – 100 GeV.

6.1.1 Cross Section Measurement

The cross section times branching fraction, $\sigma(pp \rightarrow h_1 + X) \times \mathcal{B}(h_1 \rightarrow b\bar{b})$, is measured using the h_1-only signal template (as sketched in Fig. 6.2, left) by means of a maximum likelihood fit. Figure 6.3 shows the best-fit value of different h_1-mass hypotheses within 30 – 100 GeV. No significant excess is observed and the largest upward fluctuation above the core part of the SM background is observed at a mass of 75 GeV (see Fig. 6.4).

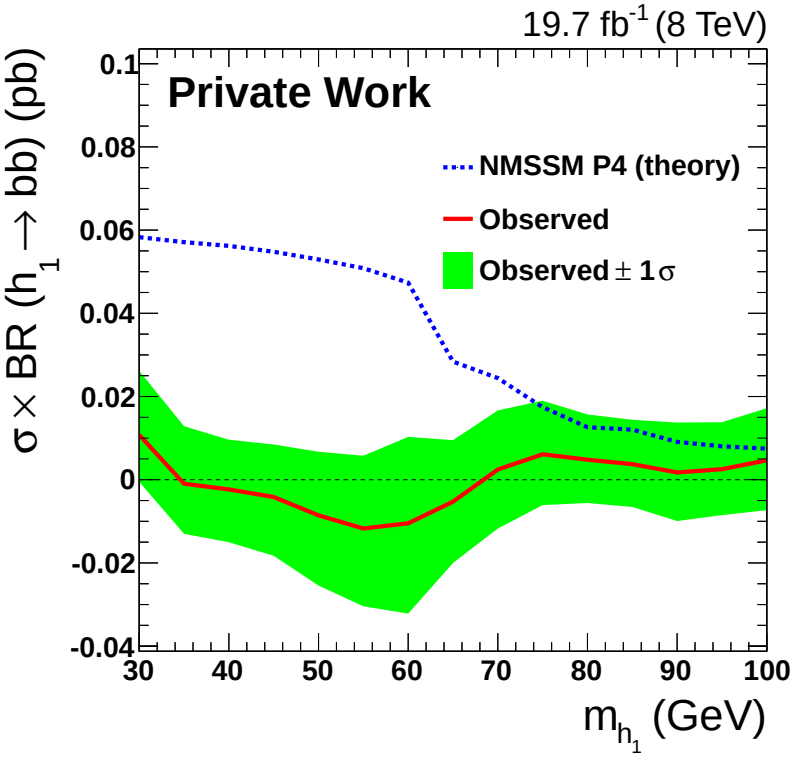

Figure 6.3: Measured cross section of a light resonance decaying into a pair of b quarks. No significant excess is observed. The detailed fitted signal strength modifiers are listed in Appendix G.2.

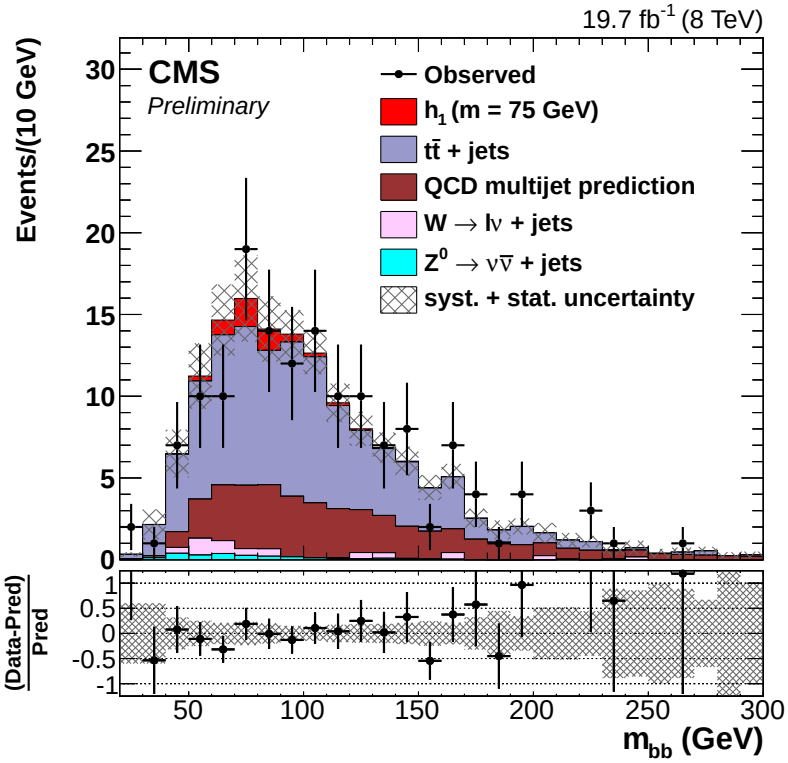

Figure 6.4: Invariant mass of the selected pair of b-tagged jets, where the SM background contributions are stacked. The h_1 template with a mass hypothesis of 75 GeV and the SM-background templates are fitted to the data.

6.1.2 Upper Limits on Cross Sections

In absence of signal, upper limits on the cross section times the branching fraction, $\sigma(pp \to h_1 + X) \times \mathcal{B}(h_1 \to b\bar{b})$, are derived. The chosen test statistic used to determine how signal- and background-like the data are and the procedure to derive upper limits are described in Section 5.6.1 and Section 5.6.2, respectively.

The upper limits on the $h_1 \to b\bar{b}$ cross section times branching fraction are derived using the signal definition of Fig. 6.2 (left) in order to obtain a minimal-model-dependent interpretation of the results. The results of this search for a singly produced Higgs boson peak using the SM expectations as Asimov data set (see Section 5.6.2) are shown in Fig. 6.5 and are compared with the expectations for $h_1 \to b\bar{b}$ from the modified P4 scenario (blue dashed line).

The h_1-production rate decreases with increasing h_1 mass due to the shrinking phase space for neutralino decays into h_1. The abrupt decline between $m_{h_1} = 60\,\text{GeV}$ and $m_{h_1} = 65\,\text{GeV}$ is due the closing phase space for the $h_2 \to h_1 h_1$ decay. The trend of the expected limits is a joint effect of signal efficiency (see Fig. 5.5, left) and the background distribution. Figure 6.1 shows that for increasing m_{h_1} the Higgs peak is shifted towards the core part of the expected SM m_{bb} spectrum, making it less prominent with respect to the background expectations. Thus, the effect of the increasing efficiency (see Fig. 5.5, left) is compensated. The observed limit lies

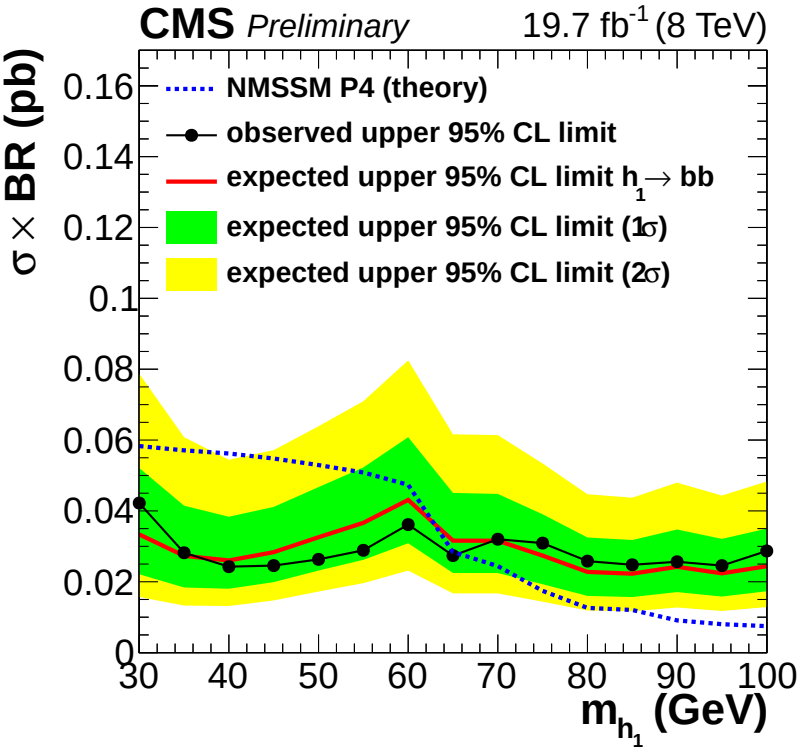

Figure 6.5: Upper limits for the light Higgs boson production cross section times branching fraction, $\sigma(\mathrm{pp} \to \mathrm{h}_1 + X) \times \mathcal{B}(\mathrm{h}_1 \to \mathrm{b\bar{b}})$, in comparison to the theoretical expectations from the modified NMSSM P4 scenario (blue dashed line). The step in the NMSSM prediction is due to the transition from the phase space where the decay $\mathrm{h}_2 \to \mathrm{h}_1\mathrm{h}_1$ is possible to the regime where this is kinematically not allowed.

within the 1σ band of the expected limit, demonstrating the agreement of the data with the SM expectations. Below $m_{\mathrm{h}_1} = 65\,\mathrm{GeV}$ these limits go below the $\mathrm{h}_1 \to \mathrm{b\bar{b}}$ rate expected from the modified P4 scenario. Thus, the analysis is sensitive to h_1 production as predicted in this scenario.

6.2 Interpretation in the Context of NMSSM Scenarios

The obtained results are also interpreted within NMSSM scenarios, thus providing information on the model parameter space. The full NMSSM $\to \mathrm{b\bar{b}}$ spectrum (see Fig. 6.2, right) is used, thus profiting from the whole m_{bb} shape information. Upper limits are derived in the modified P4 scenario and the decoupled squarks scenario (see Section 1.4.2).

6.2.1 Modified P4 Scenario

In the modified P4 scenario, different h_1-mass hypotheses are tested, thereby also influencing the other NMSSM contributions. Upper limits are derived in the h_1-mass range of 30 – 100 GeV.

Figure 6.6 shows that the analysis excludes over the full h_1-mass range the modified P4 scenario with a mass scale of 1 TeV for gluinos and squarks ($M_3 = M_{\mathrm{SUSY}} = 1\,\mathrm{TeV}$). Except for a small step between $m_{h_1} = 60\,\mathrm{GeV}$ and $m_{h_1} = 65\,\mathrm{GeV}$, the limits depend only mildly on m_{h_1}. The decrease of the limit below $m_{h_1} = 65\,\mathrm{GeV}$ is caused by the opening of the $h_2 \to h_1 h_1$ decay. This results on average in a larger number of bosons per event leading to a slightly different event topology and a higher rate of h_1 production for lower h_1 masses. Thus, the overall m_{bb} profile is sharpened.

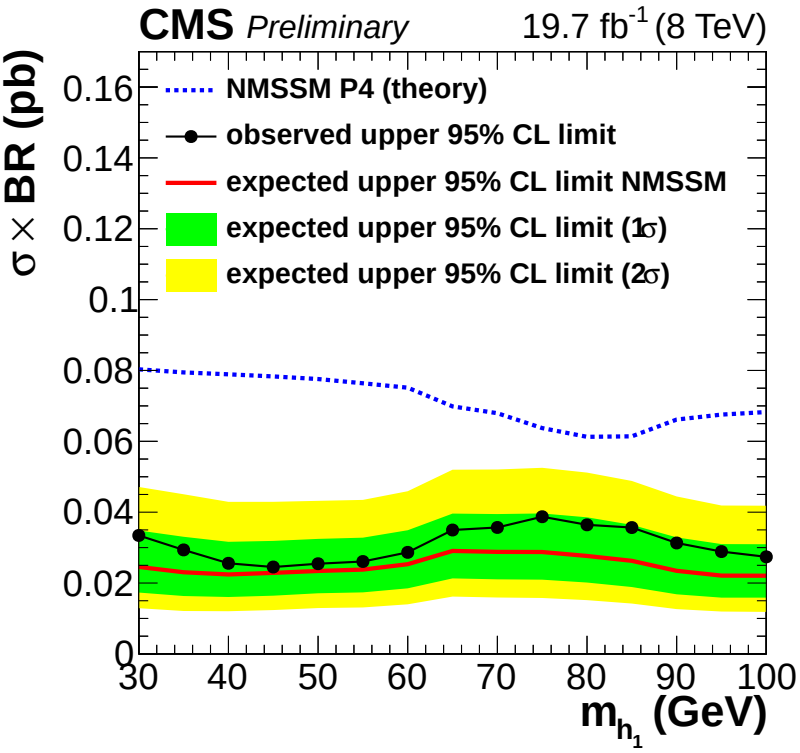

Figure 6.6: Upper limits for the NMSSM cross section times branching fraction into $b\bar{b}$ in comparison to the theoretical expectations from the NMSSM P4 scenario (blue dashed line).

The otherwise small dependence on the h_1 mass can be explained with the rather small h_1-mass range compared to the probed m_{bb} range (see Fig. 5.22). Therefore, the variation of m_{h_1} merely changes the left flank of the NMSSM $\to b\bar{b}$ signal template. The observed limit is slightly higher, but still below the expectation from the modified P4 scenario.

The rise in the theoretical expectation above 85 GeV is due to the opening phase space for additional production of a_1 bosons, whose mass is inversely proportional to m_{h_1}. Hence, the overall Higgs boson production in neutralino decays increases.

Under the assumption that the signal efficiency is independent of the gluino- and squark-mass parameters M_3 and M_{SUSY}, the cross section limits in Fig. 6.6 are translated into exclusion limits on a common mass parameter for M_{SUSY} and M_3. The results in Fig. 6.7 ignore possible shape dependencies on the mass of the coloured SUSY particles. However, the signal efficiencies in the decoupled squarks scenario (see Fig. 5.7) show that the efficiency significantly increases for higher M_3 values. Therefore, the exclusion limits in Fig. 6.7 are considered as a conservative lower bound on the common mass of coloured SUSY particles.

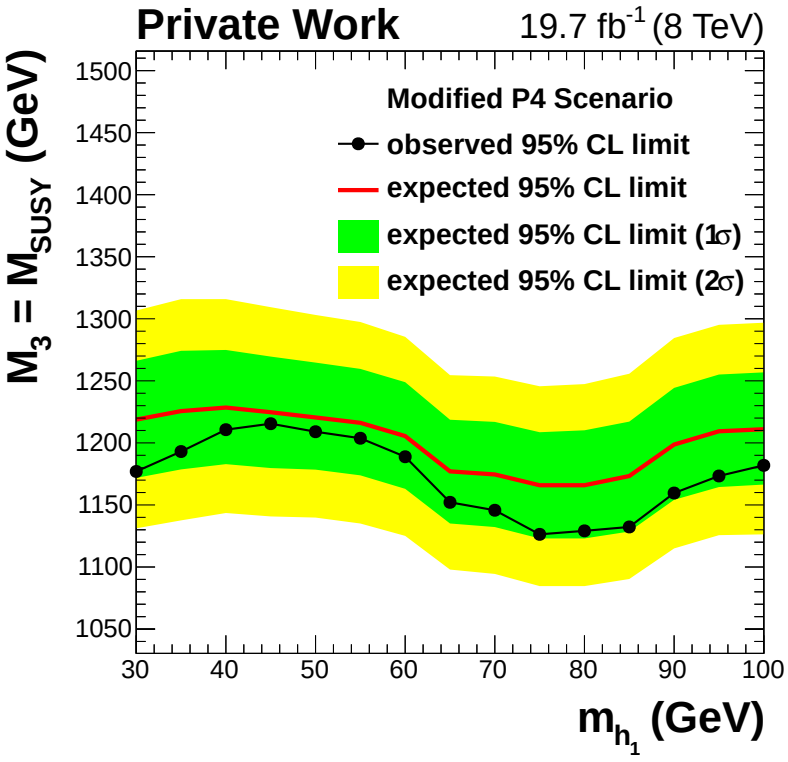

Figure 6.7: Exclusion limits on a common mass parameter M_3/M_{SUSY} in the modified P4 scenario. These limits are derived under the assumption that the selection efficiency is independent of M_3/M_{SUSY}.

6.2.2 Decoupled Squarks Scenario

The decoupled squarks scenario is investigated in a larger region of the NMSSM phase space. The four parameters that change the cascade decay structure most with respect to light Higgs searches are M_1, M_2, M_3, and m_{h_1}, i.e. the two gaugino mass parameters, the gluino mass parameter, and the mass of the light Higgs boson. The probed parameter points are listed in detail in Appendix C.1. This NMSSM parameter scan is performed based on the outcome of the studies presented in Chapter 2.

The exclusion limits on the signal strength modifier within the decoupled squarks scenario are translated into exclusion limits in the M_3–m_{h_1} plane. From the summary

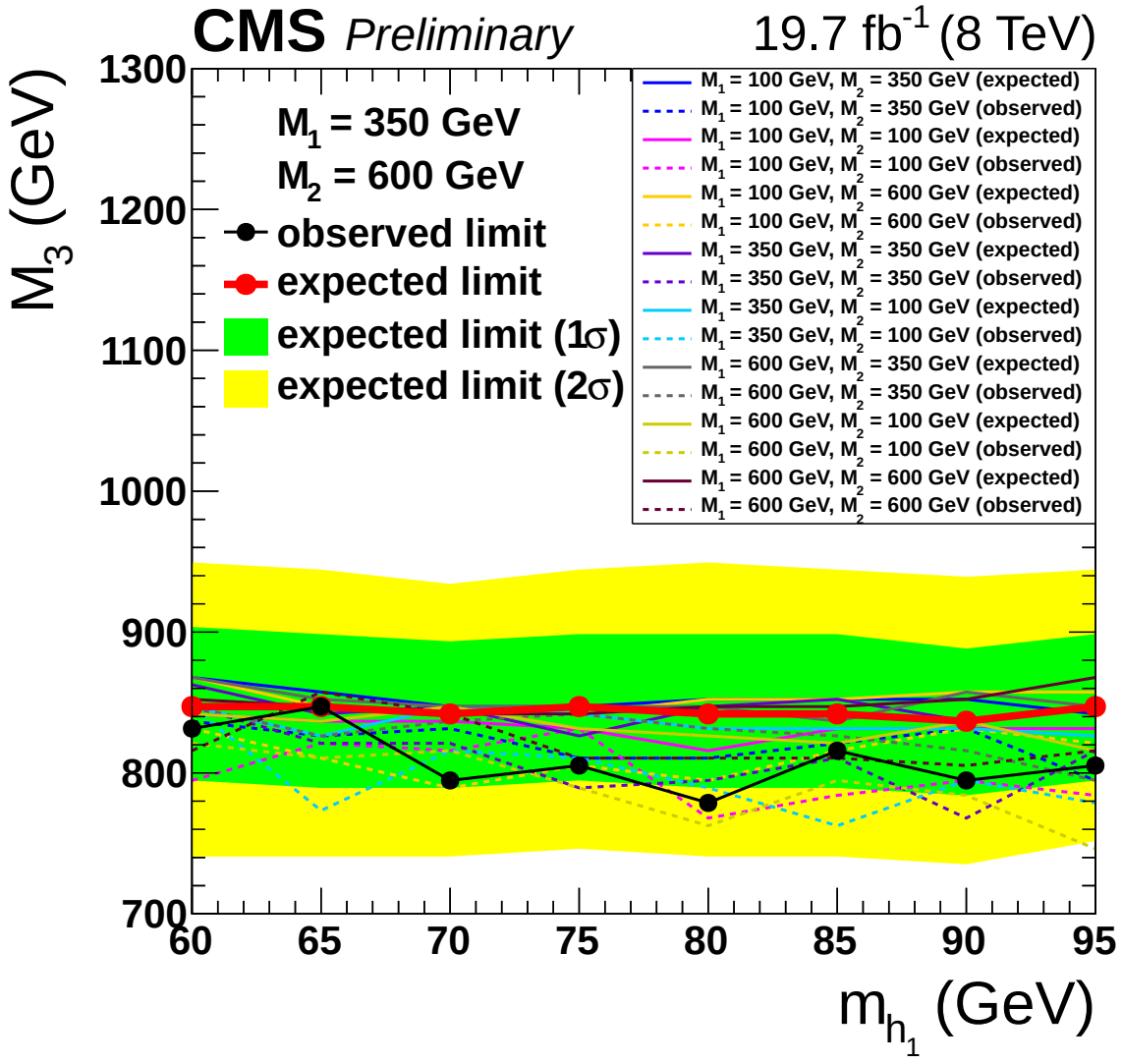

Figure 6.8: Exclusion limits in the M_3–m_{h_1} plane of the NMSSM in the scenario of decoupled squarks for various combinations of M_1 and M_2. In this approach the full NMSSM spectrum is used (see Fig. 6.2, right).

plot of the parameter scan in Fig. 6.8 one can infer little dependence of the limits on the gaugino mass parameters M_1, M_2, and on the h_1 mass. The latter shows qualitatively a similar behaviour as in the modified P4 scenario, i.e. a step between $m_{\mathrm{h}_1} = 60\,\mathrm{GeV}$ and $m_{\mathrm{h}_1} = 65\,\mathrm{GeV}$ is observed (note the different h_1 mass range).

The production cross section of the initial particles of the SUSY cascades has the largest impact on the results, leading to a strong dependence on the gluino mass parameter M_3. The sensitivity of the results on the varied masses, production cross sections, and branching fractions of the individual particles is at most mild. Since, for the signal extraction all NMSSM $\rightarrow \mathrm{b\bar{b}}$ contributions are used, the effect of the parameters M_1, M_2 and m_{h_1} is compensated. In addition, the event topology is slightly different than in the modified P4 scenario, where squark production dominates (see

Table 1.7). Since the only contributing production mode in the decoupled squarks scenario is gluino pair production, the fraction of the non-resonant contributions in Fig. 5.1 (bottom left) is increased. Thus, non-resonant contributions play a more important role, leading to a further dilution of the effects on the individual resonances in the decoupled squarks scenario. A detailed overview of all combinations of M_1 and M_2 is shown in Fig. 6.9; the observed limits are as in the modified P4 scenario consistent with the expected limits.

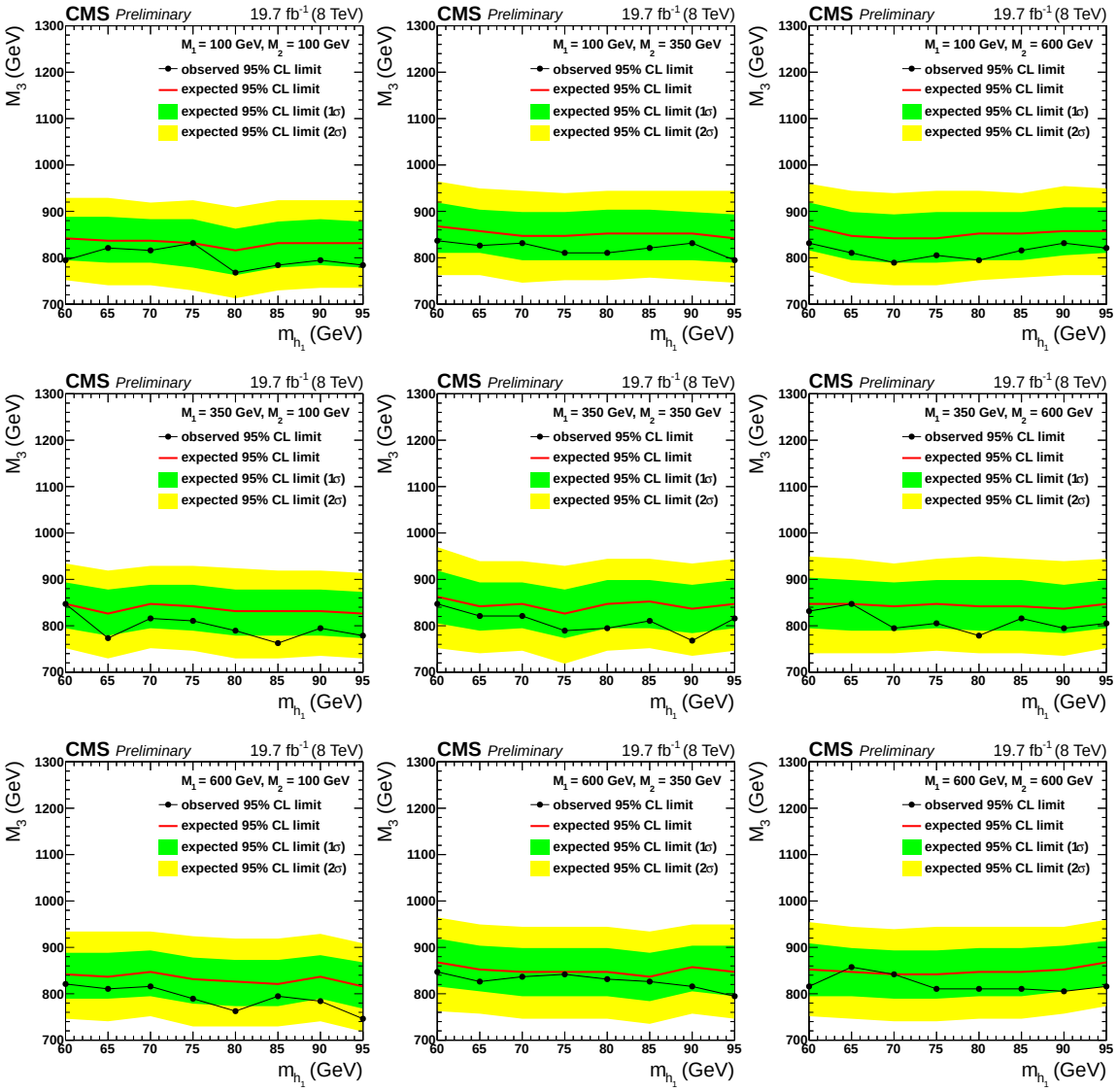

Figure 6.9: Exclusion limits in the M_3–m_{h_1} plane of the NMSSM in the decoupled squarks scenario for various combinations of M_1 and M_2. The exclusion curves show little dependence on the value of the h_1 mass and on the combination of (M_1, M_2).

7 Summary and Outlook

A search for a light Higgs boson with a mass below 100 GeV, as predicted in certain NMSSM scenarios, is performed. At the LHC, only islands have been probed previously in this mass range [245, 246], thus the results of the presented analysis are unique.

The theoretical production properties of light Higgs bosons in SUSY cascades are studied in more detail, analysing in particular the branching fractions of squarks into light Higgs bosons via intermediate neutralinos and charginos. The study shows that decay chains with two or more intermediate particles are most likely. This motivates a full-model scan for the interpretation of the results because it is more suitable than simplified models.

A search is performed for a light Higgs boson produced in cascade decays of strongly interacting SUSY particles, where the light Higgs boson decays into b quarks. The data recorded by the CMS experiment in 2012 at a centre-of-mass energy of 8 TeV, corresponding to an integrated luminosity of 19.7 fb^{-1}, are analysed. Events with two b-tagged jets, two untagged jets, and large $E_{\mathrm{T}}^{\mathrm{miss}}$ in the final state are selected. The contributions from SM processes are estimated from simulation as for the dominant background $\mathrm{t\bar{t}}$, and with a data-driven method in case of the QCD-multijet contributions.

The signal is searched for in the invariant mass distribution of the two selected b-tagged jets. No significant deviation from the expected background from SM processes is observed. The results are interpreted in terms of upper exclusion limits on the light-Higgs-boson production cross-section times the branching fraction into b quarks, $\sigma(\mathrm{pp} \to \mathrm{h}_1 + X) \times \mathcal{B}(\mathrm{h}_1 \to \mathrm{b\bar{b}})$, which is observed to be less than 0.042 pb at 95 % confidence level. The analysis excludes Higgs-production cross-sections that are below the expected h_1 contribution in the modified NMSSM P4 benchmark scenario for h_1 masses below 65 GeV. The model-specific interpretation utilising the full NMSSM spectrum excludes the modified P4 scenario over the whole probed mass range of 30 – 100 GeV assuming a mass of 1 TeV for the coloured SUSY particles. Exclusion limits are derived in a parameter scan within the decoupled squarks scenario of the NMSSM. Values of the gluino mass parameter M_3 below ≈ 800 GeV are excluded for all combinations of the other scanned parameters. The results of this analysis are published in [209].

While this thesis is written, the LHC is restarted for its second run. A foreseen total integrated luminosity at the order of 200 – 500 fb^{-1} [247] for the whole run will improve the precision of measurements of the Higgs-boson properties. Possible deviations from the SM expectation could then indicate new phenomena. In the context of direct searches for new heavy particles, the increased centre-of-mass energy of 13 TeV is even more important. These particles might be additional heavy Higgs

bosons, or SUSY particles. Although, the analysis presented in this thesis searches for light Higgs bosons, the higher accessible mass scale is beneficial in this regard, too. The increased production cross section of coloured SUSY particles would increase the cross section of the dominant production channel of a light h_1 boson. Thus, prospects are good and answers to the important questions can be expected.

Appendix A

Superfield Formalism

When dealing with superfields the notation of [74, 76] is used.

A.1 Graßmann Variables

Graßmann variables θ_i are elements of an Algebra with the inner product

$$\{\theta_i, \theta_j\} := \theta_i\theta_j + \theta_j\theta_i = 0 \; . \tag{A.1}$$

The algebra can be extended by an involution $\mathcal{I}$

$$\mathcal{I} : \theta_i \to \mathcal{I}(\theta_i) \equiv \bar{\theta}_i \; , \qquad\qquad \mathcal{I}(\mathcal{I}(\theta_i)) = \theta_i \; , \tag{A.2}$$

where the conjugate Graßmann variables $\bar{\theta}_i$ fulfil the following relations

$$\{\bar{\theta}_i, \bar{\theta}_j\} = 0 \qquad \text{and} \qquad \{\theta_i, \bar{\theta}_j\} = 0 \; . \tag{A.3}$$

Because of (A.1) and (A.2) every function, which depends on Graßmann variables, can be expanded as follows

$$f(\theta_i, \bar{\theta}_j) = a_0 + \sum_i a_i\theta_i + \sum_i \bar{a}_i\bar{\theta}_i + \cdots \; . \tag{A.4}$$

Furthermore, differentiation and integration can be defined on a Graßmann algebra

$$\frac{\partial}{\partial\theta_i}\theta_j := \delta_{ij} \; , \tag{A.5a}$$

$$\frac{\partial}{\partial\theta_i}\theta_j\theta_k := \delta_{ij}\theta_k - \delta_{ik}\theta_j \; , \tag{A.5b}$$

$$\int \mathrm{d}\theta_i := 0 \; , \tag{A.5c}$$

$$\int \mathrm{d}\theta_i \; \theta_j := \delta_{ij} \; , \tag{A.5d}$$

with $\{\mathrm{d}\theta_i, \mathrm{d}\theta_j\} = 0$.

A.2 Weyl Spinor Notation

In order to formulate supersymmetric quantum field theories in an elegant and as convenient as possible way, one decomposes the four-component Dirac spinors ψ in two-component Weyl spinors,

$$\psi_a = \begin{pmatrix} \xi_A \\ \bar{\chi}^{\dot{A}} \end{pmatrix} , \qquad (a = 1,2,3,4) , (A = 1,2) , (\dot{A} = \dot{1}, \dot{2}) , \tag{A.6}$$

since the spinor components ξ_A, $\bar{\chi}^{\dot{A}}$ transform as fundamental representations of the Lorentz group $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$, respectively. The following relations hold for these spinors:

$$\xi_A = (\bar{\xi}_{\dot{A}})^\dagger , \qquad \xi^A = (\bar{\xi}^{\dot{A}})^\dagger , \tag{A.7a}$$
$$\xi^A = \epsilon^{AB}\xi_B , \qquad \xi_A = \epsilon_{AB}\xi^B , \tag{A.7b}$$
$$\bar{\chi}^{\dot{A}} = \epsilon^{\dot{A}\dot{B}}\bar{\chi}_{\dot{B}} , \qquad \bar{\chi}_{\dot{A}} = \epsilon_{\dot{A}\dot{B}}\bar{\chi}^{\dot{B}} , \tag{A.7c}$$

where the antisymmetric tensors ϵ^{AB} and ϵ_{AB} are given as

$$\epsilon^{AB} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} , \qquad \epsilon_{AB} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} , \tag{A.8}$$

and it holds $\epsilon^{AB}\epsilon_{BC} = \delta^A_C$. Furthermore, one defines that the spinor components ξ_A and $\bar{\chi}^{\dot{A}}$ are Graßmann variables, i.e.

$$\{\xi_A, \xi_B\} = \{\bar{\chi}^{\dot{A}}, \bar{\chi}^{\dot{B}}\} = \{\xi_A, \bar{\chi}^{\dot{B}}\} = 0 . \tag{A.9}$$

One now constructs an $SL(2,\mathbb{C})$ invariant spinor product,

$$\xi\chi := \xi^A\chi_A , \tag{A.10a}$$
$$\bar{\chi}\bar{\xi} := \bar{\chi}_{\dot{A}}\bar{\xi}^{\dot{A}} = (\xi\chi)^\dagger , \tag{A.10b}$$

which is symmetric, due to the antisymmetry of ϵ_{AB} and (A.9),

$$\xi\chi = \chi\xi = (\bar{\xi}\bar{\chi})^\dagger = (\bar{\chi}\bar{\xi})^\dagger . \tag{A.11}$$

Furthermore, one defines differential operators,

$$\partial_A := \frac{\partial}{\partial\theta^A} , \qquad \partial^A := \frac{\partial}{\partial\theta_A} , \qquad \bar{\partial}^{\dot{A}} := \frac{\partial}{\partial\bar{\theta}_{\dot{A}}} , \qquad \bar{\partial}_{\dot{A}} := \frac{\partial}{\partial\bar{\theta}^{\dot{A}}} , \tag{A.12}$$

with the following anticommutator relations:

$$\{\partial_A, \partial_B\} = \{\bar{\partial}_{\dot{A}}, \bar{\partial}_{\dot{B}}\} = \{\partial_A, \bar{\partial}_{\dot{B}}\} = 0 . \tag{A.13}$$

By definition these act on Graßmann variables as follows,

$$\partial_A \theta^B := \delta_A^B \,, \qquad \partial^A \theta_B := \delta_B^A \,, \qquad \bar{\partial}_{\dot{A}} \bar{\theta}^{\dot{B}} := \delta_{\dot{A}}^{\dot{B}} \,, \qquad \bar{\partial}^{\dot{A}} \bar{\theta}_{\dot{B}} := \delta_{\dot{B}}^{\dot{A}} \,, \tag{A.14a}$$

$$\partial^A \theta^B := -\epsilon^{AB} \,, \quad \partial_A \theta_B := -\epsilon_{AB} \,, \quad \bar{\partial}_{\dot{A}} \bar{\theta}_{\dot{B}} := -\epsilon_{\dot{A}\dot{B}} \,, \quad \bar{\partial}^{\dot{A}} \bar{\theta}^{\dot{B}} := -\epsilon^{\dot{A}\dot{B}} \,, \tag{A.14b}$$

and transform according to

$$\epsilon^{AB} \partial_B = -\partial^A \,, \quad \epsilon_{AB} \partial^B = -\partial_A \,, \quad \epsilon_{\dot{A}\dot{B}} \bar{\partial}^{\dot{B}} = -\bar{\partial}_{\dot{A}} \,, \quad \epsilon^{\dot{A}\dot{B}} \bar{\partial}_{\dot{B}} = -\bar{\partial}^{\dot{A}} \,. \tag{A.15}$$

Finally, a few shorthand notations are introduced,

$$\mathrm{d}^2\theta := -\frac{1}{4} \mathrm{d}\theta^A \mathrm{d}\theta_A \,, \qquad \mathrm{d}^2\bar{\theta} := -\frac{1}{4} \mathrm{d}\bar{\theta}_{\dot{A}} \mathrm{d}\bar{\theta}^{\dot{A}} \,, \qquad \mathrm{d}^4\theta := \mathrm{d}^2\theta \, \mathrm{d}^2\bar{\theta} \,, \tag{A.16}$$

in order to construct projectors for the individual components of a superfield (1.22). This yields

$$\int \mathrm{d}^2\theta \, \mathcal{F}(z) = M(x) \,, \tag{A.17a}$$

$$\int \mathrm{d}^2\bar{\theta} \, \mathcal{F}(z) = N(x) \,, \tag{A.17b}$$

$$\int \mathrm{d}^4\theta \, \mathcal{F}(z) = \frac{1}{2} D(x) \,. \tag{A.17c}$$

Appendix B

Feynman Diagrams of Squark and Gluino Production

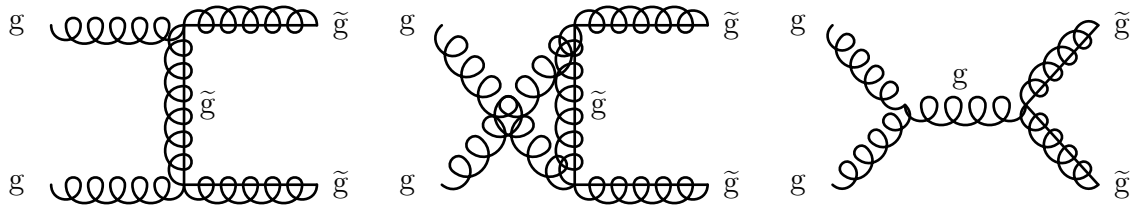

Figure B.1: Gluino production in gluon scattering.

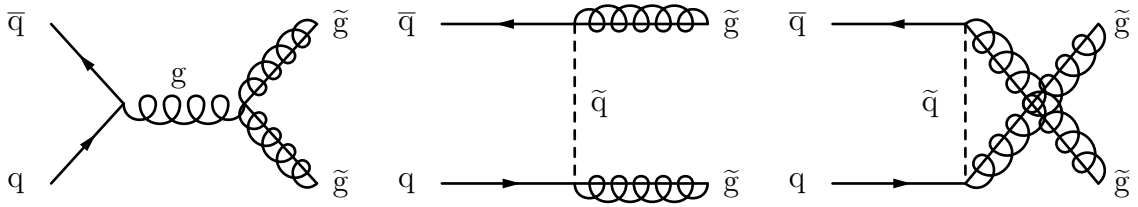

Figure B.2: Gluino production in quark scattering.

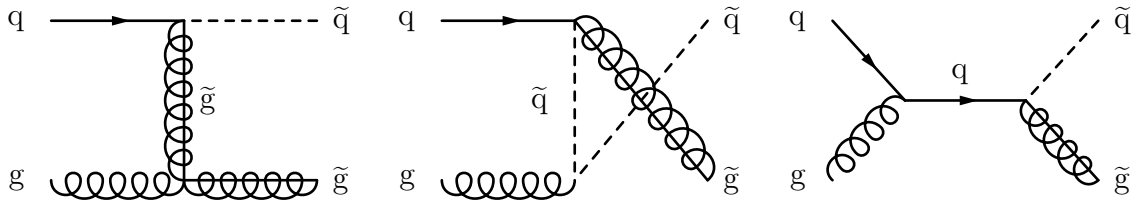

Figure B.3: Squark-gluino production in quark-gluon scattering.

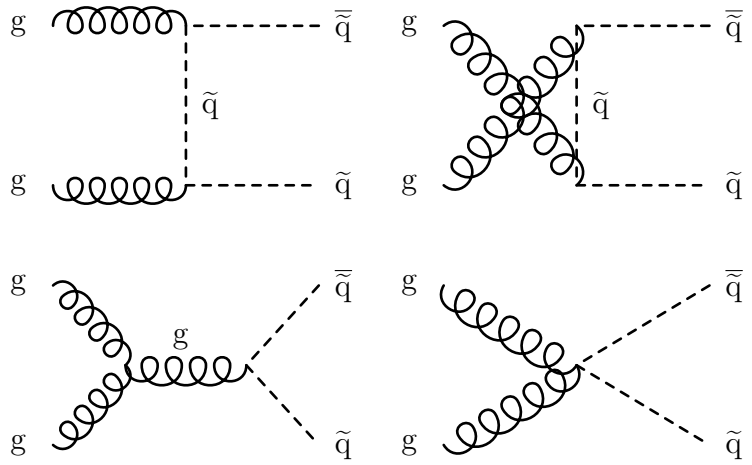

Figure B.4: Squark production in gluon scattering.

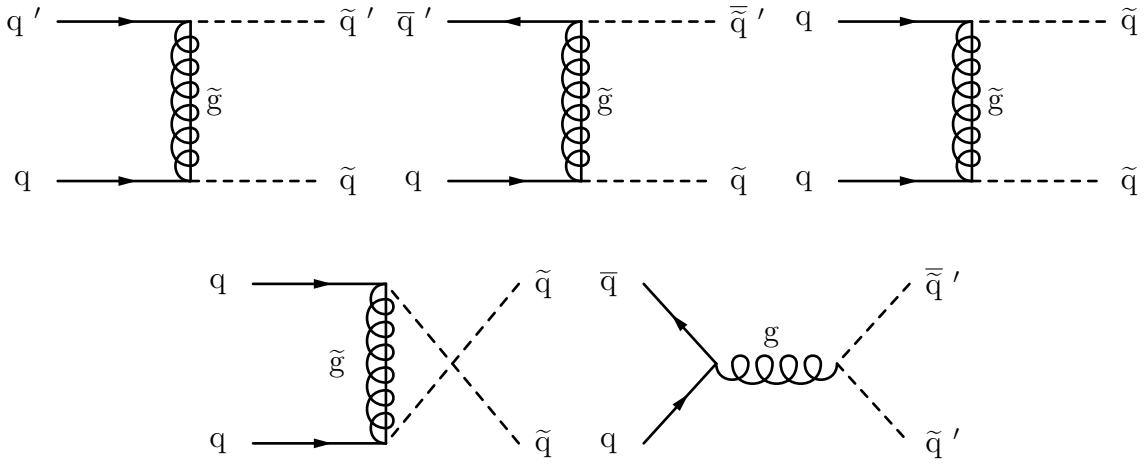

Figure B.5: Squark production in quark scattering.

Appendix C

Decoupled Squarks Scenario of the NMSSM

C.1 NMSSM Parameter Scan Points

The NMSSM parameter scan (see Table C.1) in the decoupled squarks scenario still has the light-Higgs features of the modified P4 scenario and is therefore beneficial for h_1 production while taking into account the possible effect that a slightly differing mass hierarchy in the neutralino sector would have.

Table C.1: Scanned values of NMSSM parameters. All other parameters are in the same regime as in the modified P4 scenario (see Table 1.6), except for the squark masses, which are set to 2 TeV.

NMSSM parameter	scanned values (GeV)
m_{h_1}	$60, 65, 70, 75, 80, 85, 90, 95$
M_1	$100, 350, 600$
M_2	$100, 350, 600$
M_3	$800, 1000, 1200, 1400$

C.2 $\mathcal{B}$(NMSSM $\to b\bar{b}$) in Decoupled Squarks Scenario

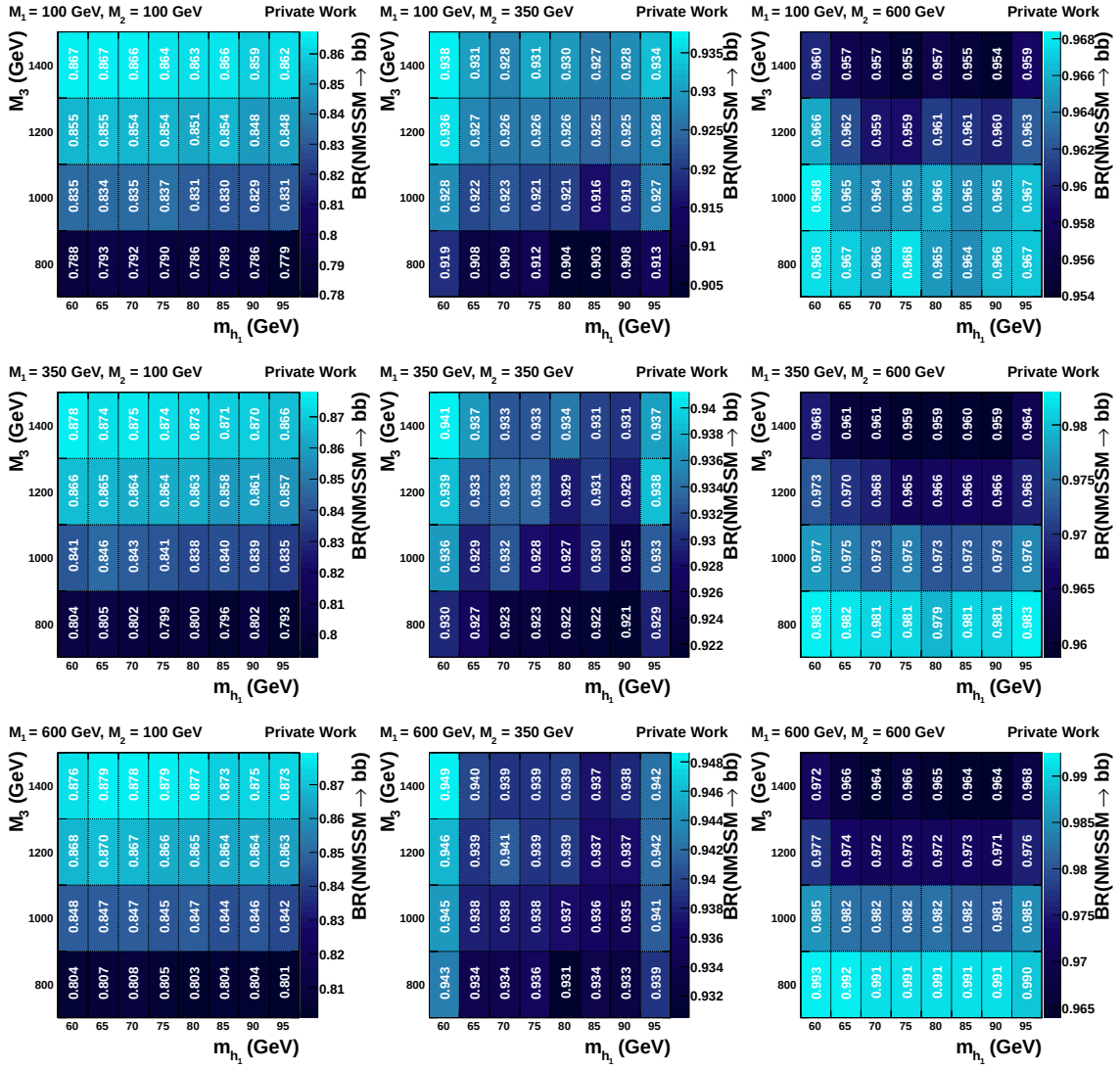

Figure C.1: Branching fractions into NMSSM $\to b\bar{b}$ for the various combinations of M_1 and M_2.

Appendix D

Offline Event Selection

D.1 Event Selection Variables

The variables used to select events are shown in Figs. D.1 – D.4 before the corresponding selection criterion is applied. For all of them the distributions, normalised to unity, are shown for signal and backgrounds to demonstrate the discriminating power of those variables. A comparison to data is also provided for variables where the background is not dominated by QCD-multijet events. Because a data-driven QCD prediction is only done for the signal region (see Section 5.4.2) QCD-multijet simulation is used for illustration and the QCD contribution is scaled to match the data normalisation. The shape of the background obtained by this procedure is in agreement with the data distribution for the variables shown in Figs. D.2 – D.4. For the variables where QCD is still dominant one cannot appreciate the difference between QCD and the other backgrounds with this simple approach. Therefore these data-to-simulation comparisons are omitted.

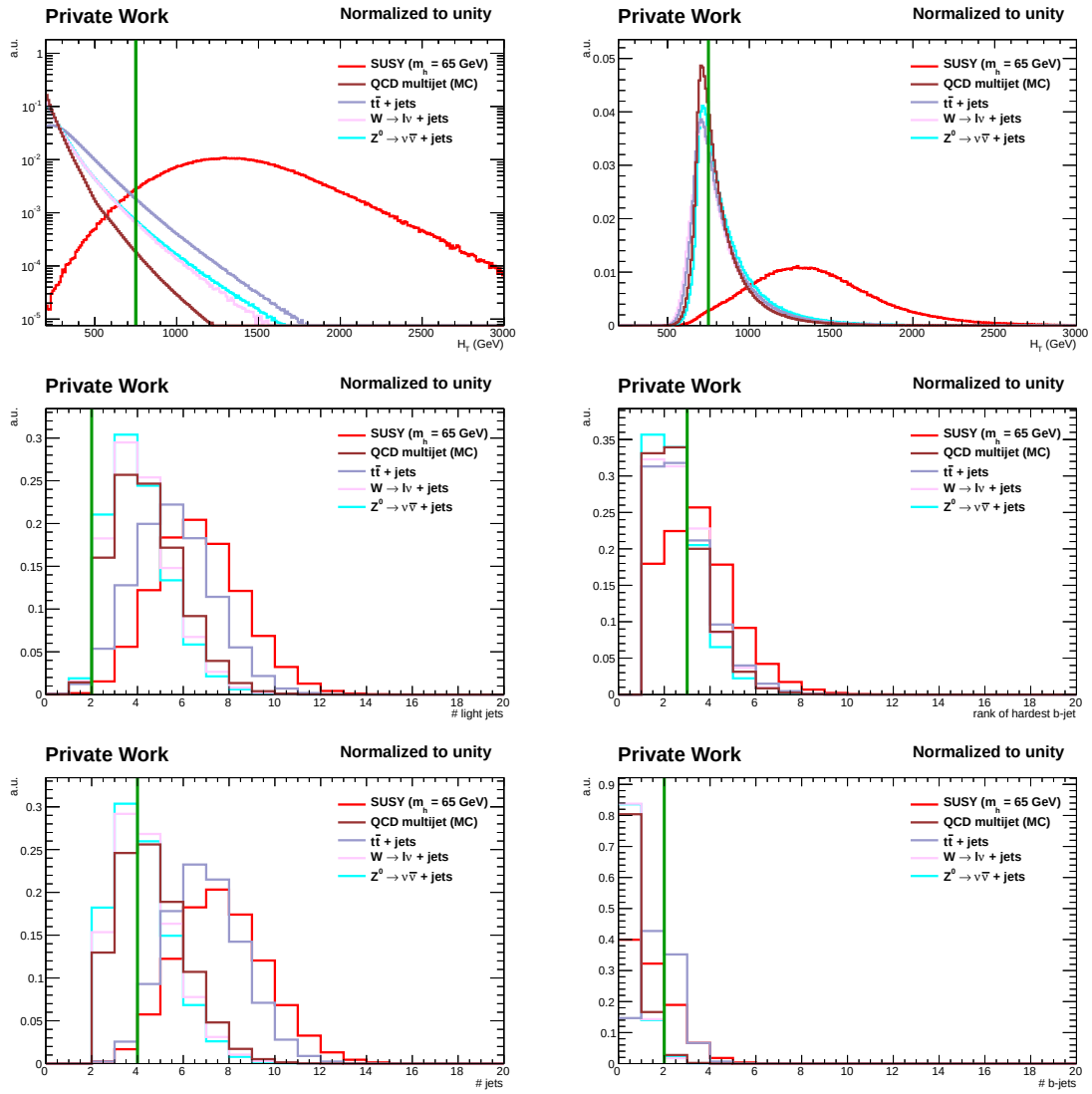

Figure D.1: A subset of the variables on which the event selection is based. For the H_{T} variable the distribution with (top left) and without (top right) trigger requirement is shown. The former illustrates the discriminating power of the chosen cut value. The multiplicity of non-b-tagged jets (centre left) shows that some of the SM backgrounds would be reduced by (#light jets $\gtrsim 4$), however the main background $t\bar{t}$ is not reduced. In addition one would introduce a higher model dependency when requiring a higher number of non-b-tagged jets. The b veto on the two leading jets, which means requiring the rank of the hardest b-tagged jet (centre right) to be greater than 2 clearly separates signal and background. The argument made for the number of non-b-tagged jets is also applicable to jet multiplicity (bottom left) and the number of b-tagged jets reduces also a sizeable fraction of the QCD-multijet background.

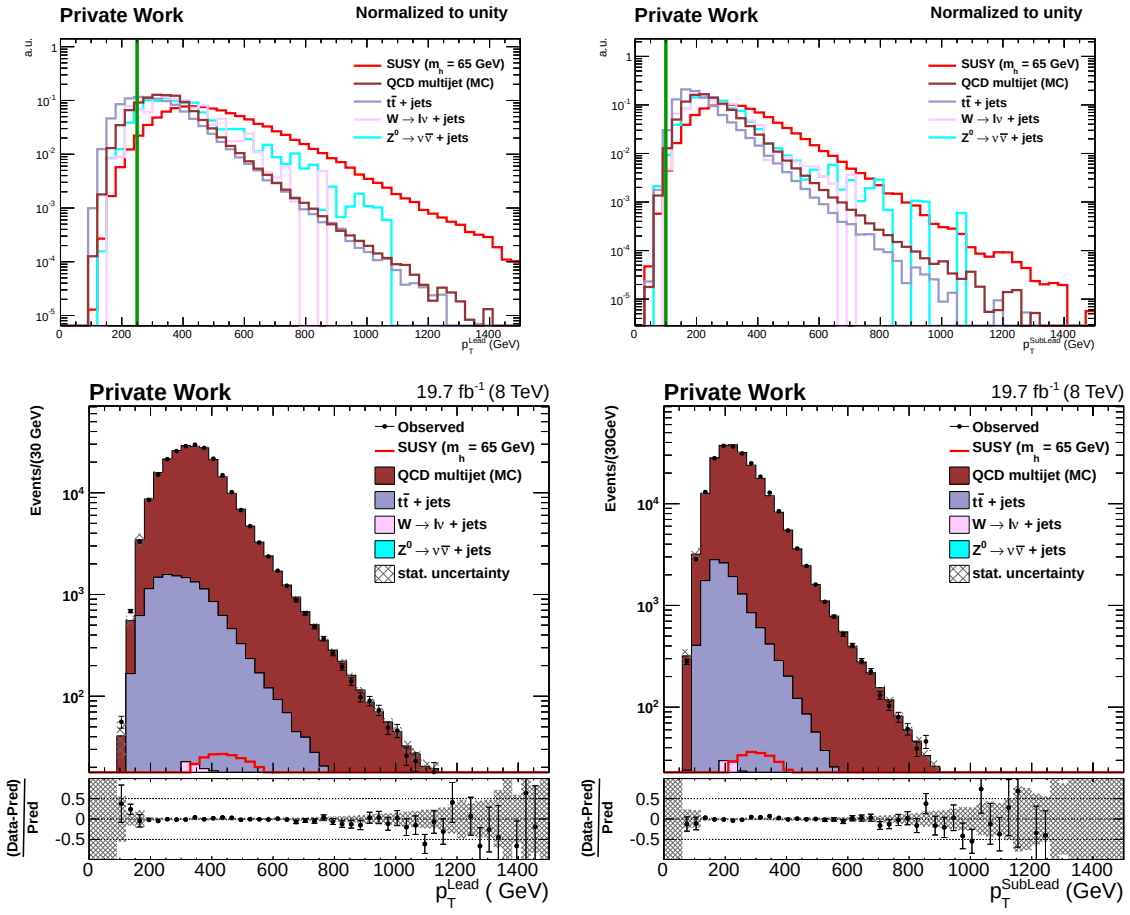

Figure D.2: The spectra of the leading (left) and subleading (right) jet are shown. The background shape is in agreement with the data distribution. QCD MC is used for illustration purposes only. The analysis uses a data-driven estimate of the QCD background.

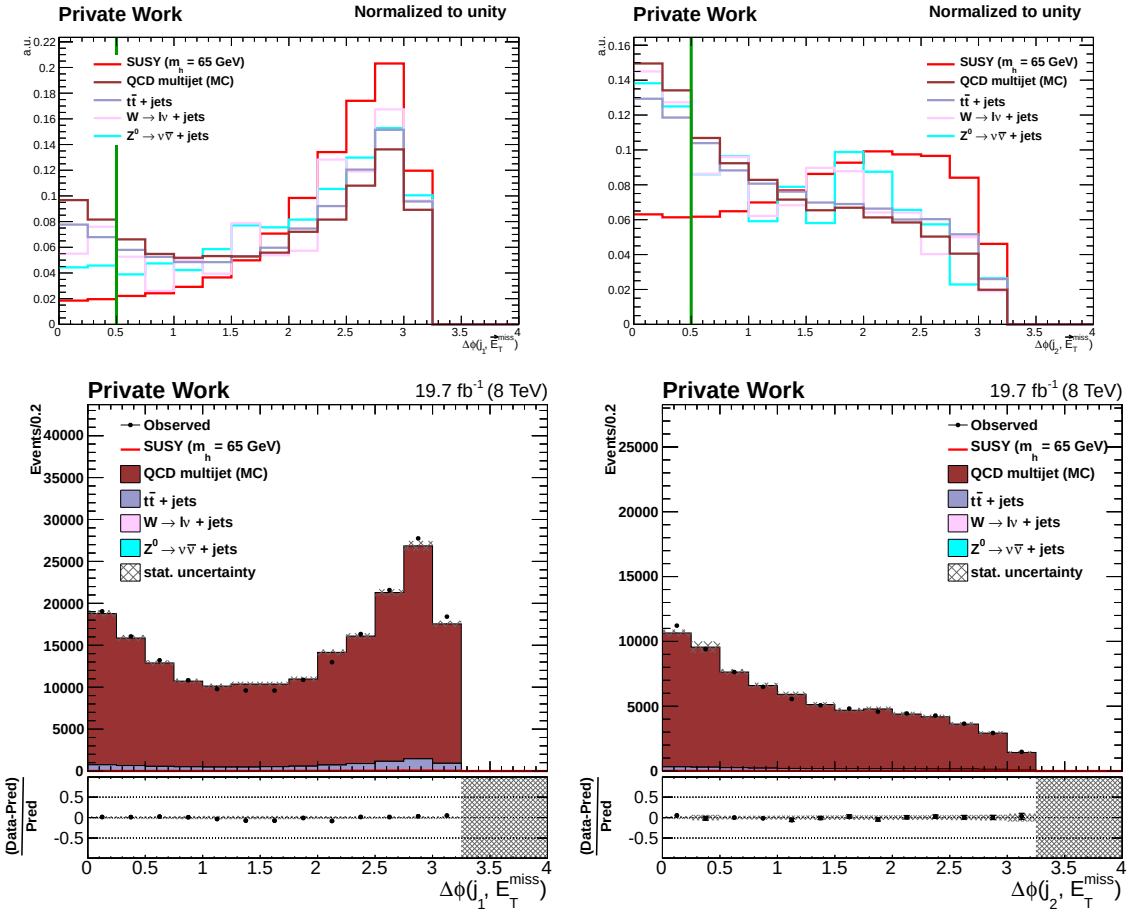

Figure D.3: The azimuthal angles between $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ and the leading (left) and subleading (right) jet are shown. The background shape is in agreement with the data distribution. QCD MC is used for illustration purposes only. The analysis uses a data-driven estimate of the QCD background.

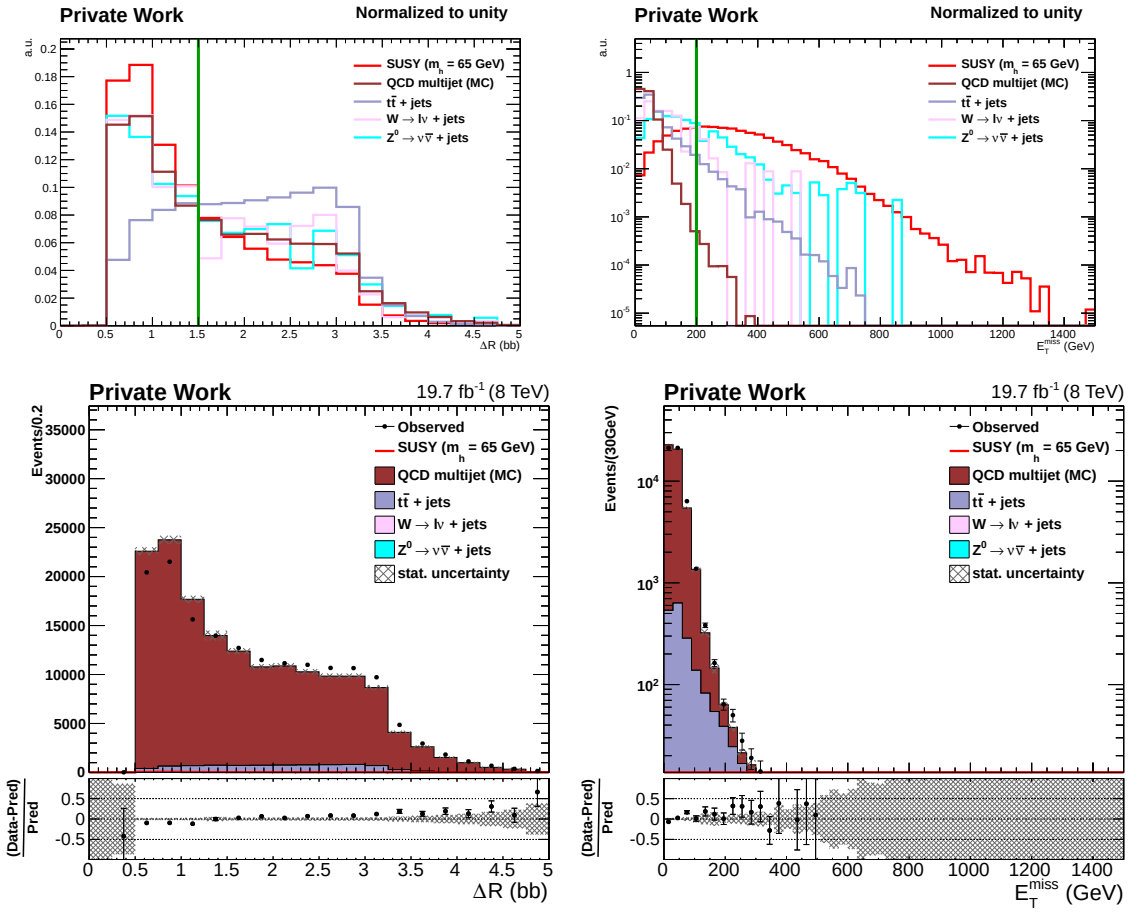

Figure D.4: The $\Delta R(\mathrm{b},\mathrm{b})$ (left) and $E_{\mathrm{T}}^{\mathrm{miss}}$ (right) distributions are shown. The background shape is in agreement with the data distribution. QCD MC is used for illustration purposes only. The analysis uses a data-driven estimate of the QCD background.

D.2 Optimisation of Selection Criteria

The offline event selection is optimised with respect to expected limits for three different h_1 mass hypotheses, 40, 65 and 80 GeV. The values of the discriminating variables are chosen such that the analysis is sensitive over the whole probed mass region. The detailed results of the procedure are shown in this section.

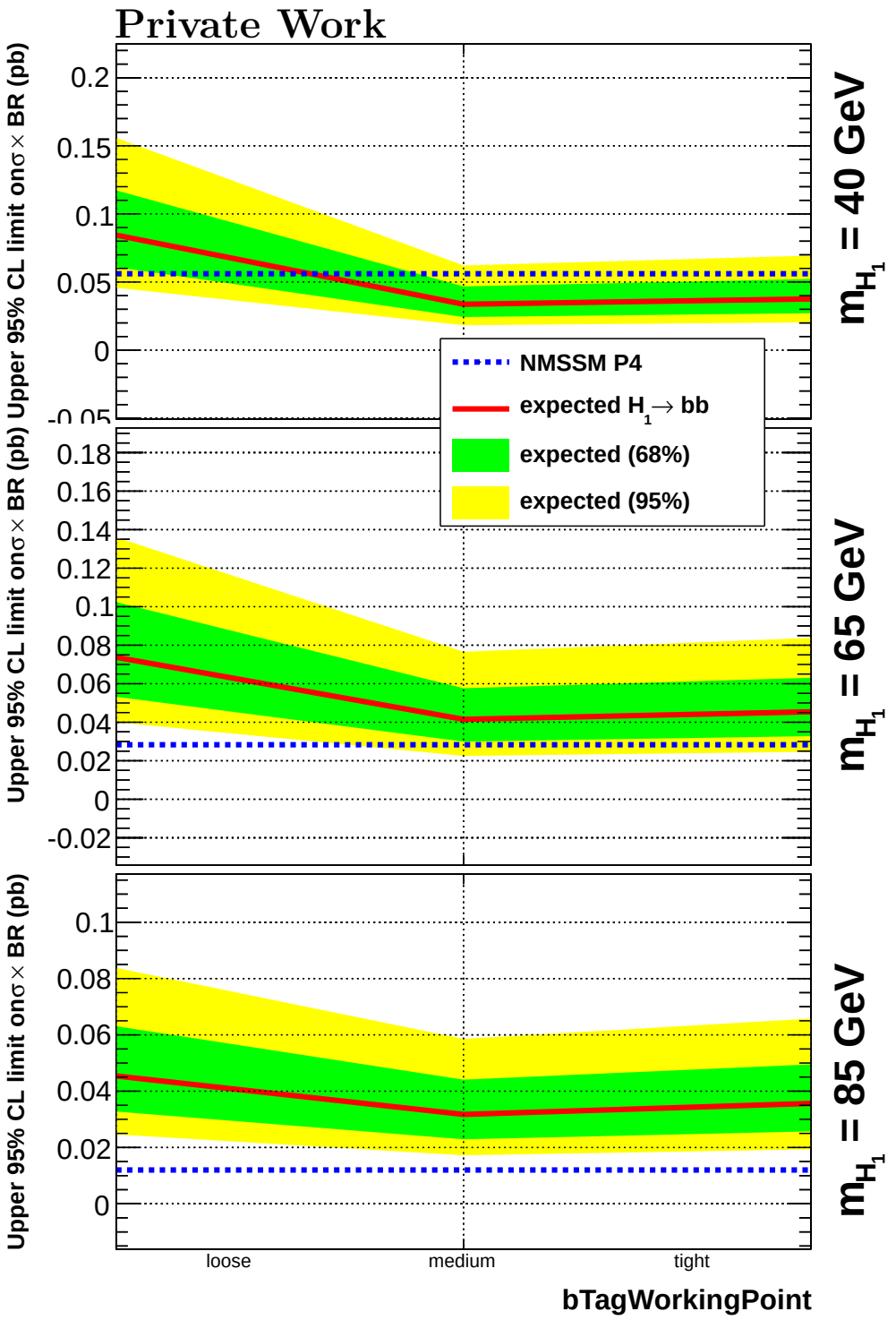

Figure D.5: Expected upper limits obtained for three different mass points using the loose, medium and tight working point of the CSV discriminator. The optimal choice for the selection is the medium working point.

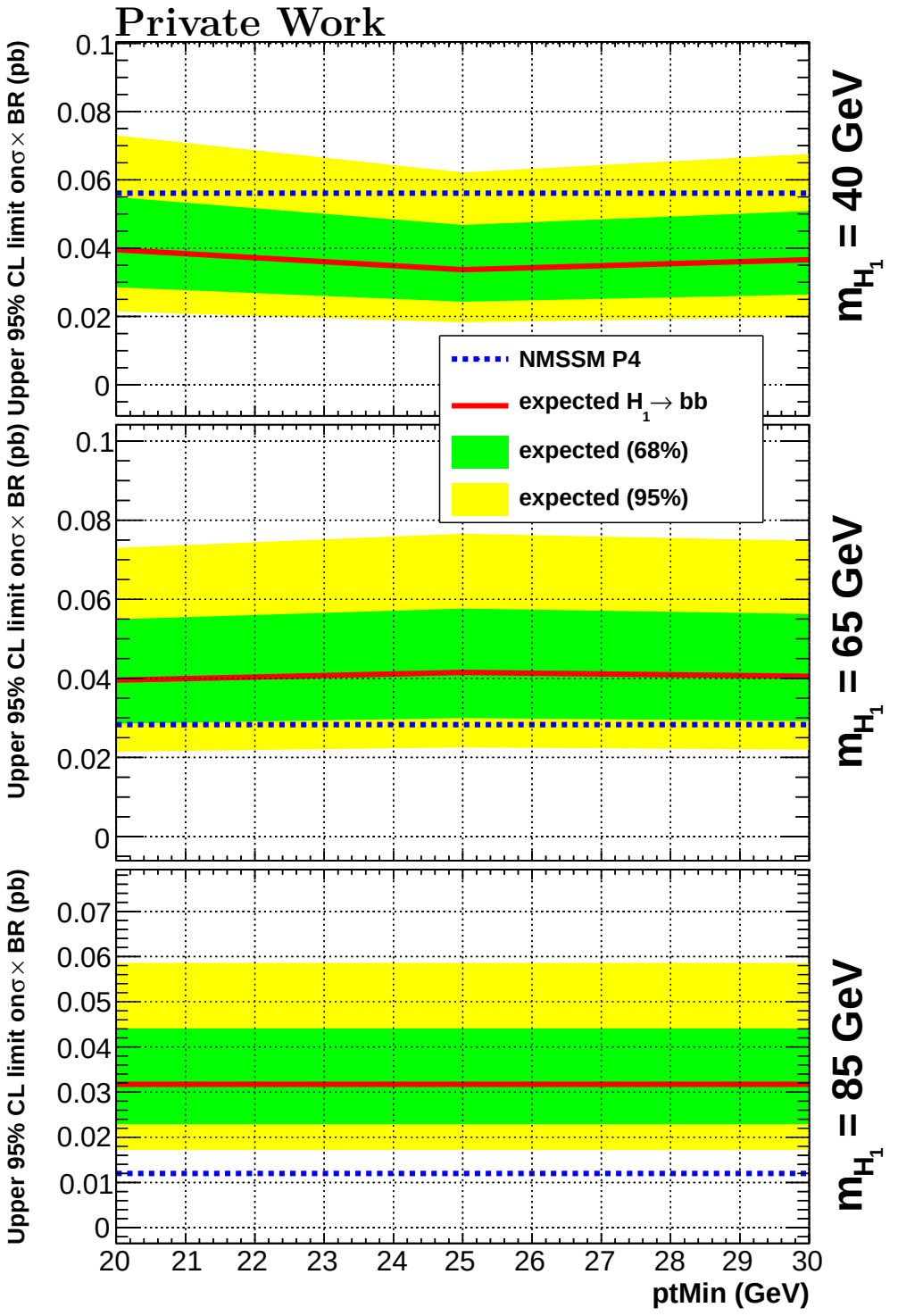

Figure D.6: Expected upper limits obtained for three different mass points using different values of the minimum jet-p_{T} criterion. The optimal choice for the selection is 25 GeV because the low mass point yields the best limit with this value and the other mass points show only slight dependence on this selection requirement.

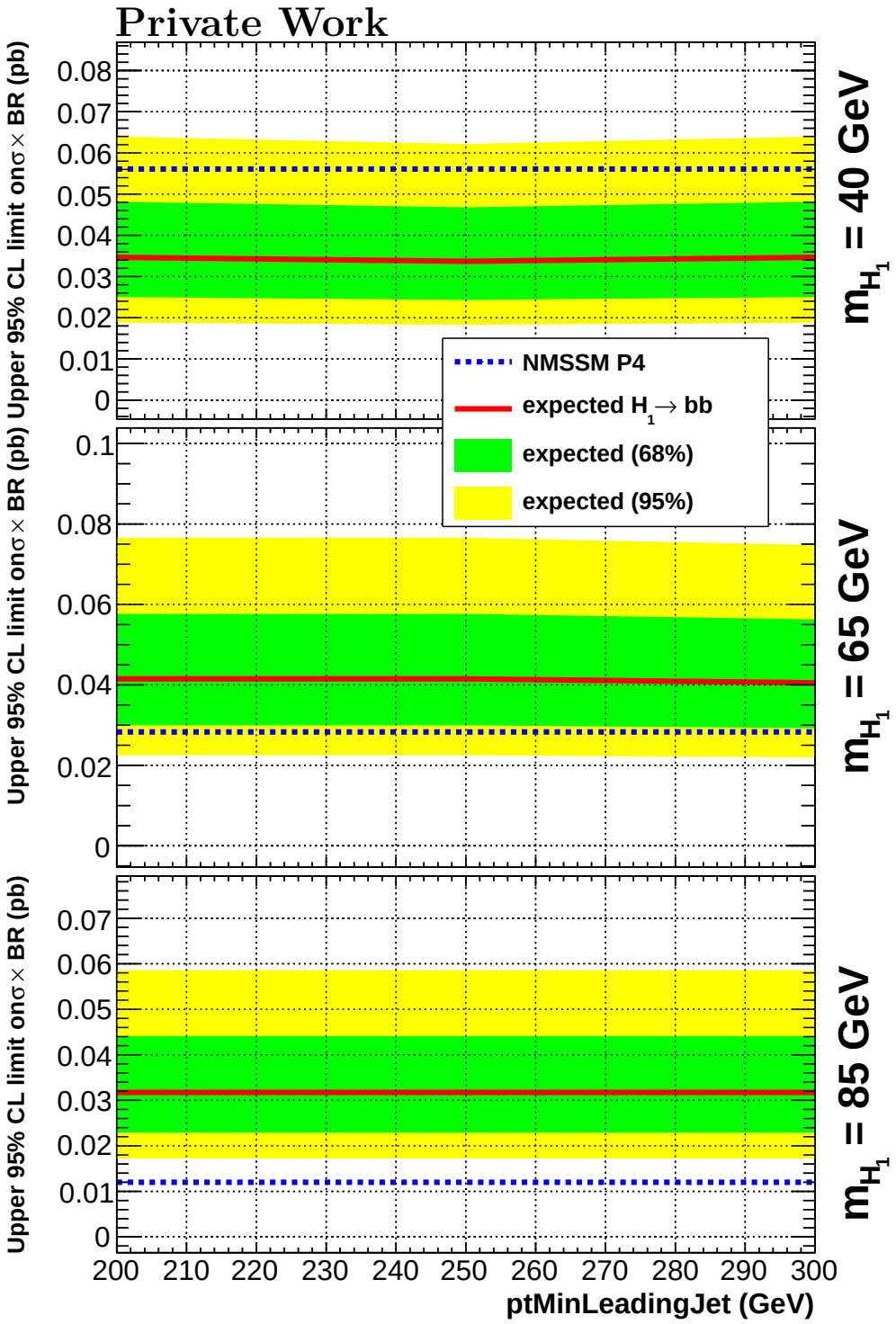

Figure D.7: Expected upper limits obtained for three different mass points using different values of the minimum p_T of the leading jet. The limits show little dependence on this selection requirement. The value 250 GeV, as used in [99], is chosen.

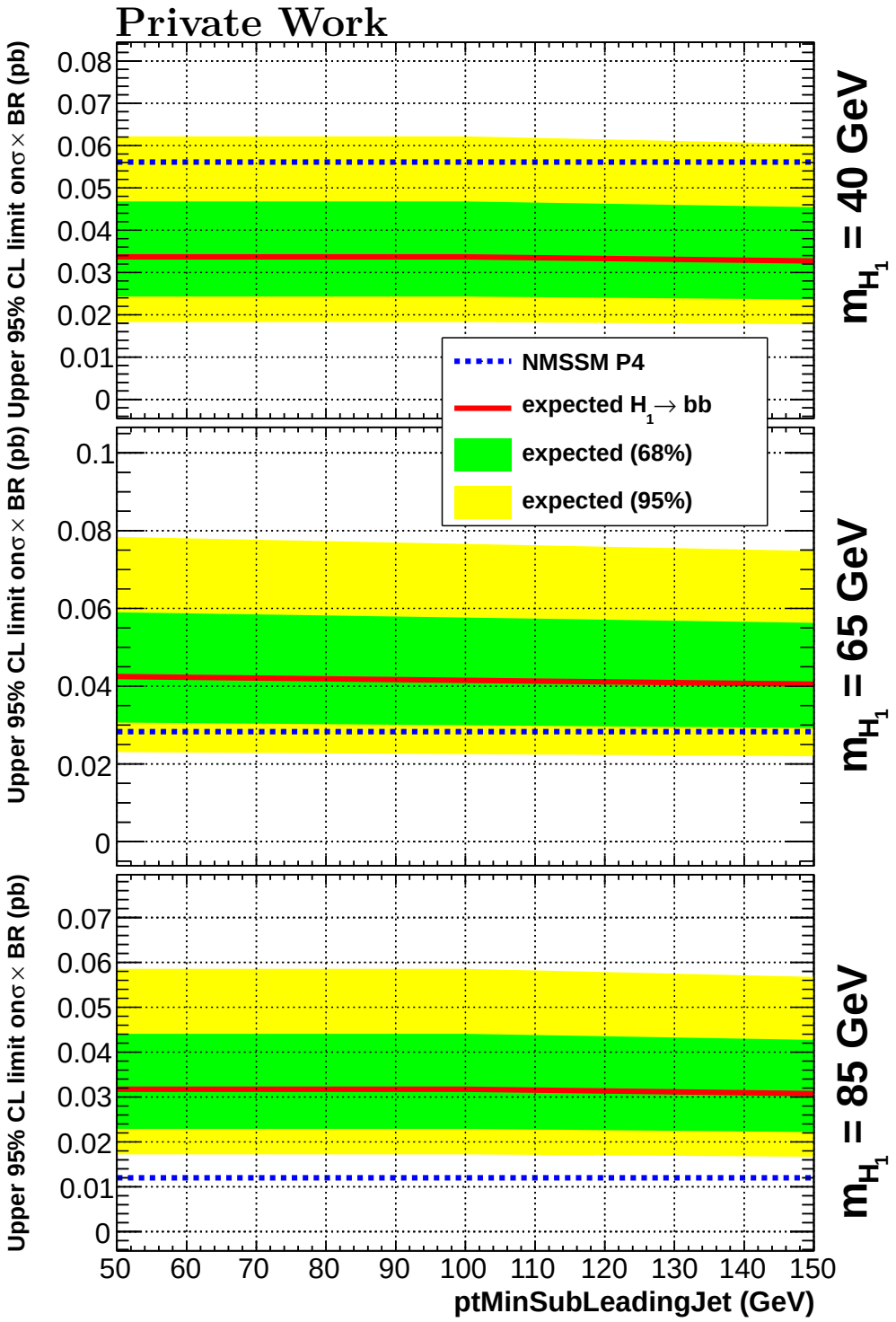

Figure D.8: Expected upper limits obtained for three different mass points using different values of the minimum p_{T} of the subleading jet. The limits show little dependence on this selection requirement. The value 100 GeV, as used in [99], is chosen.

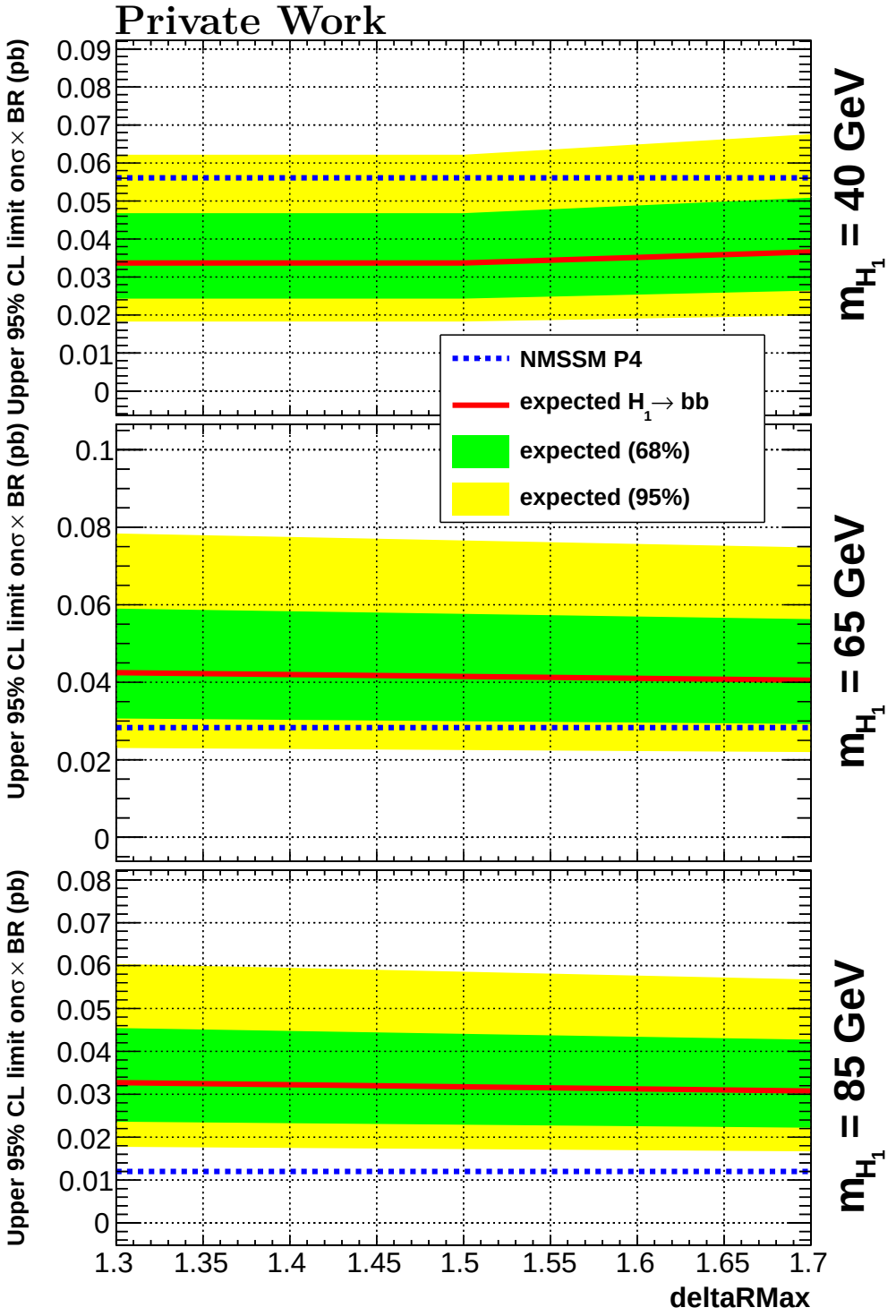

Figure D.9: Expected upper limits obtained for three different mass points using different values of the maximum separation between the two selected jets. A good compromise between low and high h_1 masses is obtained with $\Delta R^{\max}(b, b) = 1.5$.

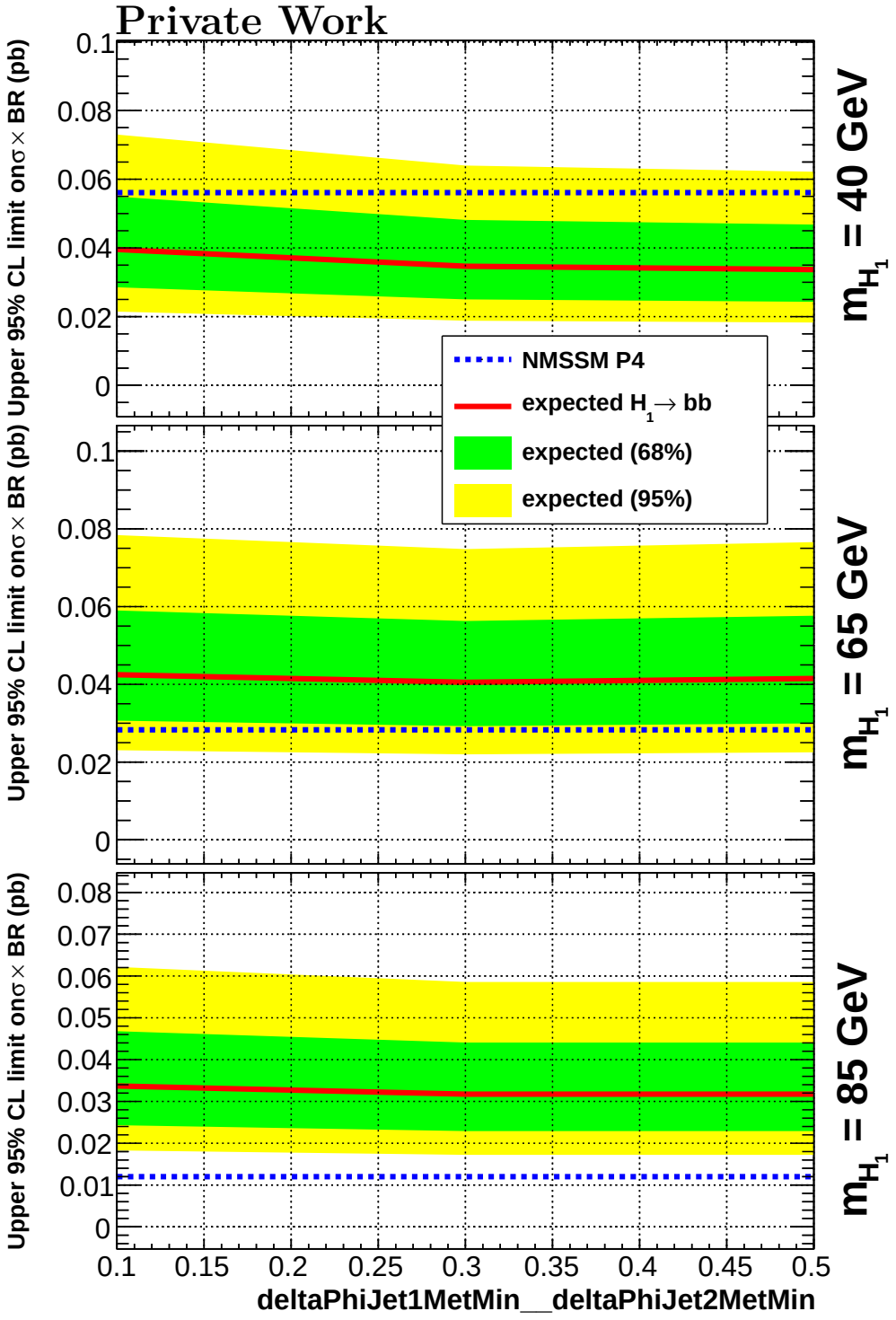

Figure D.10: Expected upper limit obtained for three different mass points using different values of the QCD suppression variables $\Delta\phi(j_{1/2}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})$. A plateau is reached at 0.5 for all mass points.

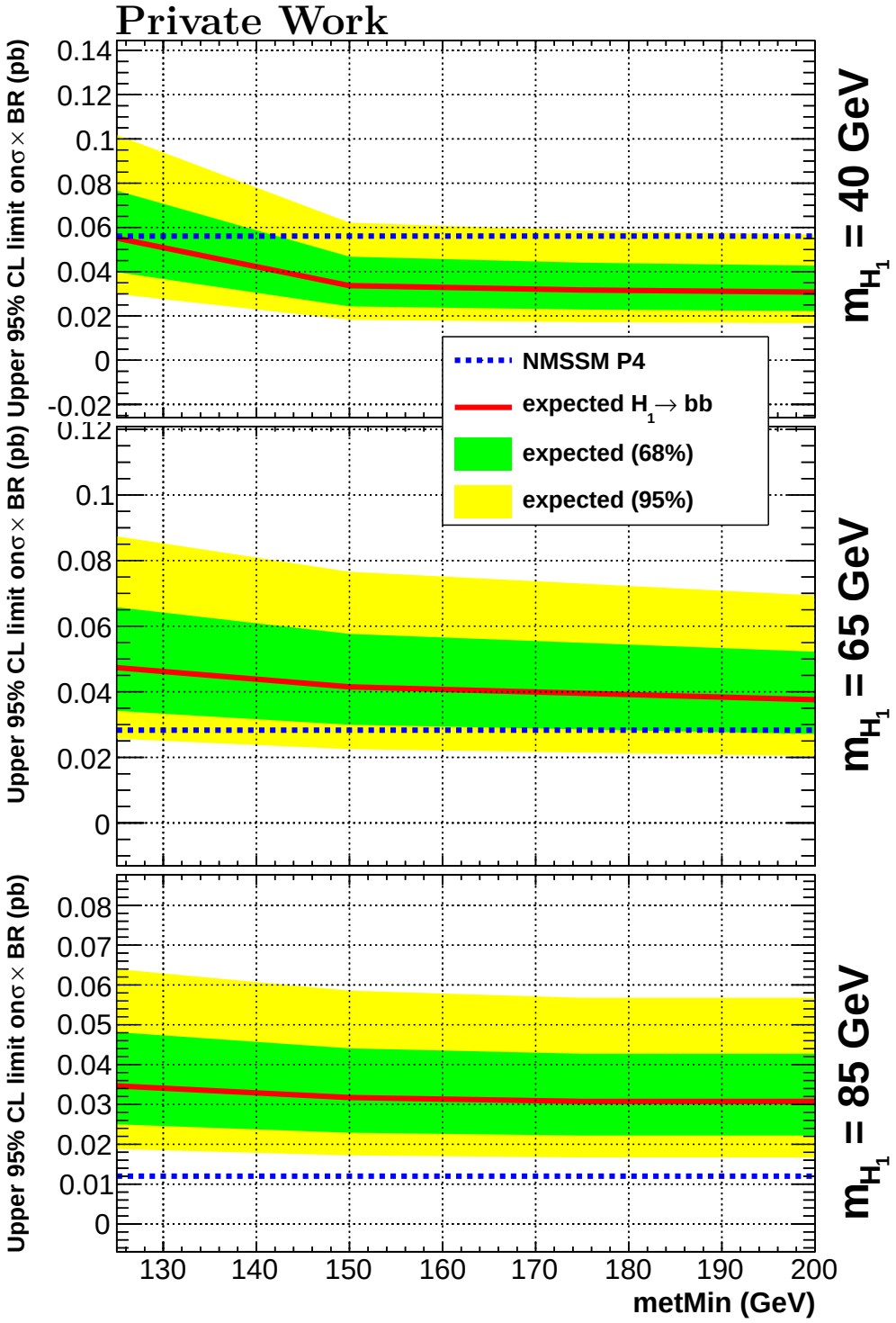

Figure D.11: Expected upper limit obtained for three different mass points using different values of the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement. A plateau is reached at 200 GeV for all mass points.

Appendix E

QCD Multijet Prediction

E.1 Sidebands of the ABCD-Like Method

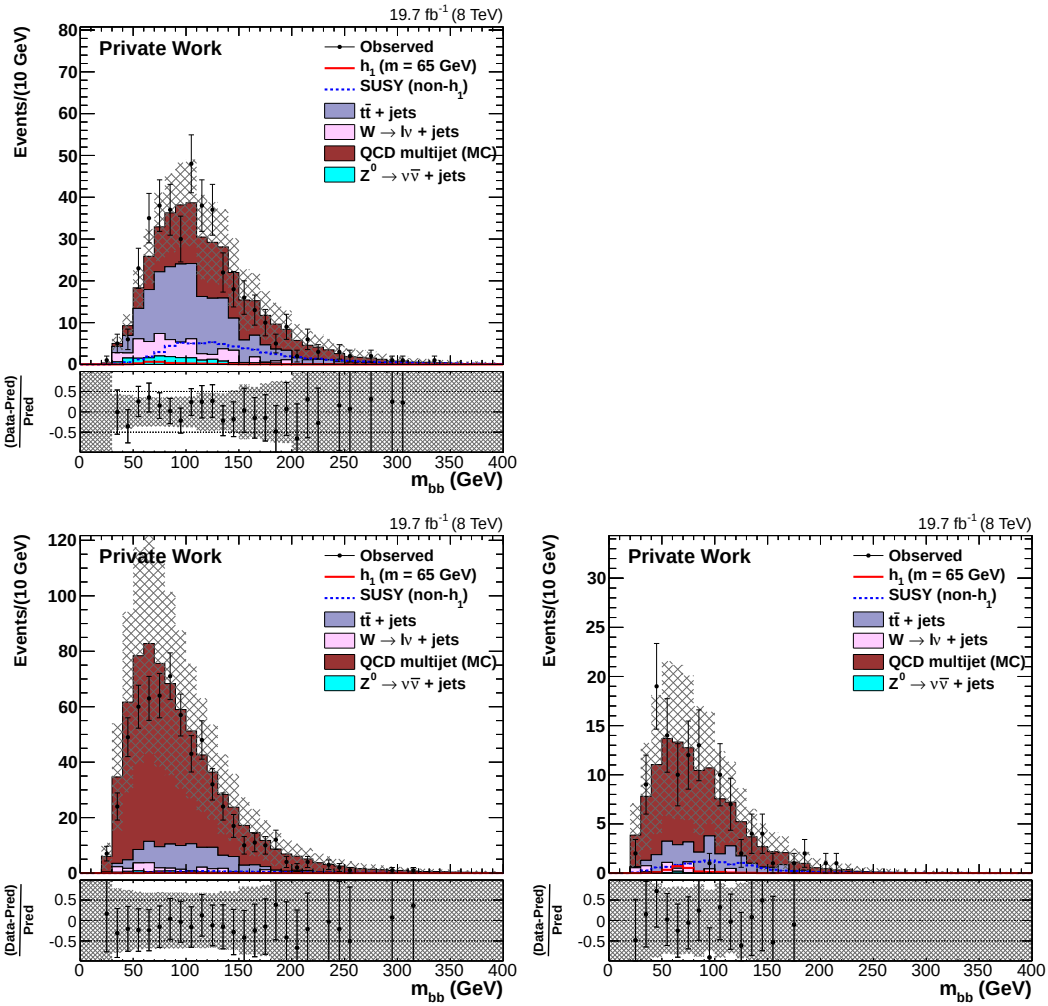

Figure E.1: Sidebands of the ABCD-like method described in Section 5.4.2. The figures are arranged according to the diagram in Fig. 5.14. The observed signal contamination (solid and dashed lines) is well contained in the uncertainty associated to the QCD-multijet normalisation. The QCD-multijet contribution is taken from simulation and scaled by the QCD-data-to-simulation ratio obtained in the signal region.

E.2 Parameterisation

E.2.1 Parameterisation Formulas

In this appendix the tested QCD parameterisation functions are listed.

- Crystal Ball function:

$$CB(m_{\mathrm{bb}}) = N \cdot \begin{cases} \exp\left\{\frac{m_{\mathrm{bb}}-m_0}{\sigma}\right\} , & \text{for } \frac{m_{\mathrm{bb}}-m_0}{\sigma} > -\alpha , \\ A \cdot \left(B - \frac{m_{\mathrm{bb}}-m_0}{\sigma}\right)^{-n} , & \text{for } \frac{m_{\mathrm{bb}}-m_0}{\sigma} \leq -\alpha , \end{cases} \tag{E.1}$$

$$\text{where } A = \left(\frac{n}{|\alpha|}\right) \cdot \exp\left\{-\frac{|\alpha|^2}{2}\right\} , \qquad B = \frac{n}{|\alpha|} - |\alpha| ,$$

$$C = \frac{n}{|\alpha|} \cdot \frac{1}{n-1} \cdot \exp\left\{-\frac{|\alpha|^2}{2}\right\} , \qquad D = \sqrt{\frac{\pi}{2}} \left(1 + \mathrm{erf}\left\{\frac{|\alpha|}{\sqrt{2}}\right\}\right) ,$$

$$N = \frac{1}{\sigma(C+D)} .$$

- Bernstein polynomials:

$$b_{\nu,n}(x) = \binom{n}{\nu} x^{\nu}(1-x)^{n-\nu} , \qquad \nu = 0, \ldots, n , \tag{E.2}$$

$$Bern_n(x) = \sum_{\nu=0}^{n} \beta_\nu b_{\nu,n}(x) , \tag{E.3}$$

$$\beta_\nu : \text{Bernstein coefficients} ,$$

$$x \in [0,1] .$$

- Novosibirsk function:

$$Novo(m_{\mathrm{bb}}) = N \cdot \exp\left\{-0.5 \cdot \left(\ln^2\left[1 + \Lambda \cdot \tau \cdot (m_{\mathrm{bb}} - m_0)\right]/\tau^2 + \tau^2\right)\right\} , \tag{E.4}$$

$$\text{where } \Lambda = \sinh(\tau\sqrt{\ln 4})/(\sigma\tau\sqrt{\ln 4}) ,$$

$$m_0 : \text{peak position} ,$$

$$\sigma : \text{width} ,$$

$$\tau : \text{tail parameter} .$$

E.2.2 Parameterisation in Control Regions

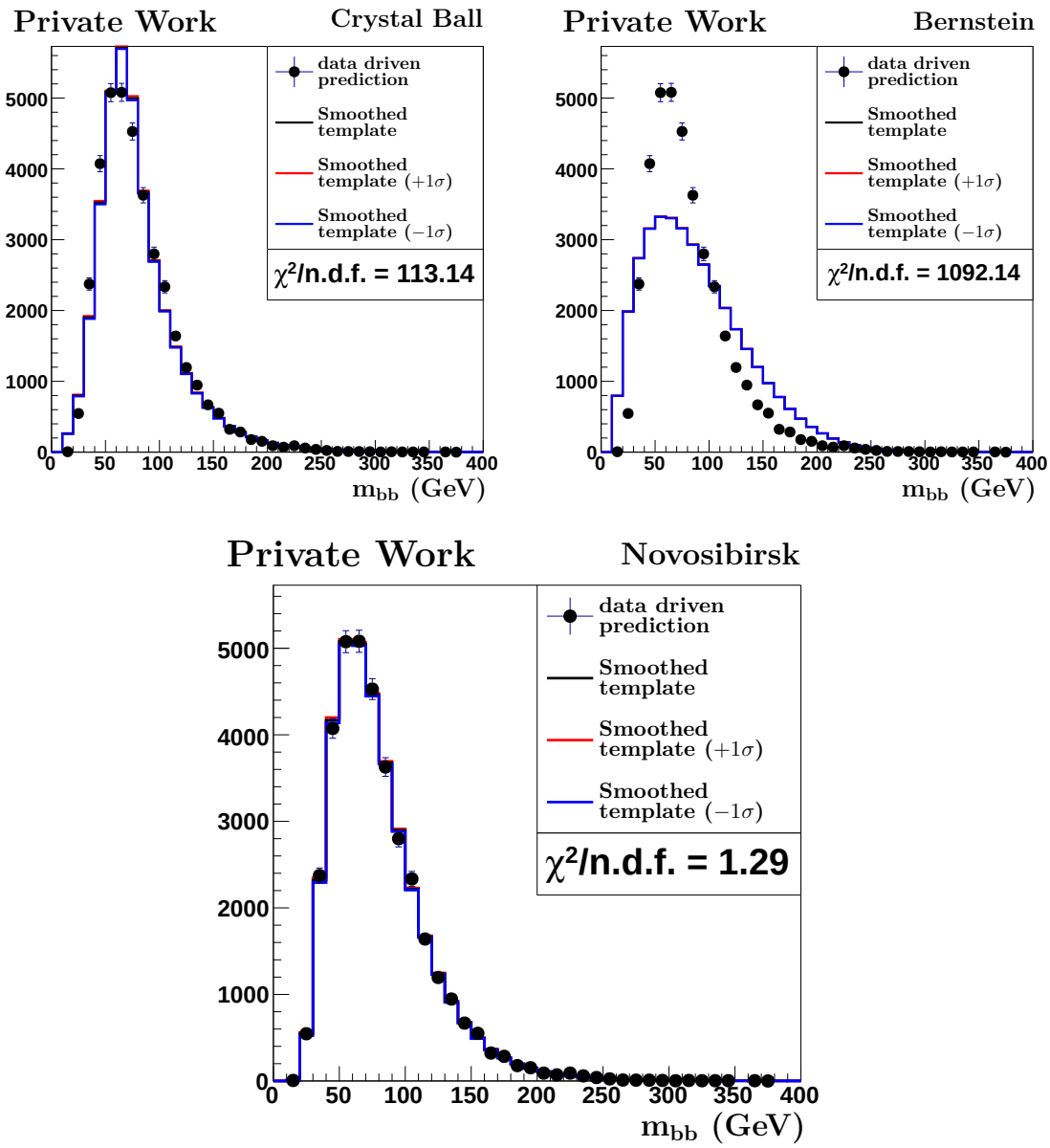

Figure E.2: Fit results of the QCD control region ($E_{\mathrm{T}}^{\mathrm{miss}} \in [0, 50]\,\mathrm{GeV}$, veto isolated leptons): Crystal Ball (top left), Bernstein (8th order, top right) and Novosibirsk (bottom). The best fit quality is achieved for the Novosibirsk function.

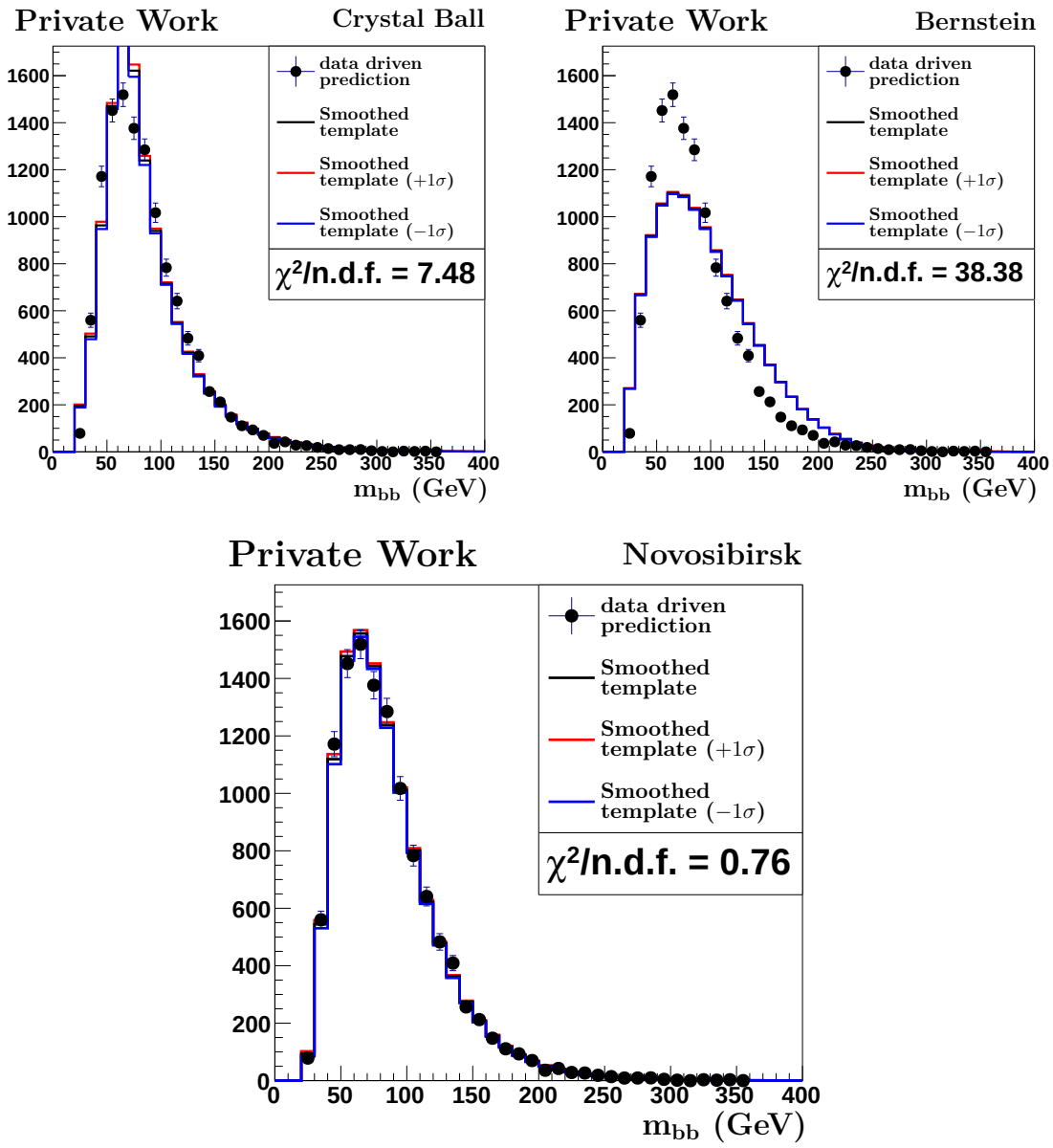

Figure E.3: Fit results of the QCD control region ($E_{\mathrm{T}}^{\mathrm{miss}} \in [50, 100]\,\mathrm{GeV}$, veto isolated leptons): Crystal Ball (top left), Bernstein (8th order, top right) and Novosibirsk (bottom). The best fit quality is achieved for the Novosibirsk function.

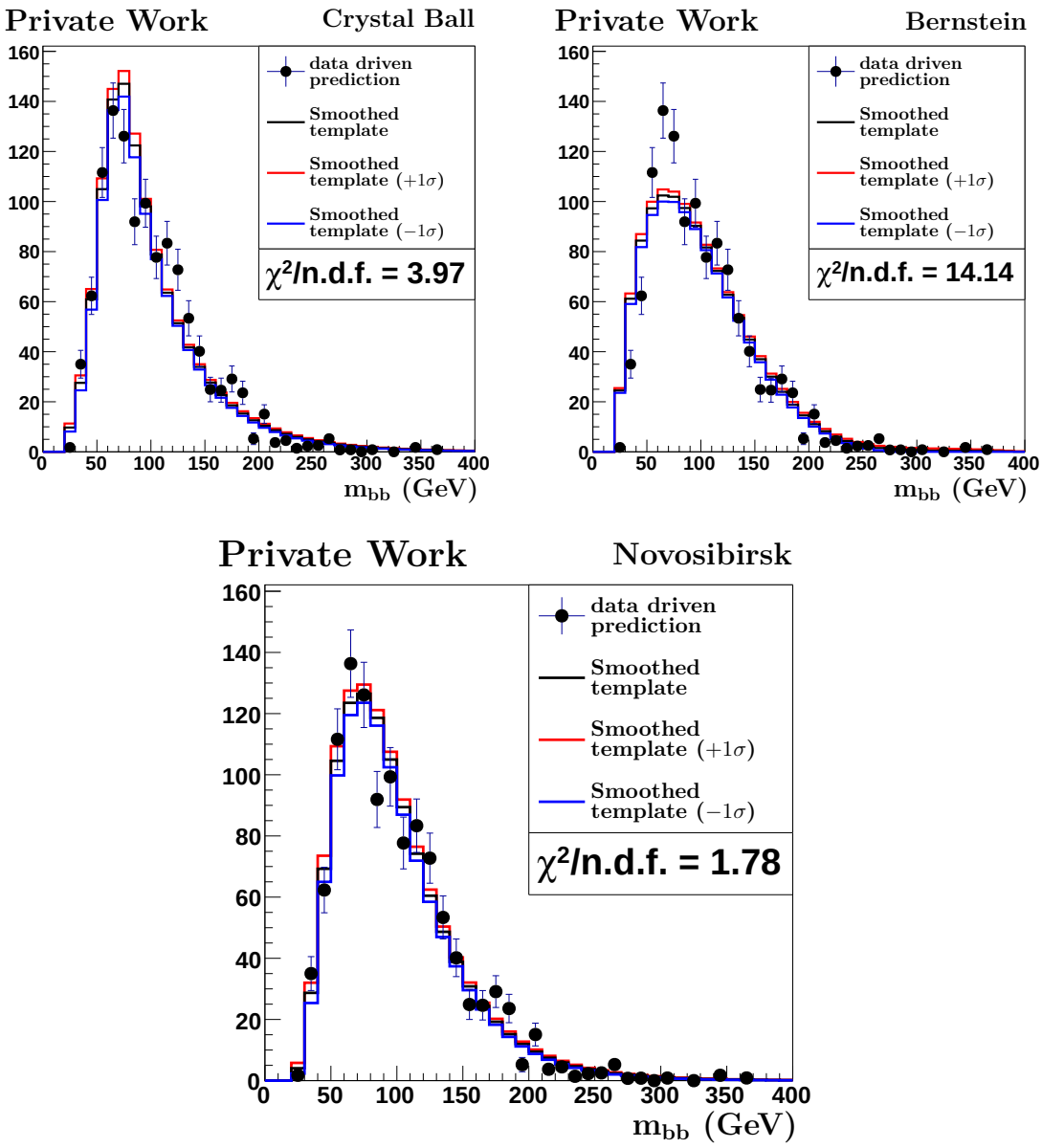

Figure E.4: Fit results of the QCD control region ($E_{\mathrm{T}}^{\mathrm{miss}} \in [100, 200]\,\mathrm{GeV}$, veto isolated leptons): Crystal Ball (top left), Bernstein (8th order, top right) and Novosibirsk (bottom). The best fit quality is achieved for the Novosibirsk function.

Appendix F

Expected Event Yields and Selection Efficiencies of Minor Backgrounds

Table F.1: Event yields of single top events after each selection step. The yields are normalised to the integrated luminosity of $19.7\,\mathrm{fb}^{-1}$ assuming the cross sections listed in Table 5.2. The efficiency of each selection step relative to the previous step is also listed. In the first line, "all" means the events which pass the trigger selection.

	events	efficiency
all	24946.90	1.000
$H_\mathrm{T} > 750\,\mathrm{GeV}$	11425.94	0.458
#non-b-tagged jets ≥ 2	11083.67	0.970
#jets ≥ 4	10015.06	0.904
#b-tagged jets ≥ 2	3307.67	0.330
b veto for leading two jets	728.53	0.220
$p_\mathrm{T}^\mathrm{Lead} > 250\,\mathrm{GeV}$, $p_\mathrm{T}^\mathrm{SubLead} > 100\,\mathrm{GeV}$	541.00	0.743
$\Delta\phi(j_1, \vec{E}_\mathrm{T}^\mathrm{miss}) > 0.5$	459.32	0.849
$\min \Delta R(\mathrm{b}_1, \mathrm{b}_2) < 1.5$	138.20	0.301
$\Delta\phi(j_2, \vec{E}_\mathrm{T}^\mathrm{miss}) > 0.5$	97.75	0.707
$E_\mathrm{T}^\mathrm{miss} > 200\,\mathrm{GeV}$	1.63	0.017
total		0.000066

Table F.2: Event yields of WZ^0 events after each selection step. The yields are normalised to the integrated luminosity of $19.7\,\mathrm{fb}^{-1}$ assuming the cross sections listed in Table 5.2. The efficiency of each selection step relative to the previous step is also listed. In the first line, "all" means the events which pass the trigger selection.

	events	efficiency
all	2285.12	1.000
$H_\mathrm{T} > 750\,\mathrm{GeV}$	1019.17	0.446
#non-b-tagged jets ≥ 2	989.14	0.971
#jets ≥ 4	676.11	0.684
#b-tagged jets ≥ 2	52.35	0.077
b veto for leading two jets	19.38	0.370
$p_\mathrm{T}^\mathrm{Lead} > 250\,\mathrm{GeV}$, $p_\mathrm{T}^\mathrm{SubLead} > 100\,\mathrm{GeV}$	14.56	0.751
$\Delta\phi(j_1, \vec{E}_\mathrm{T}^\mathrm{miss}) > 0.5$	12.48	0.857
$\min \Delta R(\mathrm{b}_1, \mathrm{b}_2) < 1.5$	6.78	0.543
$\Delta\phi(j_2, \vec{E}_\mathrm{T}^\mathrm{miss}) > 0.5$	4.88	0.720
$E_\mathrm{T}^\mathrm{miss} > 200\,\mathrm{GeV}$	0.00	0.000
total		0.000000

Table F.3: Event yields of Z^0Z^0 events after each selection step. The yields are normalised to the integrated luminosity of $19.7\,\mathrm{fb}^{-1}$ assuming the cross sections listed in Table 5.2. The efficiency of each selection step relative to the previous step is also listed. In the first line, "all" means the events which pass the trigger selection.

	events	efficiency
all	520.95	1.000
$H_\mathrm{T} > 750\,\mathrm{GeV}$	226.92	0.436
#non-b-tagged jets ≥ 2	216.37	0.954
#jets ≥ 4	155.08	0.717
#b-tagged jets ≥ 2	20.41	0.132
b veto for leading two jets	6.66	0.326
$p_\mathrm{T}^\mathrm{Lead} > 250\,\mathrm{GeV}$, $p_\mathrm{T}^\mathrm{SubLead} > 100\,\mathrm{GeV}$	4.79	0.720
$\Delta\phi(j_1, \vec{E}_\mathrm{T}^\mathrm{miss}) > 0.5$	3.95	0.824
$\min \Delta R(\mathrm{b}_1, \mathrm{b}_2) < 1.5$	2.07	0.525
$\Delta\phi(j_2, \vec{E}_\mathrm{T}^\mathrm{miss}) > 0.5$	1.50	0.724
$E_\mathrm{T}^\mathrm{miss} > 200\,\mathrm{GeV}$	0.07	0.043
total		0.000125

Table F.4: Conservative upper limits on event yields for $Z^0 \to b\bar{b}$ events after each selection step. The yields are normalised to the integrated luminosity of $19.7\,\mathrm{fb}^{-1}$ assuming the cross sections listed in Table 5.2 and the branching fraction correction described in Section 5.4.3. The efficiency of each selection step relative to the previous step is also listed. In the first line, "all" means the events which pass the trigger selection.

	events	efficiency
all	41342.11	1.000
$H_\mathrm{T} > 750\,\mathrm{GeV}$	10554.07	0.255
#non-b-tagged jets ≥ 2	10387.52	0.984
b veto for leading two jets	2375.48	0.229
#jets ≥ 4	2375.48	1.000
#b-tagged jets ≥ 2	2211.04	0.931
$p_\mathrm{T}^\mathrm{Lead} > 250\,\mathrm{GeV}$, $p_\mathrm{T}^\mathrm{SubLead} > 100\,\mathrm{GeV}$	2077.47	0.940
$\Delta\phi(j_1, \vec{E}_\mathrm{T}^\mathrm{miss}) > 0.5$	1653.26	0.796
$\min \Delta R(\mathrm{b}_1, \mathrm{b}_2) < 1.5$	894.97	0.541
$\Delta\phi(j_2, \vec{E}_\mathrm{T}^\mathrm{miss}) > 0.5$	586.78	0.656
$E_\mathrm{T}^\mathrm{miss} > 200\,\mathrm{GeV}$	3.85	0.007
total		0.000093

Appendix G

Extended Results of the Light Higgs Boson Search

G.1 Pre-Fit Background Prediction

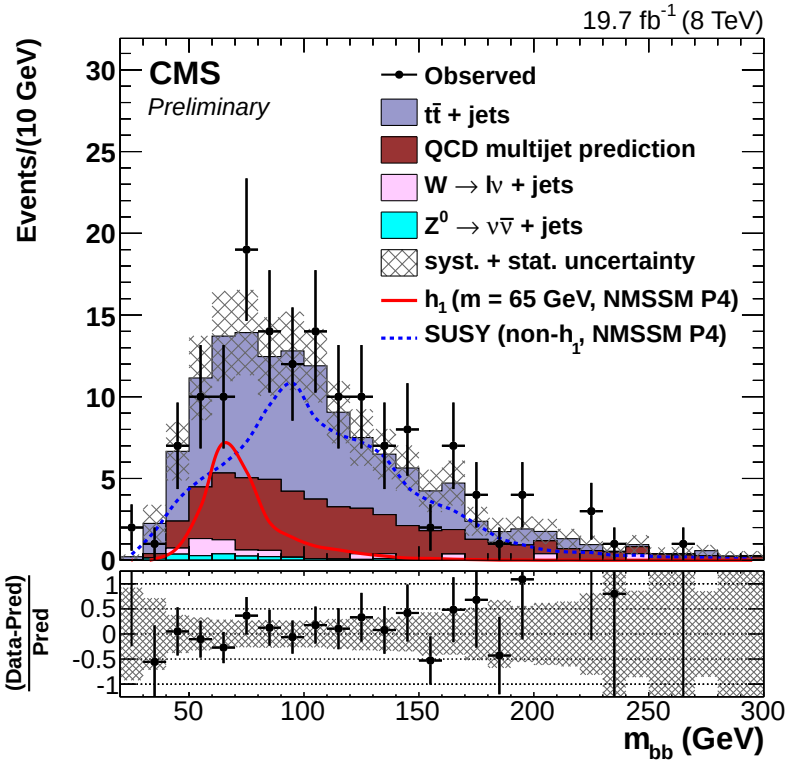

Figure G.1: Pre-fit background prediction of the invariant mass of the selected b-tagged jets, where the SM background contributions are stacked. The h_1 and non-h_1 contributions are overlaid to illustrate the sensitivity of the analysis to the modified P4 scenario.

G.2 Cross Section Measurement

Table G.1: Fitted signal strength modifiers of the search for a resonant structure in the probed h_1 mass hypotheses. A maximum upward deviation of 0.49σ with respect to the SM background is observed at $m_{h_1} = 75\,\text{GeV}$.

m_{h_1} (GeV)	μ	$\mu/\sigma(\mu)$
30	$0.18^{+0.26}_{-0.19}$	0.81
35	$-0.02^{+0.24}_{-0.21}$	-0.07
40	$-0.04^{+0.21}_{-0.23}$	-0.19
45	$-0.08^{+0.23}_{-0.26}$	-0.31
50	$-0.16^{+0.29}_{-0.32}$	-0.53
55	$-0.23^{+0.34}_{-0.37}$	-0.65
60	$-0.22^{+0.44}_{-0.46}$	-0.49
65	$-0.19^{+0.52}_{-0.52}$	-0.36
70	$0.10^{+0.58}_{-0.58}$	0.17
75	$0.35^{+0.74}_{-0.70}$	0.49
80	$0.38^{+0.86}_{-0.82}$	0.45
85	$0.31^{+0.89}_{-0.86}$	0.36
90	$0.19^{+1.32}_{-1.28}$	0.15
95	$0.32^{+1.41}_{-1.38}$	0.23
100	$0.62^{+1.68}_{-1.60}$	0.38

G.3 NMSSM Cross Section in the Modified P4 Scenario

An NMSSM cross-section measurement in the modified P4 scenario is performed with the full NMSSM spectrum as signal. No significant excess is observed (see Table G.2, Fig.G.2).

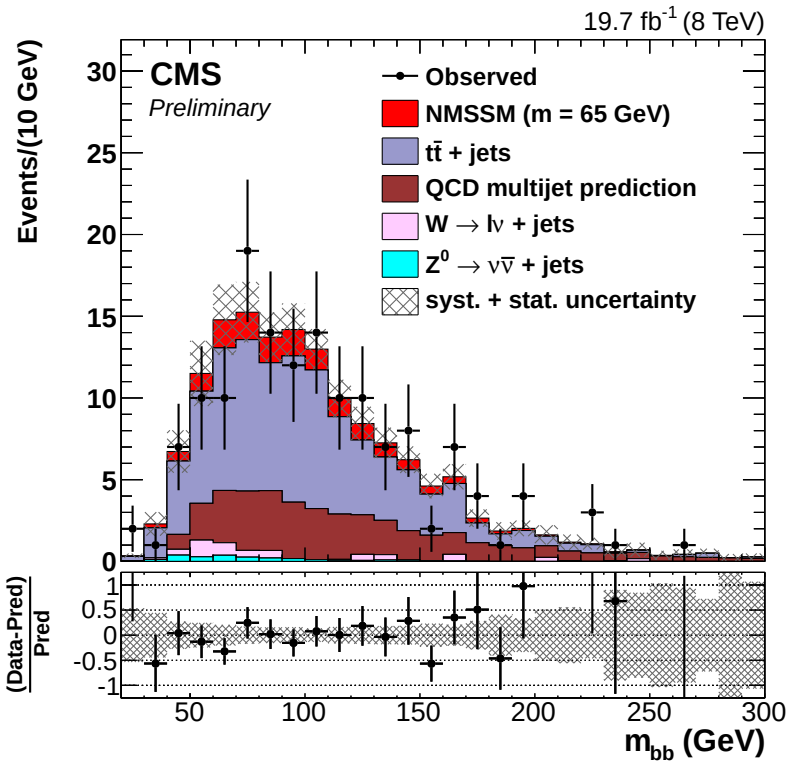

Figure G.2: Invariant mass of the selected b-tagged jets, where the SM background contributions are stacked. The NMSSM P4 signal template with a h_1 mass hypothesis of 65 GeV and the SM background are fitted to the data.

Table G.2: Fitted signal strength modifiers of the full NMSSM P4 spectrum for different probed h_1 mass hypotheses. A maximum deviation of 0.99σ with respect to the core of the SM background is observed at $m_{h_1} = 75\,\text{GeV}$.

m_{h_1} (GeV)	μ	$\mu/\sigma(\mu)$
30	$0.15^{+0.15}_{-0.13}$	1.03
35	$0.09^{+0.15}_{-0.15}$	0.60
40	$0.05^{+0.14}_{-0.14}$	0.33
45	$0.03^{+0.15}_{-0.15}$	0.22
50	$0.04^{+0.15}_{-0.16}$	0.28
55	$0.05^{+0.16}_{-0.17}$	0.30
60	$0.08^{+0.16}_{-0.17}$	0.47
65	$0.13^{+0.21}_{-0.21}$	0.63
70	$0.15^{+0.22}_{-0.21}$	0.70
75	$0.22^{+0.22}_{-0.22}$	0.99
80	$0.19^{+0.23}_{-0.23}$	0.83
85	$0.21^{+0.21}_{-0.21}$	0.99
90	$0.15^{+0.17}_{-0.18}$	0.85
95	$0.13^{+0.17}_{-0.17}$	0.75
100	$0.10^{+0.17}_{-0.17}$	0.61

Bibliography

[1] S. L. Glashow, Partial-symmetries of weak interactions, *Nucl.Phys.* **22**, 579 (1961).

[2] S. Weinberg, A Model of Leptons, *Phys.Rev.Lett.* **19**, 1264 (1967).

[3] A. Salam, Weak and electromagnetic interactions, in *Elementary particle theory*, edited by N. Svartholm, pp. 367–377, Almquist & Wiksell, 1968.

[4] Particle Data Group, K. A. Olive *et al.*, Review of Particle Physics (RPP), *Chin.Phys.* **C38**, 090001 (2014).

[5] ATLAS, G. Aad *et al.*, Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC, *Phys.Lett.* **B716**, 1 (2012), 1207.7214.

[6] CMS, S. Chatrchyan *et al.*, Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC, *Phys.Lett.* **B716**, 30 (2012), 1207.7235.

[7] CMS, V. Khachatryan *et al.*, Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV, *Eur. Phys. J.* **C75**, 212 (2015), 1412.8662.

[8] ATLAS, G. Aad *et al.*, Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC, *Phys.Lett.* **B726**, 88 (2013), 1307.1427.

[9] ATLAS, G. Aad *et al.*, Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels with the ATLAS detector using 25 fb^{-1} of pp collision data, *Phys.Rev.* **D90**, 052004 (2014), 1406.3827.

[10] L. M. J. S. Volders, Neutral hydrogen in M 33 and M 101, *Bulletin of the Astronomical Institutes of the Netherlands* **14**, 323 (1959).

[11] A. Refregier, Weak gravitational lensing by large scale structure, *Ann.Rev.Astron.Astrophys.* **41**, 645 (2003), astro-ph/0307212.

[12] A. Vikhlinin *et al.*, Chandra sample of nearby relaxed galaxy clusters: Mass, gas fraction, and mass-temperature relation, *Astrophys.J.* **640**, 691 (2006), astro-ph/0507092.

[13] D. Clowe *et al.*, A direct empirical proof of the existence of dark matter, *Astrophys.J.* **648**, L109 (2006), astro-ph/0608407.

[14] N. Jarosik *et al.*, Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Sky Maps, Systematic Errors, and Basic Results, *Astrophys.J.Suppl.* **192**, 14 (2011), 1001.4744.

[15] WMAP Collaboration, E. Komatsu *et al.*, Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation, *Astrophys.J.Suppl.* **192**, 18 (2011), 1001.4538.

[16] S. Weinberg, Implications of dynamical symmetry breaking, *Phys.Rev.* **D13**, 974 (1976).

[17] S. Weinberg, Implications of dynamical symmetry breaking: An addendum, *Phys.Rev.* **D19**, 1277 (1979).

[18] E. Gildener, Gauge-symmetry hierarchies, *Phys.Rev.* **D14**, 1667 (1976).

[19] L. Susskind, Dynamics of spontaneous symmetry breaking in the Weinberg-Salam theory, *Phys.Rev.* **D20**, 2619 (1979).

[20] P. Van Nieuwenhuizen, Supergravity, *Phys.Rept.* **68**, 189 (1981).

[21] L. Evans and P. Bryant, LHC Machine, *JINST* **3**, S08001 (2008).

[22] CMS Collaboration, S. Chatrchyan *et al.*, The CMS experiment at the CERN LHC, *JINST* **3**, S08004 (2008).

[23] P. Fayet, Supergauge Invariant Extension of the Higgs Mechanism and a Model for the electron and Its Neutrino, *Nucl.Phys.* **B90**, 104 (1975).

[24] M. Dine, W. Fischler, and M. Srednicki, A Simple Solution to the Strong CP Problem with a Harmless Axion, *Phys.Lett.* **B104**, 199 (1981).

[25] H. P. Nilles, M. Srednicki, and D. Wyler, Weak Interaction Breakdown Induced by Supergravity, *Phys.Lett.* **B120**, 346 (1983).

[26] J. Frere, D. Jones, and S. Raby, Fermion Masses and Induction of the Weak Scale by Supergravity, *Nucl.Phys.* **B222**, 11 (1983).

[27] J. Derendinger and C. A. Savoy, Quantum Effects and SU(2) x U(1) Breaking in Supergravity Gauge Theories, *Nucl.Phys.* **B237**, 307 (1984).

[28] Muon G-2 Collaboration, G. Bennett *et al.*, Final Report of the Muon E821 Anomalous Magnetic Moment Measurement at BNL, *Phys.Rev.* **D73**, 072003 (2006), hep-ex/0602035.

[29] H. P. Nilles, Supersymmetry, Supergravity and Particle Physics, *Phys.Rept.* **110**, 1 (1984).

[30] H. E. Haber and G. L. Kane, The Search for Supersymmetry: Probing Physics Beyond the Standard Model, *Phys.Rept.* **117**, 75 (1985).

[31] S. Tomonaga, On a relativistically invariant formulation of the quantum theory of wave fields, *Prog.Theor.Phys.* **1**, 27 (1946).

[32] J. S. Schwinger, On Quantum electrodynamics and the magnetic moment of the electron, *Phys.Rev.* **73**, 416 (1948).

[33] J. S. Schwinger, Quantum Electrodynamics. I. A Covariant Formulation, *Phys.Rev.* **74**, 1439 (1948).

[34] F. Dyson, The Radiation theories of Tomonaga, Schwinger, and Feynman, *Phys.Rev.* **75**, 486 (1949).

[35] F. Dyson, The S matrix in quantum electrodynamics, *Phys.Rev.* **75**, 1736 (1949).

[36] R. Feynman, The Theory of positrons, *Phys.Rev.* **76**, 749 (1949).

[37] R. Feynman, Space - time approach to quantum electrodynamics, *Phys.Rev.* **76**, 769 (1949).

[38] R. Feynman, Mathematical formulation of the quantum theory of electromagnetic interaction, *Phys.Rev.* **80**, 440 (1950).

[39] C. Wu, E. Ambler, R. Hayward, D. Hoppes, and R. Hudson, Experimental Test of Parity Conservation in Beta Decay, *Phys.Rev.* **105**, 1413 (1957).

[40] J. Christenson, J. Cronin, V. Fitch, and R. Turlay, Evidence for the 2 pi Decay of the k(2)0 Meson, *Phys.Rev.Lett.* **13**, 138 (1964).

[41] J. Pumplin *et al.*, New generation of parton distributions with uncertainties from global QCD analysis, *JHEP* **0207**, 012 (2002), hep-ph/0201195.

[42] D. J. Gross and F. Wilczek, Ultraviolet Behavior of Nonabelian Gauge Theories, *Phys.Rev.Lett.* **30**, 1343 (1973).

[43] H. D. Politzer, Reliable Perturbative Results for Strong Interactions?, *Phys.Rev.Lett.* **30**, 1346 (1973).

[44] H. Georgi, Lie Algebras In Particle Physics. From Isospin To Unified Theories, *Front.Phys.* **54**, 1 (1982).

[45] M. Böhm, A. Denner, and H. Joos, *Gauge Theories of the Strong and Electroweak Interaction*, 3. ed. (Teubner, Stuttgart, 2001).

[46] C. N. Yang and R. Mills, Conservation of Isotopic Spin and Isotopic Gauge Invariance, *Phys.Rev.* **96**, 191 (1954).

[47] A. Einstein, Die Grundlage der allgemeinen Relativitätstheorie, *Annalen der Physik* **354**, 769 (1916).

[48] D. Kaiser, Physics and Feynman's Diagrams, *American Scientist* **93**, 156 (2005).

[49] J. C. Collins, *Renormalization: an introduction to renormalization, the renormalization group, and the operator-product expansion* Cambridge monographs on mathematical physics (Cambridge Univ. Press, Cambridge, 1984).

[50] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory* (Westview Press, 1995).

[51] UA1, G. Arnison *et al.*, Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ GeV, *Phys.Lett.* **B122**, 103 (1983).

[52] UA1, G. Arnison *et al.*, Experimental observation of lepton pairs of invariant mass around 95 GeV/c^2 at the CERN SPS collider, *Phys.Lett.* **B126**, 398 (1983).

[53] UA1, G. Arnison *et al.*, Observation of Muonic Z0 Decay at the anti-p p Collider, *Phys.Lett.* **B147**, 241 (1984).

[54] P. W. Higgs, Broken symmetries and the masses of gauge bosons, *Phys.Rev.Lett.* **13**, 508 (1964).

[55] F. Englert and R. Brout, Broken symmetry and the mass of gauge vector mesons, *Phys.Rev.Lett.* **13**, 321 (1964).

[56] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, Global conservation laws and massless particles, *Phys.Rev.Lett.* **13**, 585 (1964).

[57] J. Goldstone, Field Theories with Superconductor Solutions, *Nuovo Cim.* **19**, 154 (1961).

[58] J. Goldstone, A. Salam, and S. Weinberg, Broken Symmetries, *Phys.Rev.* **127**, 965 (1962).

[59] T. Nakano and K. Nishijima, Charge Independence for V-particles, *Prog.Theor.Phys.* **10**, 581 (1953).

[60] M. Gell-Mann, The interpretation of the new particles as displaced charge multiplets, *Nuovo Cim.* **4**, 848 (1956).

[61] N. Cabibbo, Unitary Symmetry and Leptonic Decays, *Phys.Rev.Lett.* **10**, 531 (1963).

[62] M. Kobayashi and T. Maskawa, CP Violation in the Renormalizable Theory of Weak Interaction, *Prog.Theor.Phys.* **49**, 652 (1973).

[63] Z. Maki, M. Nakagawa, and S. Sakata, Remarks on the unified model of elementary particles, *Prog.Theor.Phys.* **28**, 870 (1962).

[64] B. Pontecorvo, Neutrino Experiments and the Problem of Conservation of Leptonic Charge, *Sov.Phys.JETP* **26**, 984 (1968).

[65] S. Coleman and J. Mandula, All Possible Symmetries of the S-Matrix, *Phys.Rev.* **159**, 1251 (1967).

[66] Y. A. Golfand and E. P. Likhtman, Extension of the Algebra of Poincare Group Generators and Violation of p Invariance, *JETP Lett.* **13**, 323 (1971).

[67] P. Ramond, Dual Theory for Free Fermions, *Phys.Rev.* **D3**, 2415 (1971).

[68] A. Neveu and J. H. Schwarz, Factorizable dual model of pions, *Nucl.Phys.* **B31**, 86 (1971).

[69] J.-L. Gervais and B. Sakita, Field theory interpretation of supergauges in dual models, *Nucl.Phys.* **B34**, 632 (1971).

[70] J. Wess and B. Zumino, Supergauge Transformations in Four-Dimensions, *Nucl.Phys.* **B70**, 39 (1974).

[71] A. Salam and J. A. Strathdee, Supergauge Transformations, *Nucl.Phys.* **B76**, 477 (1974).

[72] R. Haag, J. T. Łopuszánski, and M. Sohnius, All Possible Generators of Supersymmetries of the s Matrix, *Nucl.Phys.* **B88**, 257 (1975).

[73] M. F. Sohnius, Introducing Supersymmetry, *Phys.Rept.* **128**, 39 (1985).

[74] A. Wiedemann and H. J. W. Müller-Kirsten, *Supersymmetry: An Introduction With Conceptual and Calculational Details*, 1. ed. (World Scientific, 1987).

[75] J. Wess and B. Zumino, Supergauge Invariant Extension of Quantum Electrodynamics, *Nucl.Phys.* **B78**, 1 (1974).

[76] M. Drees, R. Godbole, and P. Roy, *Theory and Phenomenology of Sparticles: An Account of four-dimensional N=1 Supersymmetry in High Energy Physics*, 1. ed. (World Scientific, New York, 2004).

[77] L. Girardello and M. T. Grisaru, Soft breaking of supersymmetry, *Nucl.Phys.* **B194**, 65 (1982).

[78] S. Ferrara, Supersymmetry And Fundamental Particle Interactions, *Phys.Rept.* **105**, 5 (1984).

[79] P. Nath, R. L. Arnowitt, and A. H. Chamseddine, APPLIED N=1 SUPERGRAVITY, **NUB-2613** (1983).

[80] S. P. Martin, A Supersymmetry primer, *Adv.Ser.Direct.High Energy Phys.* **21**, 1 (2010), hep-ph/9709356.

[81] F. Zwirner, Observable Delta B=2 Transitions Without Nucleon Decay in a Minimal Supersymmetric Extension of the Standard Model, *Phys.Lett.* **B132**, 103 (1983).

[82] S. Dawson, R-Parity Breaking in Supersymmetric Theories, *Nucl.Phys.* **B261**, 297 (1985).

[83] R. Barbieri and A. Masiero, Supersymmetric Models with Low-Energy Baryon Number Violation, *Nucl.Phys.* **B267**, 679 (1986).

[84] S. Dimopoulos and L. J. Hall, Lepton and Baryon Number Violating Collider Signatures from Supersymmetry, *Phys.Lett.* **B207**, 210 (1988).

[85] V. D. Barger, G. Giudice, and T. Han, Some New Aspects of Supersymmetry R-Parity Violating Interactions, *Phys.Rev.* **D40**, 2987 (1989).

[86] R. M. Godbole, P. Roy, and X. Tata, Tau signals of R-parity breaking at LEP-200, *Nucl.Phys.* **B401**, 67 (1993), hep-ph/9209251.

[87] G. Bhattacharyya and D. Choudhury, D and tau decays: Placing new bounds on R-parity violating supersymmetric coupling, *Mod.Phys.Lett.* **A10**, 1699 (1995), hep-ph/9503263.

[88] G. Bhattacharyya, R-parity violating supersymmetric Yukawa couplings: A Minireview, *Nucl.Phys.Proc.Suppl.* **52A**, 83 (1997), hep-ph/9608415.

[89] H. K. Dreiner, An Introduction to explicit R-parity violation, *Adv.Ser.Direct.High Energy Phys.* **21**, 565 (2010), hep-ph/9707435.

[90] B. Allanach, A. Dedes, and H. Dreiner, R parity violating minimal supergravity model, *Phys.Rev.* **D69**, 115002 (2004), hep-ph/0309196.

[91] M. Chemtob, Phenomenological constraints on broken R parity symmetry in supersymmetry models, *Prog.Part.Nucl.Phys.* **54**, 71 (2005), hep-ph/0406029.

[92] R. Barbier *et al.*, R-parity violating supersymmetry, *Phys.Rept.* **420**, 1 (2005), hep-ph/0406039.

[93] G. 't Hooft, Symmetry Breaking Through Bell-Jackiw Anomalies, *Phys.Rev.Lett.* **37**, 8 (1976).

[94] J. E. Kim and H. P. Nilles, The μ Problem and the Strong CP Problem, *Phys.Lett.* **B138**, 150 (1984).

[95] U. Ellwanger, C. Hugonie, and A. M. Teixeira, The Next-to-Minimal Supersymmetric Standard Model, *Phys.Rept.* **496**, 1 (2010), 0910.1785.

[96] M. Maniatis, The Next-to-Minimal Supersymmetric extension of the Standard Model reviewed, *Int.J.Mod.Phys.* **A25**, 3505 (2010), 0906.0777.

[97] ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, LEP Working Group for Higgs Boson Searches, S. Schael *et al.*, Search for neutral MSSM Higgs bosons at LEP, *Eur.Phys.J.* **C47**, 547 (2006), hep-ex/0602042.

[98] A. Djouadi *et al.*, Benchmark scenarios for the NMSSM, *JHEP* **0807**, 002 (2008), 0801.4321.

[99] O. Stål and G. Weiglein, Light NMSSM Higgs bosons in SUSY cascade decays at the LHC, *JHEP* **1201**, 071 (2012), 1108.0595.

[100] W. Beenakker, R. Hopker, M. Spira, and P. Zerwas, Squark and gluino production at hadron colliders, *Nucl.Phys.* **B492**, 51 (1997), hep-ph/9610490.

[101] A. Kulesza and L. Motyka, Threshold resummation for squark-antisquark and gluino-pair production at the LHC, *Phys.Rev.Lett.* **102**, 111802 (2009), 0807.2405.

[102] A. Kulesza and L. Motyka, Soft gluon resummation for the production of gluino-gluino and squark-antisquark pairs at the LHC, *Phys.Rev.* **D80**, 095004 (2009), 0905.4749.

[103] W. Beenakker *et al.*, Soft-gluon resummation for squark and gluino hadroproduction, *JHEP* **0912**, 041 (2009), 0909.4418.

[104] W. Beenakker *et al.*, Squark and Gluino Hadroproduction, *Int.J.Mod.Phys.* **A26**, 2637 (2011), 1105.1110.

[105] W. Beenakker, M. Kramer, T. Plehn, M. Spira, and P. Zerwas, Stop production at hadron colliders, *Nucl.Phys.* **B515**, 3 (1998), hep-ph/9710451.

[106] W. Beenakker *et al.*, Supersymmetric top and bottom squark production at hadron colliders, *JHEP* **1008**, 098 (2010), 1006.4771.

[107] ATLAS Collaboration, G. Aad *et al.*, Search for neutral Higgs bosons of the minimal supersymmetric standard model in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector, *JHEP* **1411**, 056 (2014), 1409.6064.

[108] CMS Collaboration, V. Khachatryan *et al.*, Search for neutral MSSM Higgs bosons decaying to a pair of tau leptons in pp collisions, *JHEP* **1410**, 160 (2014), 1408.3316.

[109] J. Alwall, P. Schuster, and N. Toro, Simplified Models for a First Characterization of New Physics at the LHC, *Phys.Rev.* **D79**, 075020 (2009), 0810.3921.

[110] LHC New Physics Working Group, D. Alves *et al.*, Simplified Models for LHC New Physics Searches, *J.Phys.* **G39**, 105005 (2012), 1105.2838.

[111] ATLAS Collaboration, G. Aad *et al.*, ATLAS Supersymmetry (SUSY) searches, `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/SupersymmetryPublicResults?rev=431`, October, 2014.

[112] CMS Collaboration, V. Khachatryan *et al.*, CMS Supersymmetry Physics Results, `https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS?rev=279`, October, 2014.

[113] S. F. King, S. Moretti, and R. Nevzorov, Theory and phenomenology of an exceptional supersymmetric standard model, *Phys.Rev.* **D73**, 035009 (2006), hep-ph/0510419.

[114] C. Sander, [private communication].

[115] U. Ellwanger, J. F. Gunion, and C. Hugonie, NMHDECAY: A Fortran code for the Higgs masses, couplings and decay widths in the NMSSM, *JHEP* **0502**, 066 (2005), hep-ph/0406215.

[116] U. Ellwanger and C. Hugonie, NMHDECAY 2.0: An Updated program for sparticle masses, Higgs masses, couplings and decay widths in the NMSSM, *Comput.Phys.Commun.* **175**, 290 (2006), hep-ph/0508022.

[117] G. Belanger, F. Boudjema, C. Hugonie, A. Pukhov, and A. Semenov, Relic density of dark matter in the NMSSM, *JCAP* **0509**, 001 (2005), hep-ph/0505142.

[118] U. Ellwanger and C. Hugonie, NMSPEC: A Fortran code for the sparticle and Higgs masses in the NMSSM with GUT scale boundary conditions, *Comput.Phys.Commun.* **177**, 399 (2007), hep-ph/0612134.

[119] U. Ellwanger, C.-C. Jean-Louis, and A. Teixeira, Phenomenology of the General NMSSM with Gauge Mediated Supersymmetry Breaking, *JHEP* **0805**, 044 (2008), 0803.2962.

[120] D. Das, U. Ellwanger, and A. M. Teixeira, NMSDECAY: A Fortran Code for Supersymmetric Particle Decays in the Next-to-Minimal Supersymmetric Standard Model, *Comput.Phys.Commun.* **183**, 774 (2012), 1106.5633.

[121] M. Mühlleitner, A. Djouadi, and Y. Mambrini, SDECAY: A Fortran code for the decays of the supersymmetric particles in the MSSM, *Comput.Phys.Commun.* **168**, 46 (2005), hep-ph/0311167.

[122] M. Matsumoto and T. Nishimura, Mersenne Twister: A 623-dimensionally Equidistributed Uniform Pseudo-random Number Generator, *ACM Trans. Model. Comput. Simul.* **8**, 3 (1998).

[123] R. Brun and F. Rademakers, ROOT: An object oriented data analysis framework, *Nucl.Instrum.Meth.* **A389**, 81 (1997).

[124] A. Buckley, PySLHA: a Pythonic interface to SUSY Les Houches Accord data, (2013), 1305.4194.

[125] P. Z. Skands *et al.*, SUSY Les Houches accord: Interfacing SUSY spectrum calculators, decay packages, and event generators, *JHEP* **0407**, 036 (2004), hep-ph/0311123.

[126] B. Allanach *et al.*, SUSY Les Houches Accord 2, *Comput.Phys.Commun.* **180**, 8 (2009), 0801.0045.

[127] C. Lefèvre, The CERN accelerator complex. Complexe des accélérateurs du CERN, December 2008.

[128] ATLAS Collaboration, G. Aad *et al.*, The ATLAS Experiment at the CERN Large Hadron Collider, *JINST* **3**, S08003 (2008).

[129] ALICE Collaboration, K. Aamodt *et al.*, The ALICE experiment at the CERN LHC, *JINST* **3**, S08002 (2008).

[130] LHCb Collaboration, J. Alves, A. Augusto *et al.*, The LHCb Detector at the LHC, *JINST* **3**, S08005 (2008).

[131] LHCf Collaboration, O. Adriani *et al.*, The LHCf detector at the CERN Large Hadron Collider, *JINST* **3**, S08006 (2008).

[132] TOTEM Collaboration, G. Anelli *et al.*, The TOTEM experiment at the CERN Large Hadron Collider, *JINST* **3**, S08007 (2008).

[133] CMS Collaboration, G. Bayatian *et al.*, CMS physics: Technical design report, **CERN-LHCC-2006-001, CMS-TDR-008-1** (2006).

[134] W. Badgett, V. Halyo, K. Maeshima, and D. Marlow, Putting Luminosity and Trigger Scaler Information into the Event Stream, **CMS-IN-2007-025** (2007), Unpublished. Restricted access at `https://cds.cern.ch/record/1364879`.

[135] CMS Collaboration, Luminosity 2012, `https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults?rev=101`, December, 2014.

[136] CMS Collaboration, S. Chatrchyan *et al.*, Description and performance of track and primary-vertex reconstruction with the CMS tracker, *JINST* **9**, P10009 (2014), 1405.6569.

[137] CMS Collaboration, S. Chatrchyan *et al.*, Alignment of the CMS tracker with LHC and cosmic ray data, *JINST* **9**, P06009 (2014), 1403.2286.

[138] S. Baccaro *et al.*, Radiation damage effect on avalanche photodiodes, *Nucl.Instrum.Meth.* **A426**, 206 (1999).

[139] Z. Antunovic *et al.*, Radiation hard avalanche photodiodes for the CMS detector, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **537**, 379 (2005), Proceedings of the 7th International Conference on Inorganic Scintillators and their Use in Scientific adn Industrial Applications.

[140] K. W. Bell *et al.*, Vacuum phototriodes for the CMS electromagnetic calorimeter endcap, *IEEE Trans.Nucl.Sci.* **51**, 2284 (2004).

[141] CMS, S. Chatrchyan *et al.*, Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in *pp* Collisions at $\sqrt{s} = 7$ TeV, *JINST* **8**, P09009 (2013), 1306.2016.

[142] P. Adzic *et al.*, Energy resolution of the barrel of the CMS electromagnetic calorimeter, *JINST* **2**, P04004 (2007).

[143] CMS Collaboration, S. Chatrchyan *et al.*, Missing transverse energy performance of the CMS detector, *JINST* **6**, P09001 (2011), 1106.5048.

[144] CMS Collaboration, CMS, the Compact Muon Solenoid. Muon technical design report, **CERN-LHCC-97-32** (1997).

[145] CMS Collaboration, Single Muon efficiencies in 2012 Data, **CMS-DP-2013-009, CERN-CMS-DP-2013-009** (2013).

[146] CMS Collaboration, Muon Identification performance: hadron mis-Id measurements and RPC Muon selections, **CMS-DP-2014-018** (2014).

[147] CMS Collaboration, S. Chatrchyan *et al.*, Performance of CMS muon reconstruction in *pp* collision events at $\sqrt{s} = 7$ TeV, *JINST* **7**, P10002 (2012), 1206.4071.

[148] CMS, S. Dasu *et al.*, CMS. The TriDAS project. Technical design report, vol. 1: The trigger systems, **CERN-LHCC-2000-038** (2000).

[149] L. Tuura, A. Meyer, I. Segoni, and G. Della Ricca, CMS data quality monitoring: Systems and experiences, *J.Phys.Conf.Ser.* **219**, 072020 (2010).

[150] S. Drell and T.-M. Yan, Massive Lepton Pair Production in Hadron-Hadron Collisions at High-Energies, *Phys.Rev.Lett.* **25**, 316 (1970).

[151] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, Parton Fragmentation and String Dynamics, *Phys.Rept.* **97**, 31 (1983).

[152] B. Webber, A QCD Model for Jet Fragmentation Including Soft Gluon Interference, *Nucl.Phys.* **B238**, 492 (1984).

[153] G. Marchesini and B. Webber, Monte Carlo Simulation of General Hard Processes with Coherent QCD Radiation, *Nucl.Phys.* **B310**, 461 (1988).

[154] T. Gleisberg *et al.*, Event generation with SHERPA 1.1, *JHEP* **0902**, 007 (2009), 0811.4622.

[155] T. Sjostrand, S. Mrenna, and P. Z. Skands, PYTHIA 6.4 Physics and Manual, *JHEP* **0605**, 026 (2006), hep-ph/0603175.

[156] R. Field, Early LHC Underlying Event Data - Findings and Surprises, (2010), 1010.3558, Invited talk at HCP2010, Toronto, August 23, 2010.

[157] CMS Collaboration, S. Chatrchyan *et al.*, Measurement of the Underlying Event Activity at the LHC with $\sqrt{s} = 7$ TeV and Comparison with $\sqrt{s} = 0.9$ TeV, *JHEP* **1109**, 109 (2011), 1107.0330.

[158] CMS Collaboration, S. Chatrchyan *et al.*, Jet and underlying event properties as a function of charged-particle multiplicity in proton–proton collisions at $\sqrt{s}$ = 7 TeV, *Eur.Phys.J.* **C73**, 2674 (2013), 1310.4554.

[159] F. Maltoni and T. Stelzer, MadEvent: Automatic event generation with MadGraph, *JHEP* **0302**, 027 (2003), hep-ph/0208156.

[160] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, MadGraph 5: Going Beyond, *JHEP* **1106**, 128 (2011), 1106.0522.

[161] J. M. Campbell and R. Ellis, MCFM for the Tevatron and the LHC, *Nucl.Phys.Proc.Suppl.* **205-206**, 10 (2010), 1007.3492.

[162] P. Nason, A New method for combining NLO QCD with shower Monte Carlo algorithms, *JHEP* **0411**, 040 (2004), hep-ph/0409146.

[163] S. Frixione, P. Nason, and C. Oleari, Matching NLO QCD computations with Parton Shower simulations: the POWHEG method, *JHEP* **0711**, 070 (2007), 0709.2092.

[164] S. Alioli, P. Nason, C. Oleari, and E. Re, NLO vector-boson production matched with shower in POWHEG, *JHEP* **0807**, 060 (2008), 0805.4802.

[165] Z. Was, TAUOLA for simulation of tau decay and production: perspectives for precision low energy and LHC applications, *Nucl.Phys.Proc.Suppl.* **218**, 249 (2011), 1101.1652.

[166] GEANT4, S. Agostinelli *et al.*, GEANT4: A Simulation toolkit, *Nucl.Instrum.Meth.* **A506**, 250 (2003).

[167] J. Allison *et al.*, Geant4 developments and applications, *IEEE Trans.Nucl.Sci.* **53**, 270 (2006).

[168] CMS Collaboration, Comparison of the Fast Simulation of CMS with the first LHC data, **CMS-DP-2010-039, CERN-CMS-DP-2010-039** (2010).

[169] D. Barney, CMS slice raw illustrator files., https://cms-docdb.cern.ch/cgi-bin/PublicDocDB//ShowDocument?docid=5581, November 2012.

[170] CMS Collaboration, Photon reconstruction and identification at sqrt(s) = 7 TeV, **CMS-PAS-EGM-10-005** (2010).

[171] R. Wigmans, Calorimetry: Energy measurement in particle physics, *Int.Ser.Monogr.Phys.* **107**, 1 (2000).

[172] CMS Collaboration, Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET, **CMS-PAS-PFT-09-001** (2009).

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

[174] CMS Collaboration, Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV, **CMS-PAS-PFT-10-002** (2010).

[175] CMS Collaboration, S. Chatrchyan *et al.*, Shape, Transverse Size, and Charged Hadron Multiplicity of Jets in pp Collisions at 7 TeV, *JHEP* **1206**, 160 (2012), 1204.3170.

[176] CMS Collaboration, Electromagnetic calorimeter calibration with 7 TeV data, **CMS-PAS-EGM-10-003** (2010).

[177] CMS Collaboration, S. Chatrchyan *et al.*, Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS, *JINST* **6**, P11002 (2011), 1107.4277.

[178] CMS Collaboration, Average Pileup in 2012, https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults?rev=101#Pileup_distribution, February, 2015.

[179] K. Rose, Deterministic Annealing for Clustering, Compression, Classification, Regression and related Optimisation Problems, *Proceedings of the IEEE* **86** (1998).

[180] CMS Collaboration, Baseline muon selections, https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonId#Tight_Muon?rev=54, October, 2014.

[181] CMS Collaboration, Cut Based Electron ID, Unpublished, restricted access at https://twiki.cern.ch/twiki/bin/viewauth/CMS/EgammaCutBasedIdentification#Electron_ID_Working_Points?rev=31, October, 2014.

[182] J. E. Huth *et al.*, Toward a standardization of jet definitions, **FERMILAB-CONF-90-249-E, FNAL-C-90-249-E** (1990).

[183] M. Cacciari, G. P. Salam, and G. Soyez, The Anti-k(t) jet clustering algorithm, *JHEP* **0804**, 063 (2008), 0802.1189.

[184] CMS, M. Schroeder, Performance of Jets in CMS, Geneva, 2014, CMS-CR-2014-065.

[185] CMS Collaboration, Pileup Removal Algorithms, **CMS-PAS-JME-14-001** (2014).

[186] CMS Collaboration, S. Chatrchyan *et al.*, Identification of b-quark jets with the CMS experiment, *JINST* **8**, P04013 (2013), 1211.4462.

[187] CMS Collaboration, Performance of b tagging at sqrt(s)=8 TeV in multijet, ttbar and boosted topology events, **CMS-PAS-BTV-13-001** (2013).

[188] CMS, V. Khachatryan *et al.*, Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8$ TeV, *JINST* **10**, P02006 (2015), 1411.0511.

[189] CMS, S. Chatrchyan *et al.*, Commissioning and Performance of the CMS Silicon Strip Tracker with Cosmic Ray Muons, *JINST* **5**, T03008 (2010), 0911.4996.

[190] CMS, S. Chatrchyan *et al.*, Commissioning and Performance of the CMS Pixel Tracker with Cosmic Ray Muons, *JINST* **5**, T03007 (2010), 0911.5434.

[191] J. Draeger, Track based alignment of the CMS silicon tracker and its implication on physics performance, **DESY-THESIS-2011-026**.

[192] P. Billoir, Progressive track recognition with a Kalman like fitting procedure, *Comput.Phys.Commun.* **57**, 390 (1989).

[193] P. Billoir and S. Qian, Simultaneous pattern recognition and track fitting by the Kalman filtering method, *Nucl.Instrum.Meth.* **A294**, 219 (1990).

[194] R. Mankel, A Concurrent track evolution algorithm for pattern recognition in the HERA-B main tracking system, *Nucl.Instrum.Meth.* **A395**, 169 (1997).

[195] R. Fruhwirth, Application of Kalman filtering to track and vertex fitting, *Nucl.Instrum.Meth.* **A262**, 444 (1987).

[196] CMS Collaboration, Iterative Tracking, `https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideIterativeTracking?rev=9`, March, 2015.

[197] P. Lenzi, C. Genta, and B. Mangano, Track reconstruction of real cosmic muon events with CMS tracker detector, *J.Phys.Conf.Ser.* **119**, 032030 (2008).

[198] V. Blobel and C. Kleinwort, A New method for the high precision alignment of track detectors, **DESY-02-077** (2002), hep-ex/0208021.

[199] V. Blobel, Software alignment for tracking detectors, *Nucl.Instrum.Meth.* **A566**, 5 (2006).

[200] V. Blobel, C. Kleinwort, and F. Meier, Fast alignment of a complex tracking detector using advanced track models, *Comput.Phys.Commun.* **182**, 1760 (2011), 1103.3909.

[201] C. Kleinwort, General Broken Lines as advanced track fitting method, *Nucl.Instrum.Meth.* **A673**, 107 (2012), 1201.4320.

[202] V. Blobel, A new fast track-fit algorithm based on broken lines, *Nucl.Instrum.Meth.* **A566**, 14 (2006).

[203] CMS, W. Adam *et al.*, Alignment of the CMS Silicon Strip Tracker during stand-alone Commissioning, *JINST* **4**, T07001 (2009), 0904.1220.

[204] D. Brown, A. Gritsan, Z. Guo, and D. Roberts, Local Alignment of the BABAR Silicon Vertex Tracking Detector, *Nucl.Instrum.Meth.* **A603**, 467 (2009), 0809.3823.

[205] H. Enderle, Momentum bias determination in the tracker alignment and first differential $t\bar{t}$ cross section measurement at CMS, **DESY-THESIS-2012-001, CERN-THESIS-2012-248**.

[206] B. Wittmer *et al.*, The laser alignment system for the CMS silicon microstrip tracker, *Nucl.Instrum.Meth.* **A581**, 351 (2007).

[207] CMS, S. Chatrchyan *et al.*, Alignment of the CMS Silicon Tracker during Commissioning with Cosmic Rays, *JINST* **5**, T03009 (2010), 0910.2505.

[208] J. Hauk, Measurement of associated Z^0-boson and b-jet production in proton-proton collisions with the CMS experiment, **DESY-THESIS-2012-030, CERN-THESIS-2012-182**.

[209] CMS Collaboration, Search for a light NMSSM Higgs boson produced in supersymmetric cascades and decaying into a b-quark pair, **CMS-PAS-HIG-14-030** (2015).

[210] CMS Collaboration, Search for a light NMSSM Higgs boson produced in supersymmetric cascades and decaying into a b-quark pair, `https://twiki.cern.ch/twiki/bin/view/CMSPublic/Hig14030TWiki?rev=2`, March, 2015.

[211] CMS, A. Apresyan, Identification and mitigation of anomalous signals in CMS hadronic calorimeter, *J.Phys.Conf.Ser.* **404**, 012044 (2012).

[212] CMS Collaboration, MET Optional Filters, Unpublished, restricted access at https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFilters?rev=63, November, 2014.

[213] CMS Collaboration, Utilities for Accessing Pileup Information for Data, Unpublished, restricted access at https://twiki.cern.ch/twiki/bin/view/CMS/PileupJSONFileforData?rev=16, November, 2014.

[214] CMS Collaboration, Methods to apply b-tagging efficiency scale factors: 1c) Event reweighting using scale factors only, Unpublished, restricted access at https://twiki.cern.ch/twiki/bin/viewauth/CMS/BTagSFMethods?rev=25#1c_Event_reweighting_using_scale, November, 2014.

[215] CMS Collaboration, 2012 FastSim Correction Factors, Unpublished, restricted access at https://twiki.cern.ch/twiki/bin/viewauth/CMS/Btag2012FastSimCorrectionFactors?rev=13, November, 2014.

[216] B. Safarzadeh Samani and S. Sekmen, b-tagging FullSim/FastSim correction factors, **CMS-AN-2013/376** (2014), Unpublished. Restricted access at http://cms.cern.ch/iCMS/jsp/openfile.jsp?tp=draft&files=AN2013_376_v2.pdf.

[217] CMS Collaboration, Top pair cross section in dileptons, **CMS-PAS-TOP-12-007** (2012).

[218] D. Tröndle, [private communication].

[219] M. J. Oreglia, *A Study of the Reactions* $\psi' \to \gamma\gamma\psi$, PhD thesis, Calif. Univ. Stanford, Stanford, CA, 1980, SLAC-0236, UMI-81-08973, SLAC-R-0236, SLAC-R-236.

[220] J. Gaiser, *Charmonium Spectroscopy From Radiative Decays of the* J/ψ *and* ψ', PhD thesis, Calif. Univ. Stanford, Stanford, CA, 1982, SLAC-0255, UMI-83-14449-MC, SLAC-R-0255, SLAC-R-255.

[221] T. Skwarnicki, DESY Report No. DESY-F31-86-02, 1986 (unpublished).

[222] Belle, H. Ikeda *et al.*, A detailed test of the CsI(Tl) calorimeter for BELLE with photon beams of energy between 20-MeV and 5.4-GeV, *Nucl.Instrum.Meth.* **A441**, 401 (2000).

[223] ATLAS Collaboration, CMS Collaboration, LHC Higgs Combination Group, 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** (2011).

[224] T. Junk, Confidence level computation for combining searches with small statistics, *Nucl.Instrum.Meth.* **A434**, 435 (1999), hep-ex/9902006.

[225] A. L. Read, Presentation of search results: The CL(s) technique, *J.Phys.* **G28**, 2693 (2002).

[226] L. Moneta *et al.*, The RooStats Project, *PoS* **ACAT2010**, 057 (2010), 1009.1003.

[227] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Asymptotic formulae for likelihood-based tests of new physics, *Eur.Phys.J.* **C71**, 1554 (2011), 1007.1727.

[228] CMS, V. Khachatryan *et al.*, Differential cross section measurements for the production of a W boson in association with jets in proton–proton collisions at $\sqrt{s} = 7$ TeV, *Phys.Lett.* **B741**, 12 (2015), 1406.7533.

[229] CMS, S. Chatrchyan *et al.*, Measurement of the production cross sections for a Z boson and one or more b jets in pp collisions at sqrt(s) = 7 TeV, *JHEP* **1406**, 120 (2014), 1402.1521.

[230] CMS Collaboration, PAT Tools, MET Systematics Tools, `https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuidePATTools?rev=54s#MET_Systematics_Tools`, November, 2013.

[231] CMS Collaboration, 2012 b-Tagging Scale Factors, Light Flavour, Unpublished, restricted access at `https://twiki.cern.ch/twiki/pub/CMS/BtagPOG/SFlightFuncs_EPS2013.C`, November, 2014.

[232] CMS Collaboration, 2012 b-Tagging Scale Factors, b/c Flavour, Unpublished, restricted access at `https://twiki.cern.ch/twiki/pub/CMS/BtagPOG/SFb-pt_WITHttbar_payload_EPS13.txt`, November, 2014.

[233] CMS Collaboration, 2012 FastSim Correction Factors, Central Light Flavour, Unpublished, restricted access at `https://twiki.cern.ch/twiki/pub/CMS/Btag2012FastSimCorrectionFactors/btagCF_Cen_l.txt`, November, 2014.

[234] CMS Collaboration, 2012 FastSim Correction Factors, Forward Light Flavour, Unpublished, restricted access at `https://twiki.cern.ch/twiki/pub/CMS/Btag2012FastSimCorrectionFactors/btagCF_For_l.txt`, November, 2014.

[235] CMS Collaboration, 2012 FastSim Correction Factors, c Flavour, Unpublished, restricted access at `https://twiki.cern.ch/twiki/pub/CMS/Btag2012FastSimCorrectionFactors/btagCF_c.txt`, November, 2014.

[236] CMS Collaboration, 2012 FastSim Correction Factors, b Flavour, Unpublished, restricted access at `https://twiki.cern.ch/twiki/pub/CMS/Btag2012FastSimCorrectionFactors/btagCF_b.txt`, November, 2014.

[237] P. M. Nadolsky *et al.*, Implications of CTEQ global analysis for collider observables, *Phys.Rev.* **D78**, 013004 (2008), 0802.0007.

[238] A. Martin, W. Stirling, R. Thorne, and G. Watt, Parton distributions for the LHC, *Eur.Phys.J.* **C63**, 189 (2009), 0901.0002.

[239] A. Martin, W. Stirling, R. Thorne, and G. Watt, Uncertainties on alpha(S) in global PDF analyses and implications for predicted hadronic cross sections, *Eur.Phys.J.* **C64**, 653 (2009), 0905.3531.

[240] A. Martin, W. Stirling, R. Thorne, and G. Watt, Heavy-quark mass dependence in global PDF analyses and 3- and 4-flavour parton distributions, *Eur.Phys.J.* **C70**, 51 (2010), 1007.2624.

[241] R. D. Ball *et al.*, A first unbiased global NLO determination of parton distributions and their uncertainties, *Nucl.Phys.* **B838**, 136 (2010), 1002.4407.

[242] M. Botje *et al.*, The PDF4LHC Working Group Interim Recommendations, (2011), 1101.0538.

[243] CMS Collaboration, CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update, **CMS-PAS-LUM-13-001** (2013).

[244] CMS Collaboration, Estimating Systematic Errors Due to Pileup Modeling, Unpublished, restricted access at `https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupSystematicErrors?rev=11`, November, 2014.

[245] ATLAS, G. Aad *et al.*, Search for Scalar Diphoton Resonances in the Mass Range $65 - 600$ GeV with the ATLAS Detector in pp Collision Data at $\sqrt{s} = 8\ TeV$, *Phys.Rev.Lett.* **113**, 171801 (2014), 1407.6583.

[246] CMS Collaboration, Search for a non-standard-model Higgs boson decaying to a pair of new light bosons in four-muon final states, **CMS-PAS-HIG-13-010** (2013).

[247] CMS, A. Dominguez *et al.*, CMS Technical Design Report for the Pixel Detector Upgrade, **CERN-LHCC-2012-016, CMS-TDR-011, FERMILAB-DESIGN-2012-02** (2012).

Acknowledgements

Now that this work has come to an end, I would like to thank a list of people that made this thesis possible and supported me during my PhD project.

First of all, I would like to thank Dr. Rainer Mankel for being my supervisor. He always supported me, especially his guidance through the CMS approval procedure was invaluable. Having him as my advisor helped me a lot, thanks to his great experience and because he asked the right questions.

I also want to express my gratitude to my referees Dr. Rainer Mankel, Prof. Peter Schleper, and Dr. Giacomo Polesello for agreeing to evaluate my thesis. In addition, I want to thank Dr. Rainer Mankel, Prof. Peter Schleper, Prof. Christian Sander, Prof. Georg Weiglein, and Prof. Jan Louis for taking part in my disputation and evaluating my work.

I am also grateful for the support from the DESY-CMS-Higgs group. In particular, Dr. Alexei Raspereza for offering me the topic and the opportunity for the PhD project, as well as advise for statistical methods. Not to mention Dr. Roberval Walsh and Dr. Matthias Schröder helped me a lot with practical aspects of my work and fruitful discussions.

Although, it was not the main part of this thesis, I also want to thank the DESY-CMS-Alignment group, which gave me the possibility to work on a subject with demands different from my physics analysis. In particular, I appreciate the work together with Dr. Gero Flucke including the nice lunch breaks.

I very much enjoyed the collaboration with the CMS-SUSY group of the University of Hamburg, most notably Prof. Christian Sander and Dr. Daniel Tröndle.

Dr. Oscar Stål and Prof. Georg Weiglein provided invaluable support regarding theoretical questions, and the correct usage of the theoretical tools and the interpretation of the output. Not to forget, the physics analysis of this thesis was stimulated by their theoretical work.

Although not directly related to this thesis, I want to appreciate my time in the group of Prof. Dominik Stöckinger, which equipped me with a solid theoretical knowledge that was very valuable even for my work in experimental particle physics.

One aspect, clearly related to this thesis, is the proof reading of the several chapters. I want to thank Peter Drechsel, Dr. Johannes Hauk, Dr. Matthias Schröder, and Dr. Roberval Walsh for their very helpful comments and suggestions.

At this point, I also would like to thank my office mates Jacob, Agni, and Luigi for the diversified atmosphere. Special thanks go to Dr. Matthias Schröder, my last office mate as a PhD student. I very much enjoyed the discussions about physics and everything, and, of course, my bachelor party, which he co-organised.

For nice lunch breaks with conversations about topics unrelated to my work, I would like to thank the ALPS group of DESY.

Finally, I would like to thank my friends and my family. My deepest gratitude goes to Katharina and Elena, not only for kindly sharing their last name with me, but even more for tolerating long working hours and a thin-skinned husband and father during the last months. And they reminded me that although physics is about matter, there are other things that matter, too. *Thank you!*

Declaration on Oath

I hereby declare, on oath, that I have written the present dissertation by my own and have not used other than the acknowledged resources and aids.

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Hamburg, 29 April 2015

List of Publications

The work during my PhD studies resulted in the following publication:

CMS Collaboration, Search for a light NMSSM Higgs boson produced in supersymmetric cascades and decaying into a b-quark pair, **CMS-PAS-HIG-14-030** (2015).